

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**

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

Primeiro, gostaria de agradecer minha família por sempre apoiar os meus estudos. Em especial, gostaria de agradecer minha prima Raquel Oliveira por me ajudar a revisar essa tese. Todos os erros continuam sendo meus. Também gostaria de agradecer meus amigos por terem tornado a jornada do Doutorado mais divertida.

Aos professores do PPGE/UFRGS, deixo um muito obrigado pelo conhecimento que me passaram em suas aulas. Em especial, gostaria de agradecer ao meu orientador João Caldeira. Finalmente, gostaria de agradecer ao CNPq pela bolsa de Doutorado concedida para eu realizar o curso de Doutorado na UFRGS.

## 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 outperforms the other methods. It presents similar Tracking Error volatility and higher mean return, which leads to a better risk-adjusted performance. 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 forward and backward iteration algorithms (simple methods that receive little attention in the literature) perform well in selecting assets for index tracking.

**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 crítica 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, conseqüentemente, 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 supera os outros métodos. Ele apresenta volatilidade do *Tracking Error* similar aos outros métodos com maior retorno médio, o que leva a um melhor desempenho ajustado pelo risco. 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 os algoritmos de iteração *forward* e *backward* (métodos simples que recebem pouca atenção na literatura) têm bom desempenho na seleção de ativos para o *index tracking*.

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

## LIST OF FIGURES

Figure 2.1 – Portfolio Utility for Different Risk-Aversions of the FF10 Data . . . . .	29
Figure 2.2 – Portfolio Utility for Different Risk-Aversions of the FF25 Data . . . . .	30
Figure 4.1 – Cumulative returns with 10 assets (IBOV) . . . . .	57
Figure 4.2 – Cumulative returns with 15 assets (IBOV) . . . . .	58
Figure 4.3 – Cumulative returns with 20 assets (IBOV) . . . . .	59
Figure 4.4 – Volatility of returns with 10 assets (IBOV) . . . . .	60
Figure 4.5 – Volatility of returns with 15 assets (IBOV) . . . . .	61
Figure 4.6 – Volatility of returns with 20 assets (IBOV) . . . . .	62
Figure 4.7 – Cumulative returns with 20 assets (SP500) . . . . .	63
Figure 4.8 – Cumulative returns with 30 assets (SP500) . . . . .	64
Figure 4.9 – Cumulative returns with 40 assets (SP500) . . . . .	65
Figure 4.10–Volatility of returns with 20 assets (SP500) . . . . .	66
Figure 4.11–Volatility of returns with 30 assets (SP500) . . . . .	67
Figure 4.12–Volatility of returns with 40 assets (SP500) . . . . .	68

## LIST OF TABLES

Table 2.1 – In Sample Results . . . . .	25
Table 2.2 – Out of Sample Results . . . . .	27
Table 2.3 – pvalues for Difference of Variances . . . . .	28
Table 2.4 – pvalues for Difference of SR . . . . .	28
Table 4.1 – Out of Sample Rebalancing for the Index Tracking Portfolios . . . . .	49
Table 4.2 – Out of Sample Rebalancing for the Index Tracking Portfolios (SP500) .	50
Table 4.3 – Results of Regressions (IBOV) . . . . .	51
Table 4.4 – Results of Regressions (SP500) . . . . .	52
Table 4.5 – Out of Sample Results for the Index Tracking Portfolios without Trans- action Costs (IBOV) . . . . .	53
Table 4.6 – Out of Sample Results for the Index Tracking Portfolios with Transaction Costs (IBOV) . . . . .	54
Table 4.7 – Out of Sample Results for the Index Tracking Portfolios without Trans- action Costs (SP500) . . . . .	55
Table 4.8 – Out of Sample Results for the Index Tracking Portfolios with Transaction Costs (SP500) . . . . .	56



## CONTENTS

<b>1</b>	<b>INTRODUCTION . . . . .</b>	<b>10</b>
<b>2</b>	<b>TANGENCY PORTFOLIO: A CRITIQUE . . . . .</b>	<b>12</b>
2.1	PORTFOLIO SELECTION . . . . .	13
2.1.1	The Optimal Unconstrained Portfolio . . . . .	15
2.1.2	Tangency Portfolio . . . . .	16
2.1.3	Minimum Variance Portfolio . . . . .	16
2.1.4	Optimal Constrained . . . . .	17
2.2	ANALYTICAL STUDY . . . . .	18
2.3	THE EXPERIMENT . . . . .	21
2.3.1	Data . . . . .	21
2.3.2	Evaluation Measures . . . . .	22
2.4	RESULTS . . . . .	23
2.5	CONCLUDING REMARKS . . . . .	25
<b>3</b>	<b>ON THE EQUIVALENCE OF MINIMUM TRACKING ERROR AND MINIMUM VARIANCE PORTFOLIOS . . . . .</b>	<b>31</b>
3.1	TWO OF A KIND: MINIMUM TRACKING ERROR AND MINIMUM VARIANCE PORTFOLIOS . . . . .	31
3.2	REGRESSION BASED PORTFOLIOS . . . . .	34
3.2.1	Memmel . . . . .	34
3.3	FINAL REMARKS . . . . .	36
<b>4</b>	<b>INDEX TRACKING WITH FEATURE SELECTION . . . . .</b>	<b>37</b>
4.1	BASIC DEFINITIONS . . . . .	39
4.2	FEATURE SELECTION ALGORITHMS . . . . .	41
4.2.1	Ordinary Least Squares (OLS) Regression . . . . .	41
4.2.2	Forward Stepwise Selection . . . . .	42
4.2.3	Backward Stepwise Selection . . . . .	42
4.2.4	Lasso . . . . .	42
4.3	THE EXPERIMENT . . . . .	43
4.3.1	Data . . . . .	43
4.3.2	Evaluation Measures . . . . .	45
4.4	RESULTS . . . . .	46
4.5	CONCLUDING REMARKS . . . . .	48

5	FINAL CONSIDERATIONS . . . . .	69
	REFERENCES . . . . .	71

## 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) and Markowitz (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) and Markowitz (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).<sup>1</sup> 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.

---

<sup>1</sup>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 paper 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. We also perform an empirical evaluation using the Kenneth French Data Library. 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.

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 Tangency Portfolio are aired by Kirby and Ostdiek (2012) when they point out that: i) if its denominator ( $A = \mathbf{e}'\Sigma^{-1}\mu$ ) is small, then the TP can display extreme weights and hence extreme returns; ii) 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; iii) 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  $\mu$ . Several papers on parameter uncertainty like Merton (1980) and Chopra and Ziemba (1993) indict  $\mu$  as the main culprit of the volatility of portfolio weights. They also point out that  $\mu$  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 ( $\mu$ ) and the covariance matrix of returns ( $\Sigma$ ), and find that the maximum utility is given when  $\mathbf{w} = \frac{\Sigma^{-1}\mu}{\gamma}$ . Later, they relax the assumption of knowing  $\mu$  and  $\Sigma$  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.

Our results show that the TP rarely outperforms other portfolios on in and out of sample evaluations. 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.

The rest of the paper is organized as follows. Section 2.1 presents the basics of portfolio selection and the portfolios that we will use throughout this paper. 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 shows an empirical evaluation of the TP against other portfolios; for this, we will use the Kenneth French data library. 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 paper. We start by considering an investment universe with  $N$  assets  $S_1, S_2, \dots, S_N$ . Then, we denote the vector of future and uncertain returns of those assets by  $R_t$ , where

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

Letting  $R_t$  be an  $N \times 1$  vector of risky-asset returns for period  $t$ , we will assume that the returns follow a multivariate normal distribution with mean vector  $\mu = E[R_t]$  and covariance matrix  $\Sigma = \text{Var}[R_t]$ .

The vector of expected returns  $\mu = E[R_t]$  has elements  $\mu_i = E[R_{it}]$ , for all  $i = 1, \dots, N$  and it is denoted by

$$\mu = [\mu_1, \mu_2, \dots, \mu_N]'$$

The covariance matrix of returns  $\Sigma = \text{Var}[R_t]$  has elements  $\sigma_{ii} = \sigma_i^2$  and  $\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j$  (for  $i \neq j$ ), where  $\sigma_i$  is the standard deviation of  $R_{it}$  and  $\rho_{ij}$  is the correlation coefficient of assets  $S_i$  and  $S_j$  (for  $i \neq j$ ). The covariance matrix  $\Sigma$  is symmetric and is denoted by

$$\Sigma = (\sigma_{ij})_{i,j=1,\dots,N} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1N} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \dots & \sigma_{NN} \end{bmatrix}.$$

Throughout this paper, we will assume that  $\Sigma$  is positive definite so it can be invertible.

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

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

and  $w_i$  is the fraction of total wealth invested in asset  $S_i$ . The vector  $\mathbf{w}$  can also be called the allocation vector. The (uncertain) portfolio return  $R_p$  is a weighted average of the asset returns selected to the portfolio with weights  $w_i$  for  $i = 1, \dots, N$ .

$$R_p = w_1 R_1 + \dots + w_N R_N = \sum_{i=1}^N w_i R_{ti} = \mathbf{w}' R_t. \quad (2.1)$$

Hence, the portfolio's expected return  $\mu_p$  is a weighted average of individual expected returns and the portfolio variance  $\sigma_p^2$  is a quadratic function of the allocation vector, as shown in:

$$\mu_p = E[R_p] = E[\mathbf{w}' R_t] = \mathbf{w}' \mu \quad (2.2)$$

$$\sigma_p^2 = \text{Var}[R_p] = \text{Var}[\mathbf{w}' R_t] = \mathbf{w}' \Sigma \mathbf{w} \quad (2.3)$$

$$\sigma_p = \sqrt{\mathbf{w}' \Sigma \mathbf{w}}. \quad (2.4)$$

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

$$U(\mathbf{w}) = \mathbf{w}' \mu - \frac{\gamma}{2} \mathbf{w}' \Sigma \mathbf{w}. \quad (2.5)$$

### 2.1.1 The Optimal Unconstrained Portfolio

Here and in the next subsections, we will show how to find optimized portfolios. The first one here is the Optimal Unconstrained portfolio. It assumes that there is a risk-free asset, so we ignore the fully-invested restriction and solve the following problem:

$$\min_{\mathbf{w}} \quad \frac{\mathbf{w}'\Sigma\mathbf{w}}{2} \quad \text{s.t.} \quad \mathbf{w}'\mu = \mu_0. \quad (2.6)$$

That problem has the following Lagrangian:

$$L = \frac{\mathbf{w}'\Sigma\mathbf{w}}{2} - \lambda(\mathbf{w}'\mu - \mu_0). \quad (2.7)$$

Solving that Lagrangian is equivalent to maximizing a quadratic utility:

$$U(\mathbf{w}) = \mathbf{w}'\mu - \frac{\gamma}{2}\mathbf{w}'\Sigma\mathbf{w}, \quad (2.8)$$

where  $\gamma$  is the coefficient of risk-aversion. In the problem (2.6) the trade-off between risk and return is given by the Lagrange Multiplier  $\lambda$  in equation (2.7), and the risk-aversion is implicit and subject to alterations given the level of return wanted by the investor.

The answer to the problem (2.8) is given when the agent's utility is maximized (first derivative set to zero) and we find the following solution:

$$\mathbf{w}_{\text{ou}} = \frac{\Sigma^{-1}\mu}{\gamma}. \quad (2.9)$$

Now, the solution to the problem (2.6) is given by Lagrange's method, where we find:

$$\mathbf{w}_{\text{ou}} = \frac{\mu_0}{\mu'\Sigma^{-1}\mu}\Sigma^{-1}\mu. \quad (2.10)$$

We will call this solution  $\mathbf{w}_{\text{ou}}$  or OU for Optimal Unconstrained. From (2.9) and (2.10), we can infer that:

$$\lambda = \frac{1}{\gamma} = \frac{\mu_0}{\mu'\Sigma^{-1}\mu}, \quad (2.11)$$

which means that the choice of  $\mu_0$  affects the  $\gamma$ , i.e. the risk-aversion coefficient. **So the choice of  $\mu_0$  controls the aggressiveness with which the problem seeks return instead of variance.** The

Kirby and Ostdiek (2012) set  $\mu_0$  equal to the return of the equal weighted portfolio ( $\mu_{\text{ew}} = \mu'\mathbf{e}/N$ ) to let the aggressiveness of the mean-variance portfolio comparable to it. In this paper however, we use 1.1 times the mean return of the minimum variance portfolio ( $\mu_{\text{mv}}$ ) as our target return. This happens because, in our data, the mean return of the equally weighted portfolio is smaller than the mean return of the minimum variance portfolio, consequently, it is not efficient to use  $\mu_{\text{ew}}$  as a target return.



### *Sum of Positions*

From (2.9), we can notice that the sum of the weights, or the portfolio's total position, equals to:

$$x_{\text{TP},t} = \frac{\mathbf{e}'\Sigma^{-1}\mu}{\gamma},$$

where  $\mathbf{e}$  is a conformable vector of ones. Observe that  $x_{\text{TP},t}$  does not necessarily sum to one (fully invested position). That happens because we did not constrain our portfolio to do that, so it is only complete with the risk-free asset, which has an implicit weight of:

$$w_{f,t} = 1 - \mathbf{w}'\mathbf{e} = 1 - x_{\text{TP},t}$$

where  $w_{f,t}$  is the weight put on the risk-free asset.

#### **2.1.2 Tangency Portfolio**

If we were to invest only on the risky part of the OU, we would have to normalize the allocation vector, turning it into what is called the Tangency Portfolio (TP):

$$\mathbf{w}_{\text{TP},t} = \frac{\Sigma^{-1}\mu}{\mathbf{e}'\Sigma^{-1}\mu}. \quad (2.12)$$

Where  $\mathbf{e}$  is a conformable vector of ones. This portfolio of risky assets has the property of maximizing the Sharpe Ratio ( $\mu_p/\sigma_p$ ). We can also find it by solving:

$$\max_{\mathbf{w}} \frac{\mu'\mathbf{w} - \mathbf{e}r_f}{\sqrt{\mathbf{w}'\Sigma\mathbf{w}}} \quad \text{s.t.} \quad \mathbf{e}'\mathbf{w} = 1. \quad (2.13)$$

The statement of the problem above makes it easy for us to see that the TP is, by definition, the portfolio with the maximum Sharpe Ratio.

#### **2.1.3 Minimum Variance Portfolio**

If the investor is only interested in minimizing variance, what might be the case when his risk aversion is very high ( $\gamma \rightarrow +\infty$ ), he solves the following problem:

$$\min_{\mathbf{w}} \frac{\mathbf{w}'\Sigma\mathbf{w}}{2} \quad \text{s.t.} \quad \mathbf{e}'\mathbf{w} = 1, \quad (2.14)$$

with the following Lagrangian:

$$L = \frac{\mathbf{w}'\Sigma\mathbf{w}}{2} - \lambda(\mathbf{e}'\mathbf{w} - 1) \quad (2.15)$$

The answer to the problem is:

$$\mathbf{w}_{mv} = \frac{\Sigma^{-1}\mathbf{e}}{\mathbf{e}'\Sigma^{-1}\mathbf{e}}, \quad (2.16)$$

where  $\mathbf{e}$  is a conformable vector of ones. We call this portfolio MV.

#### 2.1.4 Optimal Constrained

In the original problem of Markowitz (1952), the investor considers only risky assets, yielding the problem:

$$\min_{\mathbf{w}} \quad \frac{\mathbf{w}'\Sigma\mathbf{w}}{2} \quad \text{s.t.} \quad \mathbf{w}'\mu = \mu_0, \quad \mathbf{w}'\mathbf{e} = 1, \quad (2.17)$$

with Lagrangean:

$$L = \frac{\mathbf{w}'\Sigma\mathbf{w}}{2} - \lambda_1(\mathbf{w}'\mu - \mu_0) - \lambda_2(\mathbf{w}'\mathbf{e} - 1). \quad (2.18)$$

This is equivalent to maximizing a quadratic utility constrained to the fully-invested condition:

$$U(\mathbf{w}) = \mathbf{w}'\mu - \frac{\gamma}{2}\mathbf{w}'\Sigma\mathbf{w} \quad \text{s.t.} \quad \mathbf{w}'\mathbf{e} = 1, \quad (2.19)$$

which has the following Lagrangian:

$$L = \mathbf{w}'\mu - \frac{\gamma}{2}\mathbf{w}'\Sigma\mathbf{w} - \lambda(\mathbf{w}'\mathbf{e} - 1). \quad (2.20)$$

Merton (1972) solves problem (2.17) and gives the answer:

$$\mathbf{w}_{oc} = \frac{C\mu_0 - A}{D}\Sigma^{-1}\mu + \frac{B - A\mu_0}{D}\Sigma^{-1}\mathbf{e}, \quad (2.21)$$

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

$$A = \mathbf{e}'\Sigma^{-1}\mu, \quad B = \mu'\Sigma^{-1}\mu, \quad C = \mathbf{e}'\Sigma^{-1}\mathbf{e}, \quad D = BC - A^2. \quad (2.22)$$

(MERTON, 1972, p. 1853) points out the fact that  $B > 0$ ,  $C > 0$  and  $D > 0$ , if  $\Sigma$  is Positive Definite. We call the portfolio in equation (2.21) OC for Optimal Constrained. Another solution to problem (2.19) is

$$\mathbf{w}_{oc} = \frac{\Sigma^{-1}\mu}{\gamma} + \left(1 - \frac{\mathbf{e}'\Sigma^{-1}\mu}{\gamma}\right) \frac{\Sigma^{-1}\mathbf{e}}{\mathbf{e}'\Sigma^{-1}\mathbf{e}}. \quad (2.23)$$

The relation between (2.23) and (2.21) appears in Bodnar, Parolya and Schmid (2013). They show for what values of  $\gamma$  those values are the same, and the answer is:

$$\gamma = \begin{cases} \frac{s}{\mu_0 - R_{\text{gmv}}}, & \text{for } \mu_0 > R_{\text{gmv}} \\ \infty, & \text{for } \mu_0 = R_{\text{gmv}}, \end{cases} \quad (2.24)$$

where,

$$R_{\text{gmv}} = \frac{A}{C}, \quad s = \mu'Q\mu, \quad Q = \Sigma^{-1} - \frac{\Sigma^{-1}\mathbf{e}\mathbf{e}'\Sigma^{-1}}{C}. \quad (2.25)$$

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

$$\gamma = \frac{D}{C\mu_0 - A}, \quad \text{for } \mu_0 > \frac{A}{C}, \quad (2.26)$$

$$\mu_0 = \frac{A}{C} + \frac{D}{\gamma C}. \quad (2.27)$$

Notice that the equation above forms an expression of  $\gamma$  in terms of  $\mu_0$ . This way, we can find the implicit  $\gamma$  for any value of  $\mu_0$ . With this expression of  $\gamma$ , we can find (2.21) in terms of the TP and the MV:

$$\begin{aligned} \mathbf{w}_{\text{oc}} &= \frac{C\mu_0 - A}{D}\Sigma^{-1}\mu + \frac{B - A\mu_0}{D}\Sigma^{-1}\mathbf{e}, \\ &= \frac{1}{\gamma}\Sigma^{-1}\mu + \frac{1}{C}\left(1 - \frac{A}{\gamma}\right)\Sigma^{-1}\mathbf{e}, \\ &= \mathbf{w}_{\text{ou}} + (1 - x_{\text{tp}})\mathbf{w}_{\text{mv}}, \\ &= \mathbf{w}_{\text{tp}}x_{\text{tp}} + \mathbf{w}_{\text{mv}} - \mathbf{w}_{\text{mv}}x_{\text{tp}}, \\ &= \mathbf{w}_{\text{mv}} + (\mathbf{w}_{\text{tp}} - \mathbf{w}_{\text{mv}})x_{\text{tp}}. \end{aligned}$$

## 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 (2.5), we may find the utilities of the portfolios listed above. This results in:

$$U_{\text{ew}} = \frac{\mathbf{e}'\mu}{N} - \frac{\gamma}{2N^2}\mathbf{e}'\Sigma\mathbf{e}, \quad (2.28)$$

$$U_{\text{ou}} = \frac{B}{2\gamma}, \quad (2.29)$$

$$U_{\text{tp}} = \frac{B}{2A^2}(2A - \gamma), \quad (2.30)$$

$$U_{\text{mv}} = \frac{1}{2C}(2A - \gamma), \quad (2.31)$$

$$U_{\text{oc}} = \frac{-\gamma^2 + 2A\gamma + D}{2\gamma C}, \quad (2.32)$$

where  $\mathbf{e}$  is a conformable vector of ones, and  $A$ ,  $B$ ,  $C$ , and  $D$  are Merton's constants, which are equal to:

$$A = \mathbf{e}'\Sigma^{-1}\mu, \quad B = \mu'\Sigma^{-1}\mu, \quad C = \mathbf{e}'\Sigma^{-1}\mathbf{e}, \quad D = BC - A^2. \quad (2.33)$$

## TP and OU

Comparing (2.30) with (2.29), we find:

$$\begin{aligned} U_{\text{tp}} &\geq U_{\text{ou}} \\ \frac{B}{2A^2}(2A - \gamma) &\geq \frac{B}{2\gamma} \\ \gamma(2A - \gamma) &\geq A^2 \\ -\gamma^2 + 2A\gamma - A^2 &\geq 0 \\ -(\gamma - A)^2 &\geq 0 \end{aligned}$$

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  $U_{\text{ou}} = U_{\text{tp}}$ .

## TP and OC

Comparing (2.30) with (2.32), we find:

$$\begin{aligned}
U_{tp} &\geq U_{oc} \\
\frac{B}{2A^2}(2A - \gamma) &\geq \frac{-\gamma^2 + 2A\gamma + D}{2\gamma C} \\
BC\gamma(2A - \gamma) &\geq -\gamma^2 A^2 + 2A^3\gamma + A^2 D \\
-\gamma^2(BC - A^2) + \gamma A(2BC - 2A^2) - A^2 D &\geq 0 \\
-D\gamma^2 + 2AD\gamma - A^2 D &\geq 0
\end{aligned}$$

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

$$\begin{aligned}
-(\gamma^2 - 2A\gamma + A^2) &\geq 0 \\
-(\gamma - A)^2 &\geq 0
\end{aligned}$$

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 (2.30) with (2.31), we find:

$$\begin{aligned}
U_{mv} &\geq U_{tp} \\
\frac{1}{2C}(2A - \gamma) &\geq \frac{B}{2A^2}(2A - \gamma) \\
A^2(2A - \gamma) &\geq BC(2A - \gamma) \\
(A^2 - BC)(2A - \gamma) &\geq 0 \\
-D(2A - \gamma) &\geq 0
\end{aligned}$$

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

$$\begin{aligned}
-2A + \gamma &\geq 0 \\
\gamma &\geq 2A
\end{aligned}$$

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

**Proposition 3:** The ex-ante utilities  $U_{mv}$  and  $U_{tp}$  depend on the values of  $\gamma$  with  $U_{mv} > U_{tp}$  for  $\gamma > 2A$  and  $U_{mv} < U_{tp}$  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  $\gamma$  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  $\gamma$  can be seen in the Figures 2.1 and 2.2 below. In the figures, we can see that when  $\gamma = A$ , the OC, the OU, and the TP have the same utilities. For values of  $\gamma$  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  $\gamma$  is lesser than  $2A$ , but the MV has higher utility when  $\gamma$  is greater than  $2A$ .

## 2.3 THE EXPERIMENT

In this section, we will do an experiment to test the performance of the Tangency Portfolio against alternative portfolios. This experiment 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 two datasets of monthly raw returns (not in excess of the risk-free asset) drawn from the Kenneth French library.<sup>1</sup> One dataset is the 10 Industry portfolios (FF10), and the other dataset is 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 ( $T = 660$ ).

To estimate the parameters inputs used in the portfolio policies of Section 2.1, we adopt a recursive estimation approach based on a *rolling window time series of returns*  $\{R_t\}_{t=i}^{J-i+1}$  with  $i$  varying from  $i = 1$  to  $i = T - J - 1$ . For this experiment, we use window length equal to  $J = 120$ .

With the rolling windows' scheme, we form the vector of expected means and the covariance matrices with the formulas:

$$\hat{\mu} = E[R_t] = J^{-1} \sum_{t=1}^J R_t \quad (2.34)$$

$$\hat{\Sigma} = \text{Var}[R_t] = (J - 1)^{-1} \sum_{t=1}^J (R_t - \hat{\mu})(R_t - \hat{\mu})'. \quad (2.35)$$

---

<sup>1</sup><[http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html)>

This results in  $T - J - 1 = 479$  estimates of covariance matrices and expected returns. Using them, we find the allocation vector using the policies of Section 2.1. By the end of this process, we generate  $T - J - 1 = 479$  allocation vectors for each policy.

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

$$R_{p,t+1} = \mathbf{w}_{p,t}' R_{t+1}. \quad (2.36)$$

The time series of the *out of sample portfolio returns*  $\{R_{p,t+1}\}_{t=J}^{T-1}$  is the subject of our analysis.

### 2.3.2 Evaluation Measures

As we have said above, what we will evaluate are the out of sample returns of the portfolios. 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.

#### 2.3.2.1 Sharpe Ratio

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:

$$\hat{S}R_p = \frac{\hat{\mu}_p}{\hat{\sigma}_p}. \quad (2.37)$$

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

$$\hat{\mu}_p = (T - J - 1)^{-1} \sum_{t=J+1}^T R_{pt}, \quad (2.38)$$

$$\hat{\sigma}_p = \sqrt{(T - J - 2)^{-1} \sum_{t=J+1}^T (R_{pt} - \hat{\mu}_p)^2}, \quad (2.39)$$

where  $R_{pt} = \mathbf{w}_{p,t-1}' R_t$  is the portfolio return.

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.

#### 2.3.2.2 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.

$$C\hat{E}Q_p = \hat{U}_p = \hat{\mu}_p - \frac{\gamma}{2}\hat{\sigma}_p^2. \quad (2.40)$$

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  $w_t$  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  $w_{i,t} \frac{1 + R_{i,t+1}}{1 + \mathbf{w}_t' R_{t+1}}$ , meaning the turnover rate in time  $t$  after rebalancing the portfolio can be defined as:

$$\text{TO}_t = \sum_{i=1}^N \left| w_{i,t+1} - w_{i,t} \frac{1 + R_{i,t+1}}{1 + R_{p,t}} \right|. \quad (2.41)$$

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:

$$R_{p,t}^{\text{net}} = (1 + R_{p,t})(1 - c \times \text{TO}), \quad (2.42)$$

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

## 2.4 RESULTS

In this subsection we present the results of our empirical analysis. 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) and the FF25 (Panel B) datasets. The statistics shown are: the annualized mean of the portfolio return ( $\mu$ ), the annualized standard deviation of the portfolio return



( $\sigma$ ), the annualized Sharpe Ratio of the portfolio returns (SR), the quadratic utility of the portfolio returns for  $\gamma = 1$  and  $\gamma = 5$  (CEQ 1 and CEQ 5), and the in sample implied risk-aversion of the portfolios ( $\gamma$ ).

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) and FF25 (Panel B) datasets. The statistics shown are the same as in Table 2.1 with the exclusion of  $\gamma$ , 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.

## FF10

The in sample results for the FF10 are presented in the Panel A of Table 2.1. We can observe that the TP and the OC are tied for the highest SR with 1.04. However, for both risk-aversions ( $\gamma = 1$  and  $\gamma = 5$ ), the OC has higher CEQ with 1.05 and 0.77, while the TP has 1.03 and 0.76.

On the out of sample results, presented in the Panel A of Table 2.2, we can observe that the TP comes at the very bottom of the comparisons, with and without transaction costs (TC). The TP's Sharpe Ratio is degraded when compared to the in sample results (from 1.04 to 0.33), what also happens in much less extent to the OC (from 1.04 to 0.99). On the other hand, for the MV and the EW, the out of sample results are barely distinguishable from the in sample ones. That happens because the objective function of the MV is to minimize volatility, and there's actually no difference from in sample and the out of sample's allocation vector of the EW. Also, the MV wins the out of sample comparisons with the best values of SR and CEQ.

After transaction costs (TC), the TP is dealt another blow to its performance and it can hardly be recognizable as an optimal portfolio (SR equal to  $-0.03$  and negative CEQs). The same cannot be said about the OC and the MV. Even though they present positive turnover (TO), they are robust to transaction costs and come on top of the EW in terms of SR and CEQ with  $\gamma = 5$ .

## FF25

Now for the FF25 dataset, the in sample results are presented in the Panel B of Table 2.1. Here, all risk-adjusted performance measures have the following order: TP, OC, MV, EW. The SR for those portfolios are 1.85, 1.45, 1.33, and 0.83.

This is different for the out of sample results. Before TC, the only performance measure that the TP wins is with the CEQ 1 with value of 1.39, tied with the OC. At all the other risk-adjusted measures the TP comes at the very bottom. The OC, on the

other hand, has all the better measures with an SR of 1.32 and CEQ 5 of 1.10. As for the lack of degradation of the EW and the MV, it is due to the same reasons presented in the FF10 results.

Once we consider transaction costs, the TP's performance goes from bad to worse and it trails further behind the other portfolios presenting negative values for all the risk-adjusted measures. The OC portfolio loses only to the EW in terms of the CEQ 1, coming on top for the other measures.

## 2.5 CONCLUDING REMARKS

In our empirical experiment, 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 5	Gamma	Gross	Short
<i>Panel A: FF10 Dataset</i>								
TP	13.12	12.57	1.04	102.79	76.47	8.31	1.98	49.23
MV	12.24	12.14	1.01	95.86	71.31	<i>Inf</i>	1.67	33.32
OC	13.46	12.95	1.04	105.21	77.27	6.01	2.15	57.27
EW	12.29	14.65	0.84	93.47	57.69	147.48	1.00	0.00
<i>Panel B: FF25 Dataset</i>								
TP	31.55	16.75	1.88	251.22	204.44	11.24	10.80	489.77
MV	15.71	11.82	1.33	125.06	101.77	<i>Inf</i>	6.15	257.63
OC	17.28	11.88	1.45	138.09	114.57	113.39	6.28	263.88
EW	14.45	17.51	0.83	107.62	56.52	-141.46	1.00	0.00
<i>Panel C: IBOV Dataset</i>								
TP	52.92	19.84	2.67	20.22	17.09	13.44	3.38	119.11
MV	27.38	14.27	1.92	10.46	8.84	<i>Inf</i>	1.96	47.89
OC	30.12	14.35	2.10	11.54	9.91	125.34	1.98	49.15
EW	25.88	23.91	1.08	9.14	4.60	-229.44	1.00	0.00

**Source:** The author.

**Note:** The Table shows in sample results of the Tangency Portfolio (TP), Minimum Variance Portfolio (MV), Optimal Constrained Portfolio (OC), and the Equally Weighted Portfolio (EW) for the FF10 (Panel A) and FF25 (Panel B) datasets. The sample period is from July 1963 to June 2018 ( $T = 660$ ). The first  $J = 120$  observations are held out to initialize the rolling estimators.  $\mu$  is the annualized mean of the portfolio return,  $\sigma$  is the annualized standard deviation of the portfolio return, SR is the annualized Sharpe Ratio of the portfolio returns, and CEQ is the quadratic utility of the portfolio returns for  $\gamma = 1$  and  $\gamma = 5$ .  $\gamma$  is the in sample implied risk-aversion of the portfolios.

Ledoit and Wolf (2008) Ledoit and Wolf (2011)

Table 2.2: Out of Sample Results

	Mean	SD	SR	CEQ 1	CEQ 5	TO	Short	Gross	SR	CEQ 1	CEQ 5
<i>Panel A: FF10 Dataset</i>											
TP	8.64	25.92	0.33	43.98	-67.96	146.56	145.86	3.92	-0.03	-98.39	-447.10
MV	12.49	12.35	1.01	97.76	72.33	16.16	67.00	2.34	0.93	89.60	64.18
OC	12.35	12.54	0.99	96.39	70.19	18.63	67.81	2.36	0.89	86.99	60.77
EW	12.29	14.65	0.84	93.47	57.69	2.46	0.00	1.00	0.83	92.23	56.47
<i>Panel B: FF25 Dataset</i>											
TP	24.52	39.63	0.62	138.86	-122.89	1276.13	952.41	20.05	-0.16	<i>NA</i>	<i>NA</i>
MV	16.79	13.23	1.27	132.63	103.45	79.32	323.83	7.48	0.90	92.44	63.21
OC	17.57	13.28	1.32	139.11	109.71	80.77	332.50	7.65	0.95	98.15	68.71
EW	14.45	17.51	0.83	107.62	56.52	1.89	0.00	1.00	0.82	106.67	55.61
<i>Panel C: IBOV Dataset</i>											
TP	-797.15	2983.93	-0.27	<i>NA</i>	<i>NA</i>	52.45	1165.84	24.32	-0.29	<i>NA</i>	<i>NA</i>
MV	25.90	15.88	1.63	9.78	7.78	4.34	66.20	2.32	1.28	7.60	5.59
OC	26.06	15.89	1.64	9.84	7.84	4.35	66.18	2.32	1.29	7.66	5.64
EW	24.14	23.83	1.01	8.45	3.95	0.41	0.00	1.00	0.99	8.25	3.75

**Source:** The author.

**Note:** The Table shows out of sample results of the Tangency Portfolio (TP), Minimum Variance Portfolio (MV), Optimal Constrained Portfolio (OC), and the Equally Weighted Portfolio (EW) for the FF10 (Panel A) and FF25 (Panel B) datasets. The sample period is from July 1963 to June 2018 ( $T = 660$ ). The first  $J = 120$  observations are held out to initialize the rolling estimators.  $\mu$  is the annualized mean of the portfolio return,  $\sigma$  is the annualized standard deviation of the portfolio return, SR is the annualized Sharpe Ratio of the portfolio returns, and CEQ is the quadratic utility of the portfolio returns for  $\gamma = 1$  and  $\gamma = 5$ . TO is the average turnover of the portfolio. The statistics to the right of the TO column are for the returns net of transaction costs with 50bp.

Table 2.3: pvalues for Difference of Variances

	Var Diff	$b = 02$	$b = 04$	$b = 06$	$b = 08$	$b = 10$
<i>Panel A: FF10 Dataset</i>						
TP	1.48	12.90	17.20	22.10	24.90	27.70
OC	0.03	3.60	4.30	4.30	4.60	4.60
EW	0.34	0.00	0.00	0.00	0.00	0.00
<i>Panel B: FF25 Dataset</i>						
TP	2.19	2.50	2.90	4.40	7.20	9.20
OC	0.01	36.50	36.80	37.40	37.10	38.20
EW	0.56	0.00	0.00	0.00	0.00	0.00
<i>Panel C: IBOV Dataset</i>						
TP	10.47	31.20	23.50	19.20	15.80	14.30
OC	0.00	54.10	53.70	53.20	52.00	53.00
EW	0.81	0.00	0.00	0.00	0.00	0.00

**Source:** The author.

**Note:** The Table shows the log difference of variances and  $p$ -values for tests of differences. In all datasets the benchmark was the GMVP. The remaining columns are values from bootstrap with  $M = 5000$ , where  $M$  is **XXXXX** and  $b$  is the average block length.

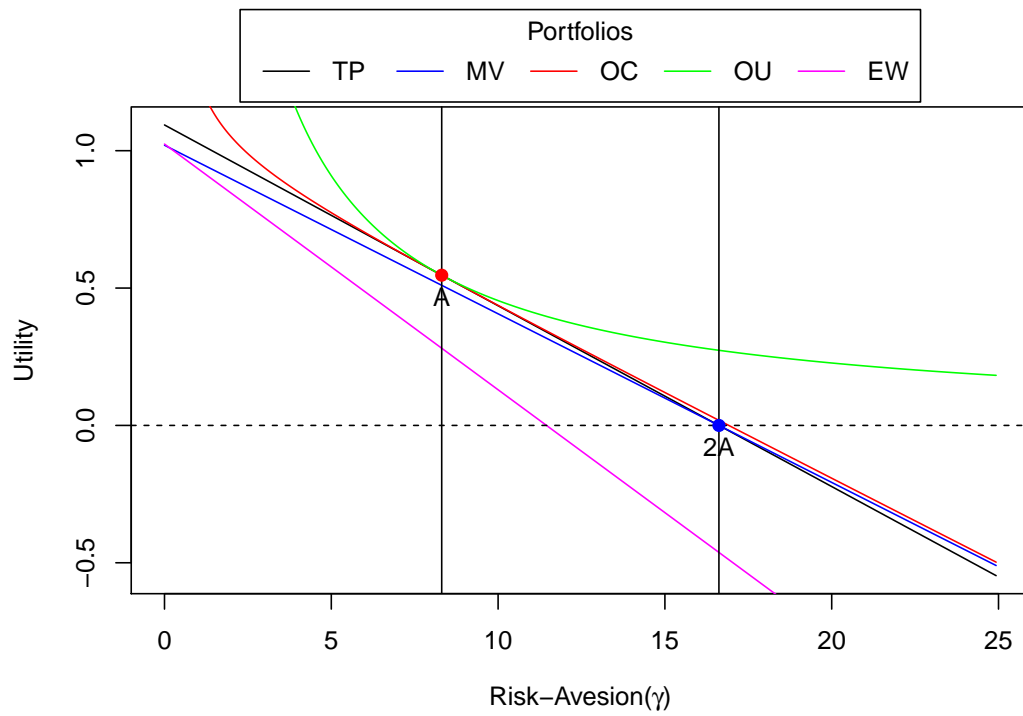
Table 2.4: pvalues for Difference of SR

	SR Diff	$b = 02$	$b = 04$	$b = 06$	$b = 08$	$b = 10$
<i>Panel A: FF10 Dataset</i>						
TP	-0.68	0.74	1.34	1.84	2.88	4.14
OC	-0.03	26.85	28.35	27.59	28.37	27.75
EW	-0.17	16.54	18.20	18.40	18.94	18.34
<i>Panel B: FF25 Dataset</i>						
TP	-0.65	3.54	3.50	4.08	4.58	5.56
OC	0.06	0.04	0.02	0.04	0.04	0.04
EW	-0.44	0.32	0.14	0.16	0.38	0.40
<i>Panel C: IBOV Dataset</i>						
TP	-1.90	1.64	1.50	1.52	1.66	1.84
OC	0.02	44.35	45.67	44.49	45.49	44.39
EW	-0.62	4.92	5.42	5.08	5.08	5.30

**Source:** The author.

**Note:** The Table shows the difference of annualized SR and  $p$ -values for tests of differences. In all datasets the benchmark was the GMVP. The remaining columns are values from bootstrap with  $M = 5000$ , where  $M$  is **XXXXX** and  $b$  is the average block length.

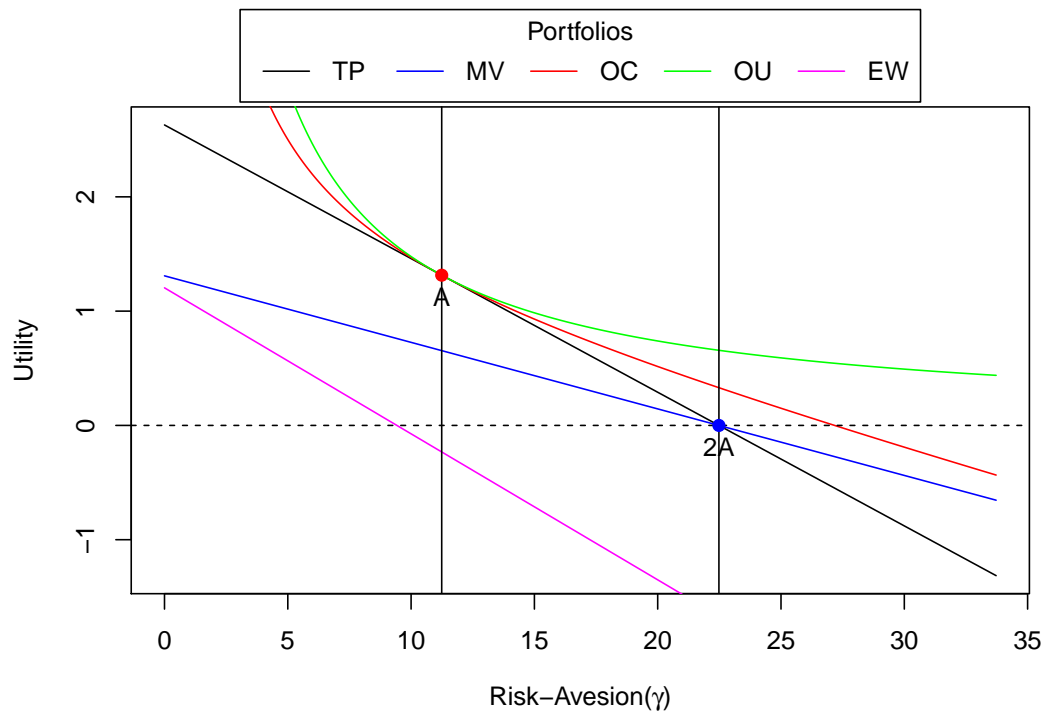
Figure 2.1: Portfolio Utility for Different Risk-Aversions of the FF10 Data



**Source:** The author.

**Note:** The figure shows the utility level rankings for different Risk-Aversion for each portfolio. TP is the Tangency Portfolio, MV is the Minimum-Variance Portfolio, OC is the Optimal Constrained Portfolio, OU is the Optimal Unconstrained Portfolio and EW is the Equally Weighted Portfolio. The OC uses the implied  $\gamma$  of having a target return of  $\mu_0 = 1.1 \times \mu_{mv}$ .  $A = e' \Sigma^{-1} \mu$  is the Merton constant shown in equation (2.33).

Figure 2.2: Portfolio Utility for Different Risk-Aversions of the FF25 Data



**Source:** The author.

**Note:** The figure shows the utility level rankings for different Risk-Aversion for each portfolio. TP is the Tangency Portfolio, MV is the Minimum-Variance Portfolio, OC is the Optimal Constrained Portfolio, OU is the Optimal Unconstrained Portfolio and EW is the Equally Weighted Portfolio. The OC uses the implied  $\gamma$  of having a target return of  $\mu_0 = 1.1 \times \mu_{mv}$ .  $A = e' \Sigma^{-1} \mu$  is the Merton constant shown in equation (2.33).

### 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 paper 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: MINIMUM 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:



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

whereas the portfolio variance is:

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

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:

$$\begin{aligned}
\text{Min}_w \quad & w'\Sigma_{XX}w + \sigma_y^2 w'ee'w - 2w'\sigma_{Xy}e'w \\
\text{s.t.} \quad & e'w = 1.
\end{aligned} \tag{3.1}$$

And has FOC:

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

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.:

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

Therefore, the MT problem can be stated as:

$$\begin{aligned}
\text{Min}_{\substack{w}} \quad & w' \Sigma_{XX} w + \sigma_y^2 - 2w' \sigma_{Xy} \\
\text{s.t.} \quad & e' w = 1.
\end{aligned} \tag{3.2}$$

And has FOC:

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

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:

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

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:

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

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:

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

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

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  $\alpha$ , slope  $\beta$  and constrain  $e' \beta = 1$ :

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

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:

$$\begin{aligned} \beta^* &= \Sigma_{XX}^{-1} \sigma_{Xy} + (1 - e' \Sigma_{XX}^{-1} \sigma_{Xy}) \frac{\Sigma_{XX}^{-1} e}{e' \Sigma_{XX}^{-1} e} \\ &= w_{ou} + (1 - e' w_{ou}) w_{mv}, \end{aligned}$$

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 Memmel

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:

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

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  $\Omega$  as the covariance matrix of regressors of equation (3.7):

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

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

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

Now, we have to show that the weights  $w_{MV}^{ex}$  of the GMVP correspond to the regression coefficient  $\beta^{ex}$ . 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 weight vector  $w_{MV}^{ex} = (w_{P,1}, \dots, w_{P,K})'$  and the stock returns  $r_t^{ex}$  and  $r_{t,N}$ :

$$\begin{aligned} r_{tP} &= w_P^{ex'} r_t^{ex} + (1 - w_P^{ex'} e) r_{tN} \\ &= r_{tN} + w_P^{ex'} (r_t^{ex} - er_{tN}) \end{aligned}$$

The return variance of this arbitrary portfolio  $P$

$$\sigma_P^2 = \sigma_N^2 + w_P^{ex'} \Omega w_P^{ex} - 2w_P^{ex'} \text{cov}(r_{tN}e - r_t^{ex}, r_{tN}) \quad (3.9)$$

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

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

and they find:

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

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 paper, 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

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 main goal of this paper is to 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).

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 paper 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.

Our results show that, in the brazilian case, the lasso selection method outperforms the other methods by presenting similar Tracking Error volatility and higher mean returns. Note that this overperformance 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 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).

However, in the American case, the lasso method presents more volatility, but this is more than offset by its superior mean returns, which ultimately delivers better risk-adjusted performance. But, because our primary goal is to minimize Tracking Error volatility, this kind of risk-return tradeoff is not desirable in our case. Even though the backward and forward selection methods might be outperformed by the lasso, there are some instances where these methods are superior or comparable to the benchmark index. 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.

Beyond this introduction, this paper is organized as follows. Section 4.1 presents the basics of portfolio selection that will be used throughout the paper. 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 paper. 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

$$R_{t,i} = \frac{P_{t,i}}{P_{t-1,i}} - 1. \quad (4.1)$$

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

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

We will assume that the returns have mean vector  $E[R_t] = \mu_X$  and covariance matrix  $V[R_t] = \Sigma$ . 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:

$$R_t^e = R_t - ey_t.$$

where  $e$  is a  $N \times 1$  vector of ones. We assume that  $R_t$  has mean vector and covariance matrix

$$\begin{aligned} E[R_t^e] &= \mu_X - e\mu_y = \mu \\ V[R_t^e] &= \Sigma + \sigma_y^2 ee' - 2Cov(X_t, y_t)e' = \Omega. \end{aligned}$$



## Portfolio 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

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

and  $w_i$  is the fraction of total wealth invested in asset  $i$ . The vector  $w$  can also be called the allocation vector. If we constrain the weights in  $w$  to sum up to one ( $e'w = 1$ ), we have the **fully bought condition**.

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

$$R_{p,t+1} = w_t' R_{t+1}. \quad (4.2)$$

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

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

## 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  $w_{t,i}(1 + R_{t+1,i})$ , where  $w_{t,i}$  is the  $i$ -th element of  $w_t$  and the  $i$ -th element of  $R_t$  is denoted as  $R_{t,i}$ . Also, in general, each dollar invested in the whole portfolio has changed from  $w_t$  to  $w_t(1 + R_{p,t+1})$ . Therefore, prior to rebalancing, the weight on asset  $i$  has changed from  $w_{i,t}$  to

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

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

$$w_t^{+1} = w_t \odot \frac{e + R_{t+1}}{1 + R_{p,t+1}}, \quad (4.3)$$

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

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

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

$$w_t^{+h} = w_t \odot \frac{(e + R_{t+1}) \odot \dots \odot (e + R_{t+h})}{(1 + R_{p,t+1}) \times \dots \times (1 + R_{p,t+h})}. \quad (4.4)$$

## Tracking Error

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

$$TE_{t+1} = w_t' R_{t+1} - y_{t+1} = R_{p,t+1} - y_{t+1}, \quad (4.5)$$

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 ( $e'w = 1$ ), then the excess return of a portfolio on a benchmark is the same as the Tracking Error:

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

Hence, if the weights sum up to one ( $e'w = 1$ ), 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  $\alpha$  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:

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

where  $\varepsilon_t$  is the residual of the regression.

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

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

where  $X$  is the  $J \times N$  matrix of asset returns:  $X = [X_{t-J}, X_{t-J+1}, \dots, X_{t-1}]'$  and  $y$  is the  $J \times 1$  vector of index returns  $y = [y_{t-J}, y_{t-J+1}, \dots, y_{t-1}]'$ .

#### 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 (LUMLEY, 2017).

#### 4.2.4 Lasso

To fit the least squares regression, we estimate the values of  $\alpha, \beta_1, \dots, \beta_p$  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:

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

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  $\lambda$ , the bigger the amount of regularization and tighter the variable selection will be. Thus, we have to fit  $\lambda$  in a convenient way to suit our objectives. In this paper we use the `glmnet` package (FRIEDMAN; HASTIE; TIBSHIRANI, 2010) of the R software (R Core Team, 2019). 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 THE EXPERIMENT

In this section, we present the methodology of our experiment. The goal of such experiment 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 ( $R_t$ ), we use the algorithms from Section 4.2 to select which assets will enter in our portfolios (asset

selection phase).

## Allocation

After that, 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 constraint ( $e'w = 1$ ). The answer to the problem is:

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

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

## Rolling Windows

To estimate the  $\hat{\Omega}_t$  we adopt a recursive estimation approach based on a *rolling window time series of returns*  $\{R_j\}_{j=t}^{t+J-1}$  with  $t$  varying from  $t = 1$  to  $t = T - J - 1$ . For this experiment, we use window length equal to  $J = 250$  for the Brazilian dataset and  $J = 500$  for the American dataset, because the American dataset has more assets, and we need  $J > N$  for the covariance matrix to be invertible. Another way to see that is by stacking  $J$  vectors of excess returns on top of each other and forming the  $J \times N$  matrix

$$R = \begin{bmatrix} R_{t-J} \\ R_{t-J+1} \\ \dots \\ R_{t-1} \end{bmatrix}$$

With the rolling windows' scheme, we form the vector of expected means and the covariance matrices of excess returns with the formulas:

$$\hat{\mu}_t = J^{-1} \sum_{j=t}^{t+J-1} R_t, \quad (4.11)$$

$$\hat{\Omega}_t = (J-1)^{-1} \sum_{j=t}^{t+J-1} (R_t - \hat{\mu})(R_t - \hat{\mu})'. \quad (4.12)$$

This results in  $T - J - 1$  estimates of expected return vectors and covariance matrices. Using them, we find the allocation vector using equation (4.10). 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  $\{R_{p,t+1}\}_{t=J}^{T-1}$ . This time series is the subject of our analysis.

#### 4.3.2 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.

The statistics that we want to evaluate are the Sharpe Ratio (SR), which is the mean of portfolio returns divided by the standard deviation of the portfolio returns; and the Information Ratio (IR), which is the mean of the Tracking Error divided by the standard deviation of the Tracking Error.

##### *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

$$TO_t = |w_{t+1} - w_t^{+1}|, \quad (4.13)$$

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:

$$R_{p,t}^{net} = (1 + R_{p,t})(1 - cTO) - 1, \quad (4.14)$$

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 short selling in the portfolios, we report their **Short Interest**. It expresses the average size of the short positions (weights smaller than zero) in the portfolio, which has the following formula:

$$SI_t = \frac{\sum_{i=1}^N |w_{i,t}| - 1}{2} \quad (4.15)$$

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

## 4.4 RESULTS

In this section we present the results of our empirical analysis. Each row of the Tables 4.5 through 4.8 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 number of active positions allowed in each portfolio.

Tables 4.5 and 4.7 show statistics without considering transaction costs. The presented statistics are: annualized Mean and Standard Deviation (SD) of the benchmark index and of the portfolio returns. The index is shown in absolute terms and the portfolios are shown in relation to the index (Mean and SD of the portfolios are divided by the Mean and the SD of the index). SR denotes the annualized Sharpe Ratio. The Tables also show the Mean and SD of the Tracking Error of the portfolios (portfolio return minus the index return). And the IR denotes the annualized Information Ratio which is the Mean of the Tracking Error divided by its SD. The last three columns of the Table show the alfa, beta, and  $R^2$  (intercept in annualized terms, the slope, and the coefficient of determination) of a univariate regression of the out of sample portfolio returns against the benchmark Index in the out of sample period.

The Figures 4.1 through 4.3 and 4.7 through 4.9, show cumulative returns of each portfolio with a specified rebalancing period and a number of assets. Those figures also

show in pointed lines the difference between those cumulative returns and the index. The Figures 4.4 through 4.6 and 4.10 through 4.12, show rolling window (250 days) annualized standard deviation. Those figures also show in pointed lines the Tracking Error standard deviation.

For the brazilian dataset, the figures show that the portfolios attain similar return and Tracking Error volatility. As for the cumulative returns, they show that the lasso presents the higher returns. This translates into a bigger Sharpe Ratio (SR) for the lasso method. This higher mean is a double-edged knife. On one hand, higher returns is good to any investor. On the other hand, the higher returns detaches the portfolio from the index, and might increase the portfolio volatility. However, because we are only minimizing Tracking Error volatility, it seems that the lasso outperforms the other methods. In terms of the Information Ratio (IR), again, the lasso method outperforms the other methods, in all rebalancing frequencies, and for the same reasons as the ones reported for the SR.

For the American dataset, the figures show that the lasso portfolios attain higher Tracking Error volatility than the other methods. And the forward and backward iteration show very similar Tracking Error volatility. As for the cumulative returns, it seems that the lasso presents the higher ones. In the risk-adjusted performance, the lasso presents the higher Sharpe Ratio (SR), but this comes in exchange for higher volatility. Once again, we remind the reader that we are only minimizing the Tracking Error volatility, so the lasso, in the American case, does not attain the goal we expected, even though it presents a higher Sharpe Ratio. In terms of the Information Ratio (IR), the story is the same as in the SR, and the lasso attains higher values of IR by exchanging more mean returns for more volatility, an undesirable transfer in this kind of portfolio.

Tables 4.6 and 4.8 show statistics of the out of sample portfolio returns net of transaction costs. We used transaction costs equal to  $50bp$  as in DeMiguel, Garlappi and Uppal (2009). The statistics shown are annualized Mean and Standard Deviation (SD). SR is the annualized Sharpe Ratio of those returns. TO is the turnover and is presented as a daily percentage average. Short interest is the size of the short position and is expressed in percentage. Active is the average number of active positions in the portfolio. The last three columns of the Table show the alfa, beta, and  $R^2$  (intercept in annualized terms, the slope, and the coefficient of determination) of a univariate regression of the out of sample portfolio returns net of transactions costs against the benchmark Index in the out of sample period.

For both the brazilian and the American dataset, the the lasso method does not always select the maximum number of assets. This might be the reason why the lasso presents lower turnover, which ultimately leads to lower transaction costs. Next, we inspect the average short positions of the portfolios. Here we can see that while the short positions increase with the active positions, for the lasso method this increase is much less pronounced than in the other methods. This, again, leads to lower turnover and



transaction costs. Another observation is that the lasso penalty diminishes the portfolio turnover. But for the penalty to have this effect, we should use the lasso weights, which we don't, as we separate the asset selection and the asset allocation phase. What might be happening here is that even without the penalty, the lasso still selects assets that will present less turnover.

Examining the risk-adjusted performance of the portfolios, we attest that the lasso, once again, has the better Sharpe Ratio (SR). And once again this is due more to the higher mean return, than to the lower volatility. We already expounded why the higher mean return has two sides to it. Other noteworthy fact is that once we consider transaction costs, the bwd and the fwd methods might present lower SR than the index itself. One last remark to clarify those results is that the portfolios only differ on its asset selection methods, the asset allocation phase is the same for all of them, so what we are examining is how well each method selects the assets that will enter our portfolio.

#### 4.5 CONCLUDING REMARKS

Here, 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).

In the results section we can observe that, in terms of risk-adjusted performance, the lasso method outperforms the backward and forward selection. In the brazilian case, the lasso presents similar return and Tracking Error volatility, but with higher mean return. This translates into higher risk-adjusted performance for the lasso. This evidence is robust to rebalancing frequency and transaction costs. In the American case, the lasso also presents higher mean return, but presents higher volatility, which ultimately leads to a higher Sharpe Ratio. However, because our primary goal is to minimize Tracking Error volatility, that exchange is undesirable to us.

Even though the backward and forward selections are outperformed by the lasso in some cases, there are some instances in which they can track the benchmark index pretty well. 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).

Ledoit and Wolf (2008) Ledoit and Wolf (2011)

Table 4.1: Out of Sample Rebalancing for the Index Tracking Portfolios

	Mean	SD	TO	Short	Active	Gross	Mean	SD
<i>Panel A: Monthly Rebalancing</i>								
bwd.10	-0.23	9.14	2.13	0.04	10.00	1.00	-2.92	9.19
fwd.10	0.77	9.16	2.18	0.12	10.00	1.00	-1.99	9.22
las.10	2.55	9.51	1.40	0.37	9.99	1.01	0.78	9.55
bwd.15	1.58	8.40	1.99	1.25	15.00	1.02	-0.93	8.45
fwd.15	1.37	8.36	2.08	0.83	15.00	1.02	-1.25	8.42
las.15	2.37	8.48	1.41	0.27	14.93	1.01	0.59	8.51
bwd.20	1.38	8.14	1.96	4.02	20.00	1.08	-1.10	8.19
fwd.20	1.62	8.06	2.00	3.30	20.00	1.07	-0.90	8.13
las.20	1.91	8.00	1.44	0.57	19.95	1.01	0.09	8.03
<i>Panel B: Bimonthly Rebalancing</i>								
bwd.10	0.53	9.07	1.46	0.00	10.00	1.00	-1.31	9.11
fwd.10	1.21	9.20	1.41	0.24	10.00	1.00	-0.57	9.25
las.10	3.05	9.62	1.01	0.48	9.99	1.01	1.77	9.66
bwd.15	2.19	8.38	1.36	1.36	15.00	1.03	0.47	8.42
fwd.15	1.79	8.39	1.41	1.04	15.00	1.02	0.00	8.42
las.15	3.65	8.55	0.98	0.30	14.97	1.01	2.41	8.57
bwd.20	1.58	8.15	1.35	4.29	20.00	1.09	-0.12	8.20
fwd.20	2.14	8.18	1.34	3.56	20.00	1.07	0.44	8.22
las.20	2.62	7.97	1.01	0.49	19.93	1.01	1.34	8.00
<i>Panel C: Quarterly Rebalancing</i>								
bwd.10	0.53	9.10	1.17	0.00	10.00	1.00	-0.95	9.14
fwd.10	-0.36	9.23	1.11	0.00	10.00	1.00	-1.77	9.26
las.10	2.74	9.52	0.81	0.40	10.00	1.01	1.71	9.54
bwd.15	1.49	8.44	1.11	1.10	15.00	1.02	0.08	8.47
fwd.15	1.36	8.40	1.06	0.84	15.00	1.02	0.01	8.44
las.15	2.39	8.65	0.79	0.22	15.00	1.00	1.38	8.67
bwd.20	1.29	8.19	1.11	4.00	20.00	1.08	-0.11	8.23
fwd.20	1.49	8.19	1.05	3.29	20.00	1.07	0.15	8.22
las.20	1.76	8.06	0.77	0.54	19.96	1.01	0.78	8.08

Table 4.2: Out of Sample Rebalancing for the Index Tracking Portfolios (SP500)

	Mean	SD	TO	Short	Active	Gross	Mean	SD
<i>Panel A: Monthly Rebalancing</i>								
bwd.10	2.01	2.75	6.42	0.00	20.00	1.00	−6.07	3.54
fwd.10	0.66	2.90	5.39	0.01	20.00	1.00	−6.12	3.50
las.10	3.39	4.10	1.82	2.93	19.93	1.06	1.10	4.16
bwd.15	0.80	2.22	6.07	0.00	30.00	1.00	−6.84	3.06
fwd.15	1.16	2.39	5.43	0.04	30.00	1.00	−5.68	3.08
las.15	2.61	3.39	1.89	4.51	29.88	1.09	0.23	3.46
bwd.20	1.36	1.93	5.66	0.00	40.00	1.00	−5.77	2.76
fwd.20	1.55	2.08	5.48	0.06	40.00	1.00	−5.35	2.86
las.20	1.52	2.93	2.07	4.46	39.88	1.09	−1.09	3.01
<i>Panel B: Bimonthly Rebalancing</i>								
bwd.10	0.91	2.86	3.43	0.00	20.00	1.00	−3.41	3.34
fwd.10	1.62	2.79	2.82	0.02	20.00	1.00	−1.94	3.13
las.10	3.21	4.06	1.18	2.95	19.92	1.06	1.73	4.11
bwd.15	−0.16	2.31	3.33	0.00	30.00	1.00	−4.35	2.86
fwd.15	2.48	2.35	2.84	0.03	30.00	1.00	−1.10	2.76
las.15	2.62	3.39	1.32	4.34	29.92	1.09	0.95	3.45
bwd.20	0.47	1.96	3.09	0.00	40.00	1.00	−3.43	2.48
fwd.20	2.35	2.03	2.81	0.04	40.00	1.00	−1.19	2.49
las.20	2.02	2.99	1.45	4.67	39.92	1.09	0.19	3.06
<i>Panel C: Quarterly Rebalancing</i>								
bwd.10	3.35	2.73	2.49	0.00	20.00	1.00	0.22	3.14
fwd.10	2.06	2.85	2.06	0.00	20.00	1.00	−0.53	3.08
las.10	2.84	4.18	1.05	2.87	19.96	1.06	1.51	4.20
bwd.15	2.99	2.22	2.30	0.00	30.00	1.00	0.09	2.65
fwd.15	2.65	2.33	2.08	0.01	30.00	1.00	0.02	2.63
las.15	2.45	3.44	1.10	4.52	29.92	1.09	1.06	3.49
bwd.20	3.09	1.92	2.11	0.00	40.00	1.00	0.43	2.33
fwd.20	2.34	1.93	2.04	0.02	40.00	1.00	−0.24	2.27
las.20	2.06	2.93	1.17	4.22	39.88	1.08	0.59	2.99

Table 4.3: Results of Regressions (IBOV)

	Alpha	pval	Beta	pval	R2	IR	Alpha	pval	Beta	pval	R2	IR
<i>Panel A: Monthly Rebalancing</i>												
bwd.10	0.80	8.18	1.00	0.00	95.17	4.59	-2.44	0.00	1.01	0.00	92.35	-10.94
fwd.10	0.19	69.27	1.01	0.00	94.82	1.04	-2.52	0.00	1.02	0.00	92.70	-11.45
las.10	1.34	4.92	1.00	0.00	89.91	5.18	0.43	53.55	1.00	0.00	89.65	1.63
bwd.15	0.33	37.62	1.00	0.00	96.80	2.33	-2.74	0.00	1.00	0.00	94.15	-14.19
fwd.15	0.42	29.40	1.01	0.00	96.38	2.76	-2.32	0.00	1.01	0.00	94.17	-11.95
las.15	1.06	5.96	1.00	0.00	92.79	4.96	0.11	84.27	1.00	0.00	92.54	0.52
bwd.20	0.56	8.17	1.00	0.00	97.54	4.59	-2.29	0.00	1.00	0.00	95.15	-13.18
fwd.20	0.58	9.24	1.01	0.00	97.23	4.43	-2.17	0.00	1.01	0.00	94.93	-12.09
las.20	0.62	20.52	1.00	0.00	94.53	3.34	-0.42	39.66	1.00	0.00	94.27	-2.23
<i>Panel B: Bimonthly Rebalancing</i>												
bwd.10	0.43	36.69	0.99	0.00	94.71	2.38	-1.29	1.97	0.99	0.00	92.90	-6.15
fwd.10	0.61	18.45	1.01	0.00	95.10	3.50	-0.80	12.20	1.01	0.00	93.94	-4.07
las.10	1.24	6.54	1.01	0.00	90.16	4.85	0.65	33.79	1.01	0.00	89.98	2.52
bwd.15	-0.01	98.47	0.99	0.00	96.49	-0.05	-1.68	0.04	0.99	0.00	94.74	-9.34
fwd.15	0.97	1.32	1.00	0.00	96.45	6.53	-0.45	32.09	1.00	0.00	95.19	-2.61
las.15	1.07	5.73	0.99	0.00	92.79	5.01	0.41	47.91	0.99	0.00	92.53	1.86
bwd.20	0.25	45.02	0.99	0.00	97.44	1.99	-1.31	0.15	0.99	0.00	95.97	-8.36
fwd.20	0.93	0.58	1.00	0.00	97.32	7.28	-0.48	24.93	1.00	0.00	96.01	-3.03
las.20	0.86	8.37	0.99	0.00	94.24	4.56	0.13	79.64	0.99	0.00	93.99	0.68
<i>Panel C: Quarterly Rebalancing</i>												
bwd.10	1.38	0.23	0.99	0.00	95.16	8.04	0.14	78.42	0.99	0.00	93.67	0.72
fwd.10	0.74	11.69	1.02	0.00	95.01	4.13	-0.28	58.20	1.01	0.00	94.19	-1.45
las.10	1.13	10.48	1.00	0.00	89.51	4.27	0.60	38.85	1.00	0.00	89.41	2.27
bwd.15	1.27	0.05	0.98	0.00	96.72	9.15	0.13	76.90	0.98	0.00	95.37	0.77
fwd.15	1.01	0.87	1.01	0.00	96.55	6.92	-0.02	95.78	1.01	0.00	95.63	-0.14
las.15	0.99	8.37	1.00	0.00	92.62	4.56	0.44	45.04	1.00	0.00	92.41	1.99
bwd.20	1.30	0.00	0.99	0.00	97.54	10.76	0.25	52.04	0.99	0.00	96.40	1.69
fwd.20	0.92	0.42	1.00	0.00	97.58	7.56	-0.10	79.45	1.00	0.00	96.68	-0.69
las.20	0.83	8.72	1.00	0.00	94.52	4.51	0.25	61.82	1.00	0.00	94.31	1.31

Table 4.4: Results of Regressions (SP500)

	Alpha	pval	Beta	pval	R2	IR	Alpha	pval	Beta	pval	R2	IR
<i>Panel A: Monthly Rebalancing</i>												
bwd.10	0.10	93.00	0.97	0.00	91.93	0.17	-0.97	38.30	0.97	0.00	91.84	-1.67
fwd.10	0.47	67.20	0.98	0.00	91.93	0.81	-0.62	57.58	0.98	0.00	91.82	-1.07
las.10	0.96	40.30	1.01	0.00	91.76	1.61	0.26	81.96	1.01	0.00	91.70	0.44
bwd.15	0.99	32.12	0.95	0.00	92.95	1.90	0.00	99.93	0.95	0.00	92.87	0.00
fwd.15	0.89	37.36	0.95	0.00	93.02	1.71	-0.15	88.06	0.95	0.00	92.93	-0.29
las.15	1.07	29.51	0.98	0.00	93.05	2.01	0.37	71.97	0.98	0.00	93.00	0.69
bwd.20	0.99	30.30	0.94	0.00	93.33	1.98	0.01	99.45	0.94	0.00	93.26	0.01
fwd.20	1.06	26.41	0.94	0.00	93.46	2.14	0.06	94.71	0.94	0.00	93.36	0.13
las.20	0.99	30.43	0.97	0.00	93.67	1.97	0.27	78.22	0.97	0.00	93.63	0.53
<i>Panel B: Bimonthly Rebalancing</i>												
bwd.10	0.40	71.61	0.97	0.00	92.04	0.70	-0.33	76.44	0.97	0.00	91.96	-0.58
fwd.10	0.62	57.71	0.98	0.00	91.90	1.07	-0.08	93.96	0.98	0.00	91.82	-0.15
las.10	1.18	31.14	1.00	0.00	91.54	1.94	0.67	56.37	1.00	0.00	91.48	1.11
bwd.15	1.26	20.47	0.95	0.00	92.97	2.43	0.58	56.06	0.95	0.00	92.90	1.12
fwd.15	1.05	29.16	0.95	0.00	92.98	2.02	0.35	72.89	0.95	0.00	92.92	0.67
las.15	1.59	12.21	0.98	0.00	92.92	2.97	1.10	28.53	0.98	0.00	92.88	2.05
bwd.20	1.08	25.81	0.94	0.00	93.31	2.17	0.41	67.12	0.94	0.00	93.24	0.81
fwd.20	1.29	17.97	0.94	0.00	93.27	2.58	0.62	52.12	0.94	0.00	93.21	1.23
las.20	1.28	18.02	0.97	0.00	93.70	2.57	0.78	41.81	0.97	0.00	93.66	1.55
<i>Panel C: Quarterly Rebalancing</i>												
bwd.10	0.40	71.84	0.97	0.00	91.99	0.69	-0.18	86.71	0.97	0.00	91.91	-0.32
fwd.10	0.06	95.93	0.97	0.00	91.76	0.10	-0.49	65.89	0.97	0.00	91.69	-0.85
las.10	1.08	34.75	1.00	0.00	91.65	1.80	0.68	55.53	1.00	0.00	91.60	1.13
bwd.15	1.00	31.86	0.94	0.00	92.86	1.91	0.45	65.60	0.94	0.00	92.80	0.86
fwd.15	0.93	35.17	0.95	0.00	92.94	1.79	0.40	68.90	0.95	0.00	92.87	0.77
las.15	1.12	28.18	0.98	0.00	92.73	2.07	0.73	48.53	0.98	0.00	92.70	1.34
bwd.20	0.99	30.14	0.93	0.00	93.24	1.98	0.44	64.59	0.93	0.00	93.18	0.88
fwd.20	1.07	26.49	0.93	0.00	93.25	2.14	0.54	57.07	0.93	0.00	93.21	1.09
las.20	0.97	31.69	0.96	0.00	93.55	1.92	0.58	54.52	0.96	0.00	93.52	1.16

Table 4.5: Out of Sample Results for the Index Tracking Portfolios without Transaction Costs (IBOV)

	Mean	SD	SR	Mean TE	SD TE	IR	alpha	beta	R2
IBOV	18.38	31.52	0.58	0.00	0.00	<i>NaN</i>	0.00	1.00	1.00
Panel A: Monthly Rebalancing									
bwd.10	0.99	1.02	0.57	-0.23	9.14	-0.03	0.24	0.97	0.92
fwd.10	1.04	1.02	0.60	0.77	9.16	0.08	1.18	0.98	0.92
las.10	1.14	1.05	0.63	2.55	9.51	0.27	2.42	1.01	0.92
bwd.15	1.09	0.99	0.64	1.58	8.40	0.19	2.50	0.95	0.93
fwd.15	1.07	0.99	0.63	1.37	8.36	0.16	2.23	0.95	0.93
las.15	1.13	1.02	0.65	2.37	8.48	0.28	2.70	0.98	0.93
bwd.20	1.07	0.97	0.64	1.38	8.14	0.17	2.49	0.94	0.93
fwd.20	1.09	0.98	0.65	1.62	8.06	0.20	2.67	0.94	0.93
las.20	1.10	1.00	0.64	1.91	8.00	0.24	2.49	0.97	0.94
Panel B: Bimonthly Rebalancing									
bwd.10	1.03	1.02	0.59	0.53	9.07	0.06	1.00	0.97	0.92
fwd.10	1.07	1.02	0.61	1.21	9.20	0.13	1.56	0.98	0.92
las.10	1.17	1.05	0.65	3.05	9.62	0.32	2.97	1.00	0.92
bwd.15	1.12	0.98	0.66	2.19	8.38	0.26	3.17	0.95	0.93
fwd.15	1.10	0.99	0.65	1.79	8.39	0.21	2.65	0.95	0.93
las.15	1.20	1.02	0.69	3.65	8.55	0.43	4.02	0.98	0.93
bwd.20	1.09	0.97	0.65	1.58	8.15	0.19	2.73	0.94	0.93
fwd.20	1.12	0.97	0.67	2.14	8.18	0.26	3.25	0.94	0.93
las.20	1.14	1.00	0.67	2.62	7.97	0.33	3.23	0.97	0.94
Panel C: Quarterly Rebalancing									
bwd.10	1.03	1.02	0.59	0.53	9.10	0.06	1.00	0.97	0.92
fwd.10	0.98	1.02	0.56	-0.36	9.23	-0.04	0.14	0.97	0.92
las.10	1.15	1.05	0.64	2.74	9.52	0.29	2.73	1.00	0.92
bwd.15	1.08	0.98	0.64	1.49	8.44	0.18	2.51	0.94	0.93
fwd.15	1.07	0.98	0.64	1.36	8.40	0.16	2.34	0.95	0.93
las.15	1.13	1.01	0.65	2.39	8.65	0.28	2.83	0.98	0.93
bwd.20	1.07	0.97	0.65	1.29	8.19	0.16	2.50	0.93	0.93
fwd.20	1.08	0.97	0.65	1.49	8.19	0.18	2.69	0.93	0.93
las.20	1.10	1.00	0.64	1.76	8.06	0.22	2.43	0.96	0.94

**Source:** The author.

**Note:** In this Table the benchmark index is the Bovespa Index (IBOV). The Table shows annualized Mean and Standard Deviation (SD) of the Index and of the portfolio returns. The index is shown in absolute terms and the portfolios are shown in relation to the index (Mean and SD of the portfolios are divided by the Mean and the SD of the index). SR denotes the annualized Sharpe Ratio. The Table also shows the Mean and SD of the Tracking Error of the portfolios (portfolio return minus the index). And the IR denotes the annualized Information Ratio which is the Mean of the Tracking Error divided by its SD. The last three columns of the Table show the alfa, beta, and  $R^2$  (intercept in annualized terms, the slope, and the coefficient of determination) of a univariate regression of the out of sample portfolio returns against the benchmark Index in the period from January 2000 to December of 2010.

Table 4.6: Out of Sample Results for the Index Tracking Portfolios with Transaction Costs (IBOV)

	Mean	SD	SR	TO	Short	Active	alpha	beta	R2
MKT	18.38	31.52	0.58	NA	NA	NA	0.00	1.00	1.00
Panel A: Monthly Rebalancing									
bwd.10	0.84	1.02	0.48	2.13	-0.08	10.00	-2.43	0.97	0.92
fwd.10	0.89	1.02	0.51	2.18	-0.24	10.00	-1.57	0.98	0.92
las.10	1.04	1.05	0.58	1.40	-0.74	9.99	0.66	1.01	0.92
bwd.15	0.95	0.99	0.56	1.99	-2.50	15.00	-0.00	0.95	0.93
fwd.15	0.93	0.99	0.55	2.08	-1.65	15.00	-0.38	0.95	0.93
las.15	1.03	1.02	0.59	1.41	-0.55	14.93	0.93	0.98	0.93
bwd.20	0.94	0.97	0.56	1.96	-8.04	20.00	0.02	0.94	0.93
fwd.20	0.95	0.98	0.57	2.00	-6.59	20.00	0.16	0.94	0.93
las.20	1.00	1.00	0.59	1.44	-1.14	19.95	0.67	0.97	0.94
Panel B: Bimonthly Rebalancing									
bwd.10	0.93	1.02	0.53	1.46	0.00	10.00	-0.83	0.97	0.92
fwd.10	0.97	1.02	0.55	1.41	-0.48	10.00	-0.21	0.98	0.92
las.10	1.10	1.05	0.61	1.01	-0.97	9.99	1.70	1.00	0.91
bwd.15	1.03	0.98	0.61	1.36	-2.72	15.00	1.46	0.95	0.93
fwd.15	1.00	0.99	0.59	1.41	-2.08	15.00	0.88	0.95	0.93
las.15	1.13	1.02	0.65	0.98	-0.59	14.97	2.78	0.98	0.93
bwd.20	0.99	0.97	0.60	1.35	-8.59	20.00	1.03	0.94	0.93
fwd.20	1.02	0.97	0.61	1.34	-7.12	20.00	1.56	0.94	0.93
las.20	1.07	1.00	0.63	1.01	-0.98	19.93	1.96	0.97	0.94
Panel C: Quarterly Rebalancing									
bwd.10	0.95	1.02	0.54	1.17	0.00	10.00	-0.46	0.97	0.92
fwd.10	0.90	1.01	0.52	1.11	0.00	10.00	-1.24	0.97	0.92
las.10	1.09	1.04	0.61	0.81	-0.80	10.00	1.72	1.00	0.92
bwd.15	1.00	0.98	0.60	1.11	-2.20	15.00	1.12	0.94	0.93
fwd.15	1.00	0.98	0.59	1.06	-1.68	15.00	1.01	0.95	0.93
las.15	1.08	1.01	0.62	0.79	-0.44	15.00	1.84	0.98	0.93
bwd.20	0.99	0.97	0.60	1.11	-7.99	20.00	1.11	0.93	0.93
fwd.20	1.01	0.97	0.61	1.05	-6.59	20.00	1.37	0.93	0.93
las.20	1.04	1.00	0.61	0.77	-1.08	19.96	1.47	0.96	0.94

**Source:** The author.

**Note:** In this Table the benchmark index is the Bovespa Index (IBOV). The Table shows annualized Mean and Standard Deviation (SD) of the out of sample portfolio returns net of transaction costs. We used  $c = 50bp$  as in DeMiguel, Garlappi and Uppal (2009). SR is the annualized Sharpe Ratio of those returns. TO is the turnover and is presented as a daily percentage average. Short interest is the size of the short position and is expressed in percentage. Active is the average number of active positions in the portfolio. The last three columns of the Table show the alfa, beta, and  $R^2$  (intercept in annualized terms, the slope, and the coefficient of determination) of a univariate regression of the out of sample portfolio returns net of transaction errors against the benchmark Index in the period from January 2000 to December of 2010.

Table 4.7: Out of Sample Results for the Index Tracking Portfolios without Transaction Costs (SP500)

	Mean	SD	SR	Mean TE	SD TE	IR	alpha	beta	R2
SP500	12.93	12.22	1.06	0.00	0.00	<i>NaN</i>	0.00	1.00	1.00
Panel A: Monthly Rebalancing									
bwd.20	1.16	1.03	1.19	2.01	2.75	0.73	2.01	1.00	0.95
fwd.20	1.05	1.04	1.07	0.66	2.90	0.23	0.48	1.01	0.95
las.20	1.26	1.06	1.26	3.39	4.10	0.83	3.38	1.00	0.90
bwd.30	1.06	1.02	1.11	0.80	2.22	0.36	0.82	1.00	0.97
fwd.30	1.09	1.03	1.12	1.16	2.39	0.49	1.05	1.01	0.96
las.30	1.20	1.03	1.23	2.61	3.39	0.77	2.67	1.00	0.93
bwd.40	1.10	1.01	1.16	1.36	1.93	0.70	1.41	1.00	0.98
fwd.40	1.12	1.02	1.16	1.55	2.08	0.75	1.46	1.01	0.97
las.40	1.12	1.03	1.15	1.52	2.93	0.52	1.55	1.00	0.95
Panel B: Bimonthly Rebalancing									
bwd.20	1.07	1.01	1.12	0.91	2.86	0.32	1.08	0.99	0.95
fwd.20	1.13	1.03	1.15	1.62	2.79	0.58	1.55	1.01	0.95
las.20	1.25	1.06	1.25	3.21	4.06	0.79	3.13	1.01	0.90
bwd.30	0.99	1.01	1.04	-0.16	2.31	-0.07	-0.02	0.99	0.96
fwd.30	1.19	1.02	1.23	2.48	2.35	1.05	2.44	1.00	0.96
las.30	1.20	1.03	1.23	2.62	3.39	0.77	2.69	0.99	0.93
bwd.40	1.04	1.00	1.10	0.47	1.96	0.24	0.62	0.99	0.97
fwd.40	1.18	1.01	1.23	2.35	2.03	1.16	2.35	1.00	0.97
las.40	1.16	1.02	1.20	2.02	2.99	0.68	2.16	0.99	0.94
Panel C: Quarterly Rebalancing									
bwd.20	1.26	1.01	1.31	3.35	2.73	1.23	3.48	0.99	0.95
fwd.20	1.16	1.04	1.18	2.06	2.85	0.72	1.87	1.02	0.95
las.20	1.22	1.06	1.22	2.84	4.18	0.68	2.84	1.00	0.90
bwd.30	1.23	1.00	1.30	2.99	2.22	1.34	3.21	0.98	0.97
fwd.30	1.20	1.02	1.24	2.65	2.33	1.14	2.55	1.01	0.97
las.30	1.19	1.04	1.22	2.45	3.44	0.71	2.49	1.00	0.93
bwd.40	1.24	1.00	1.31	3.09	1.92	1.61	3.27	0.99	0.98
fwd.40	1.18	1.01	1.23	2.34	1.93	1.21	2.31	1.00	0.98
las.40	1.16	1.03	1.20	2.06	2.93	0.70	2.10	1.00	0.95

**Source:** The author.

**Note:** In this Table the benchmark index is the SP500 Index (SP500). The Table shows annualized Mean and Standard Deviation (SD) of the Index and of the portfolio returns. The index is shown in absolute terms and the portfolios are shown in relation to the index (Mean and SD of the portfolios are divided by the Mean and the SD of the index). SR denotes the annualized Sharpe Ratio. The Table also shows the Mean and SD of the Tracking Error of the portfolios (portfolio return minus the index). And the IR denotes the annualized Information Ratio which is the Mean of the Tracking Error divided by its SD. The last three columns of the Table show the alpha, beta, and  $R^2$  (intercept in annualized terms, the slope, and the coefficient of determination) of a univariate regression of the out of sample portfolio returns against the benchmark Index in the period from January 2012 to October of 2017.



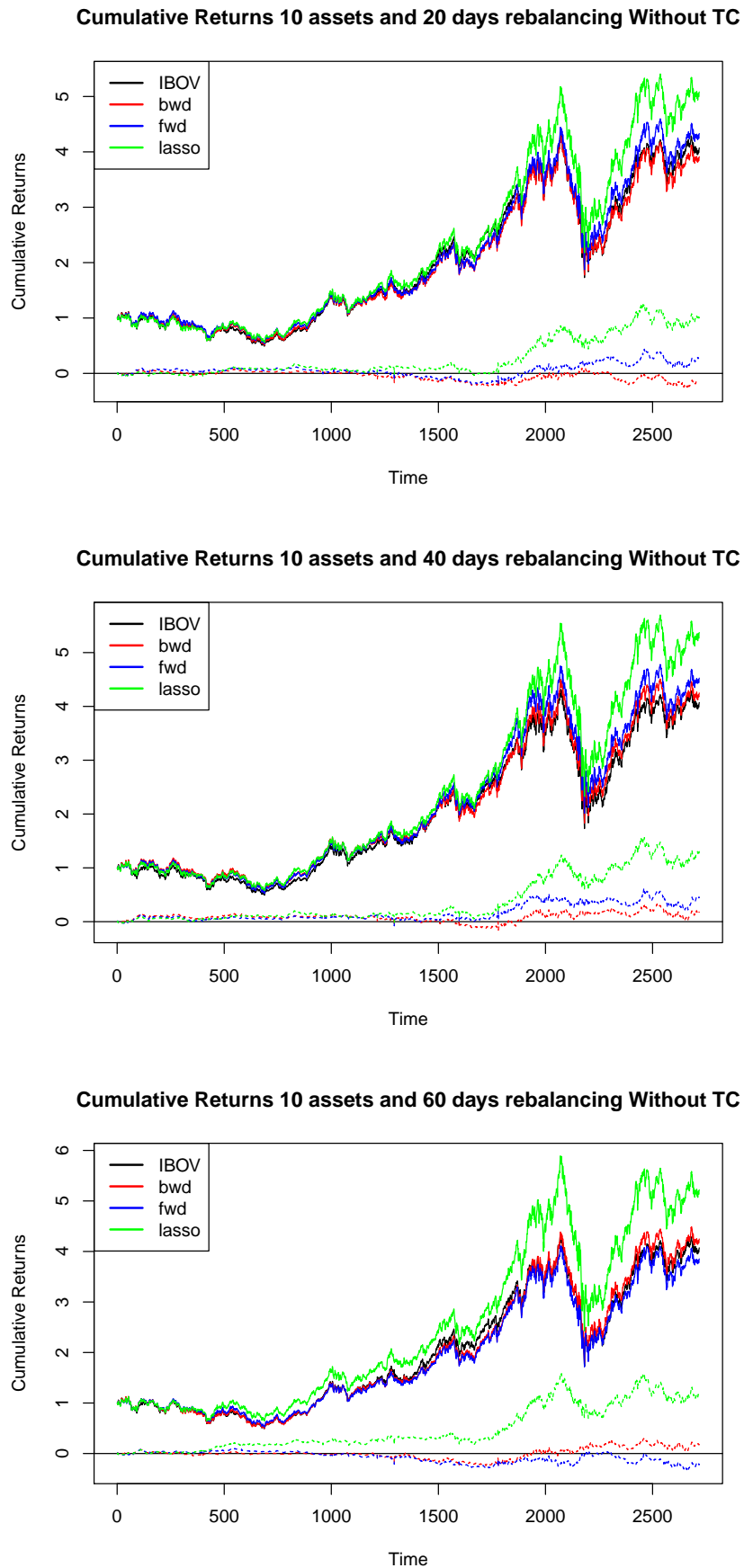
Table 4.8: Out of Sample Results for the Index Tracking Portfolios with Transaction Costs (SP500)

	Mean	SD	SR	TO	Short	Active	alpha	beta	R2
SP500	12.93	12.22	1.06	NA	NA	NA	0.00	1.00	1.00
Panel A: Monthly Rebalancing									
bwd.20	0.53	1.05	0.54	6.42	0.00	20.00	-6.15	1.01	0.92
fwd.20	0.53	1.06	0.53	5.39	-0.03	20.00	-6.35	1.02	0.93
las.20	1.09	1.06	1.09	1.82	-5.85	19.93	1.08	1.00	0.90
bwd.30	0.47	1.03	0.48	6.07	0.00	30.00	-6.89	1.00	0.94
fwd.30	0.56	1.04	0.57	5.43	-0.08	30.00	-5.84	1.01	0.94
las.30	1.02	1.04	1.04	1.89	-9.01	29.88	0.29	1.00	0.93
bwd.40	0.55	1.03	0.57	5.66	0.00	40.00	-5.77	1.00	0.95
fwd.40	0.59	1.04	0.60	5.48	-0.13	40.00	-5.48	1.01	0.95
las.40	0.92	1.03	0.94	2.07	-8.91	39.88	-1.07	1.00	0.94
Panel B: Bimonthly Rebalancing									
bwd.20	0.74	1.03	0.76	3.43	0.00	20.00	-3.26	0.99	0.93
fwd.20	0.85	1.04	0.87	2.82	-0.05	20.00	-2.02	1.01	0.94
las.20	1.13	1.06	1.13	1.18	-5.89	19.92	1.65	1.01	0.90
bwd.30	0.66	1.02	0.69	3.33	0.00	30.00	-4.23	0.99	0.95
fwd.30	0.92	1.03	0.94	2.84	-0.06	30.00	-1.14	1.00	0.95
las.30	1.07	1.03	1.10	1.32	-8.68	29.92	1.02	0.99	0.93
bwd.40	0.74	1.01	0.77	3.09	0.00	40.00	-3.29	0.99	0.96
fwd.40	0.91	1.02	0.94	2.81	-0.08	40.00	-1.20	1.00	0.96
las.40	1.01	1.02	1.05	1.45	-9.34	39.92	0.33	0.99	0.94
Panel C: Quarterly Rebalancing									
bwd.20	1.02	1.02	1.05	2.49	0.00	20.00	0.36	0.99	0.94
fwd.20	0.96	1.04	0.97	2.06	0.00	20.00	-0.71	1.01	0.94
las.20	1.12	1.06	1.12	1.05	-5.74	19.96	1.52	1.00	0.89
bwd.30	1.01	1.01	1.06	2.30	0.00	30.00	0.32	0.98	0.95
fwd.30	1.00	1.03	1.03	2.08	-0.03	30.00	-0.06	1.01	0.96
las.30	1.08	1.04	1.10	1.10	-9.04	29.92	1.10	1.00	0.92
bwd.40	1.03	1.00	1.09	2.11	0.00	40.00	0.62	0.99	0.96
fwd.40	0.98	1.02	1.02	2.04	-0.04	40.00	-0.25	1.00	0.97
las.40	1.05	1.03	1.08	1.17	-8.45	39.88	0.62	1.00	0.94

**Source:** The author.

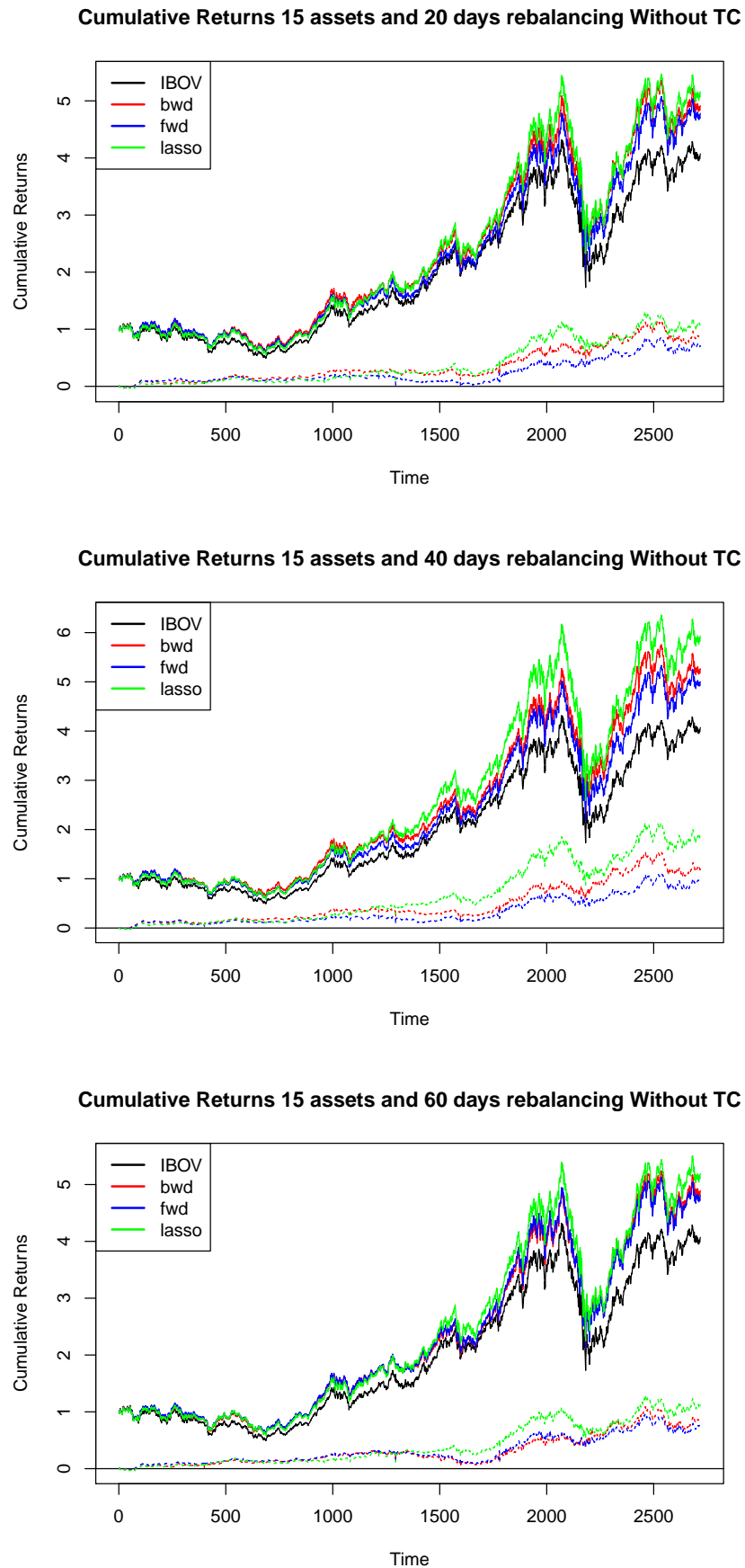
**Note:** In this Table the benchmark index is the SP500 Index (SP500). The Table shows annualized Mean and Standard Deviation (SD) of the out of sample portfolio returns net of transaction costs. We used  $c = 50bp$  as in DeMiguel, Garlappi and Uppal (2009). SR is the annualized Sharpe Ratio of those returns. TO is the turnover and is presented as a daily percentage average. Short interest is the size of the short position and is expressed in percentage. Active is the average number of active positions in the portfolio. The last three columns of the Table shows the alfa, beta, and  $R^2$  (intercept in annualized terms, the slope, and the coefficient of determination) of a univariate regression of the out of sample portfolio returns against the benchmark Index in the period from January 2012 to October of 2017.

Figure 4.1: Cumulative returns with 10 assets (IBOV)



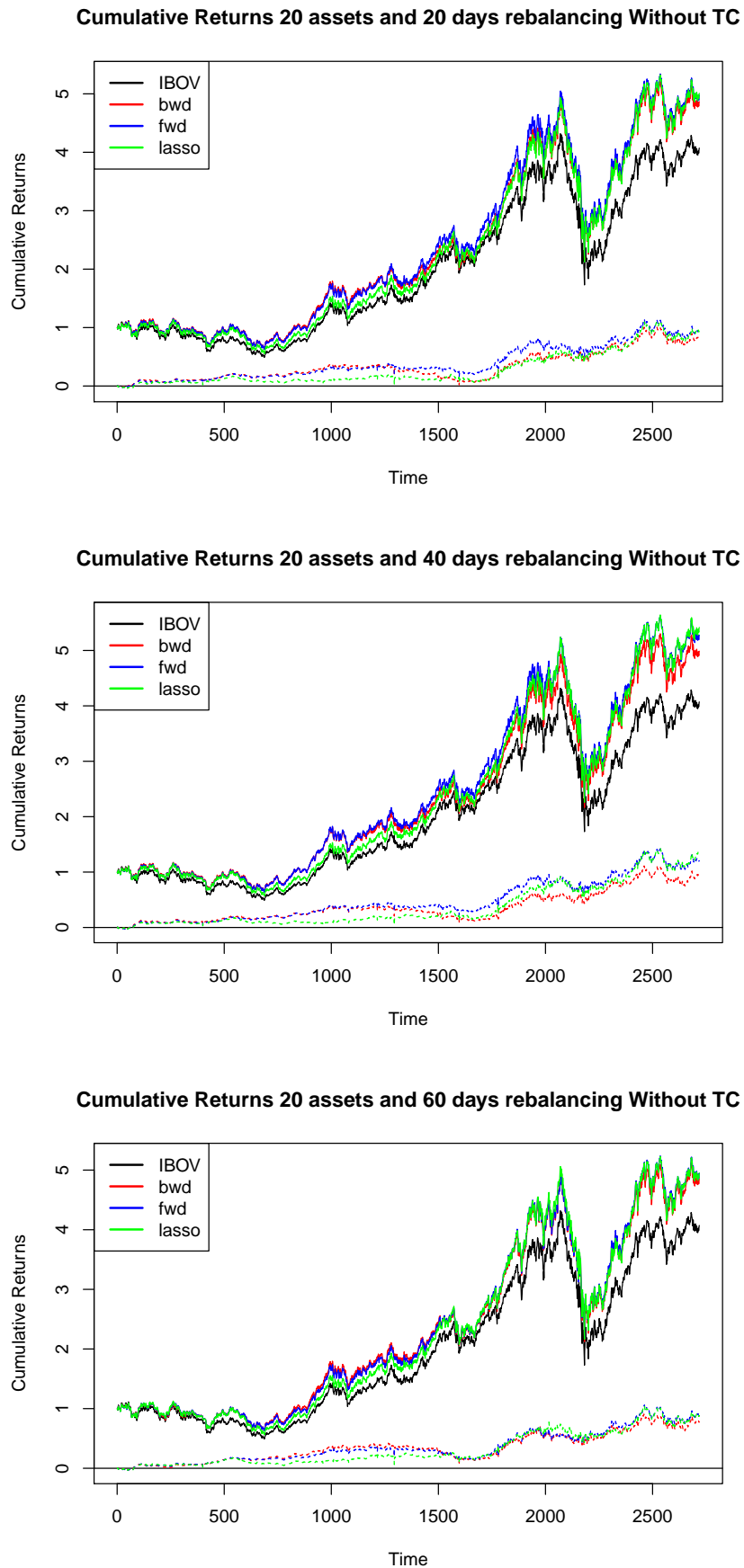
Source: The author.

Figure 4.2: Cumulative returns with 15 assets (IBOV)



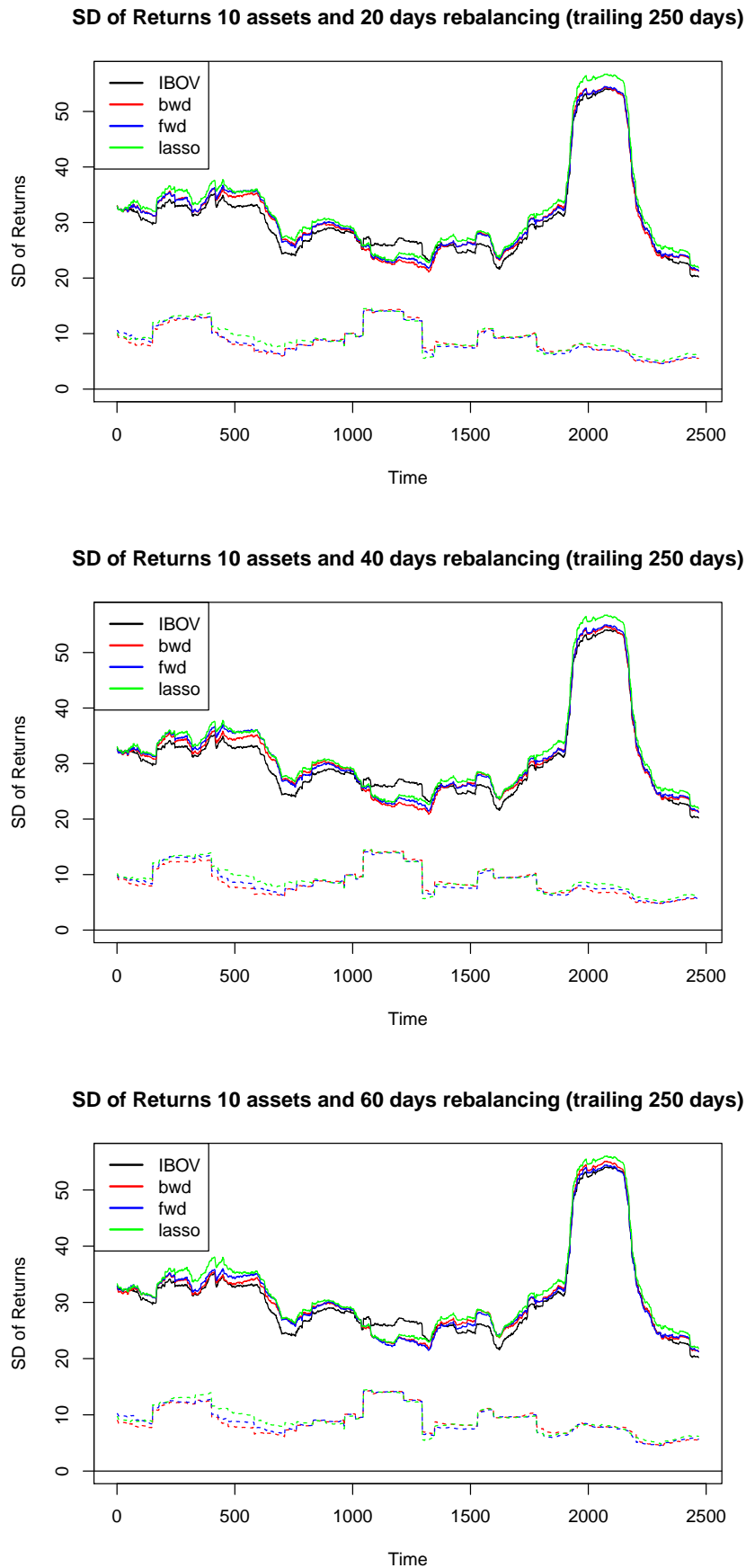
Source: The author.

Figure 4.3: Cumulative returns with 20 assets (IBOV)



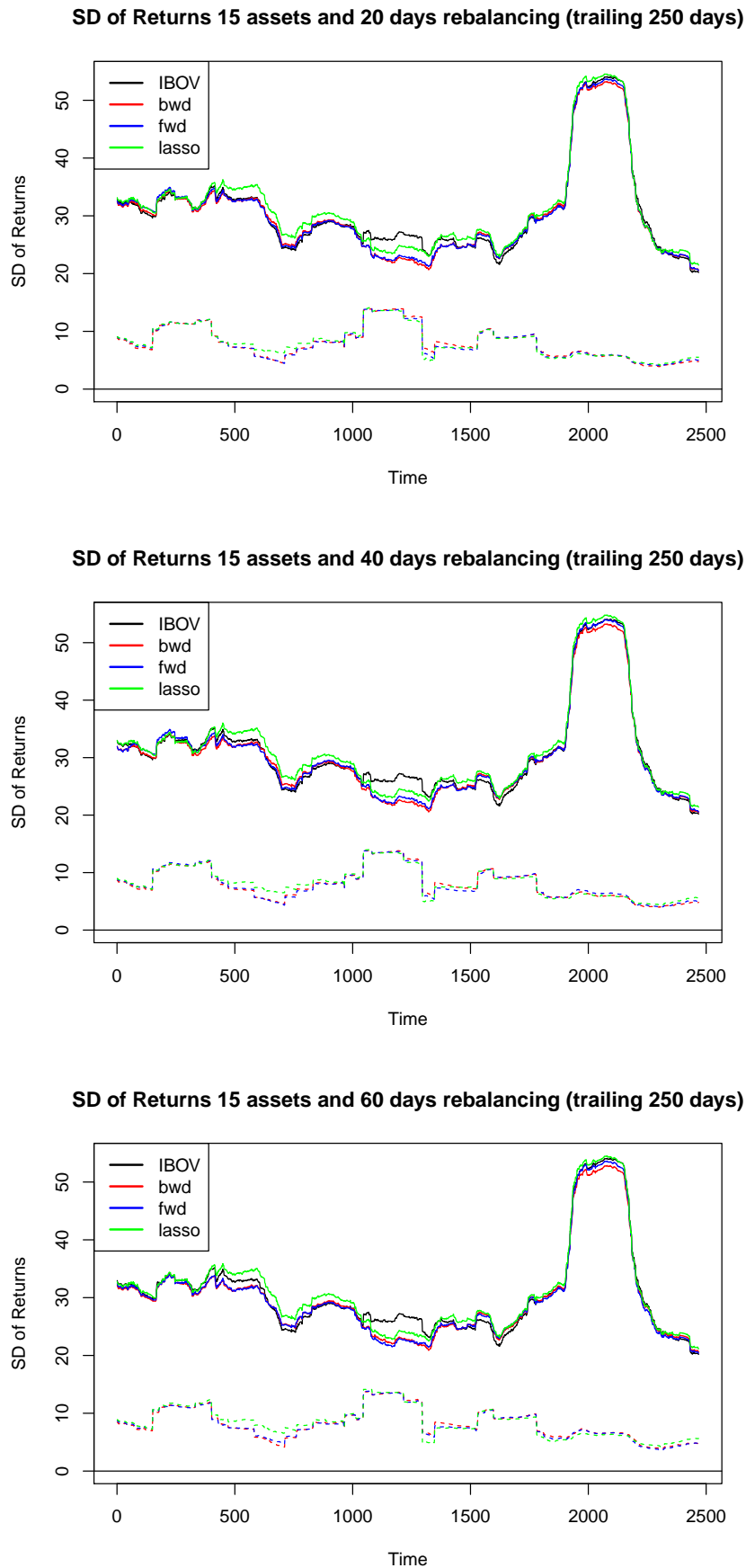
Source: The author.

Figure 4.4: Volatility of returns with 10 assets (IBOV)



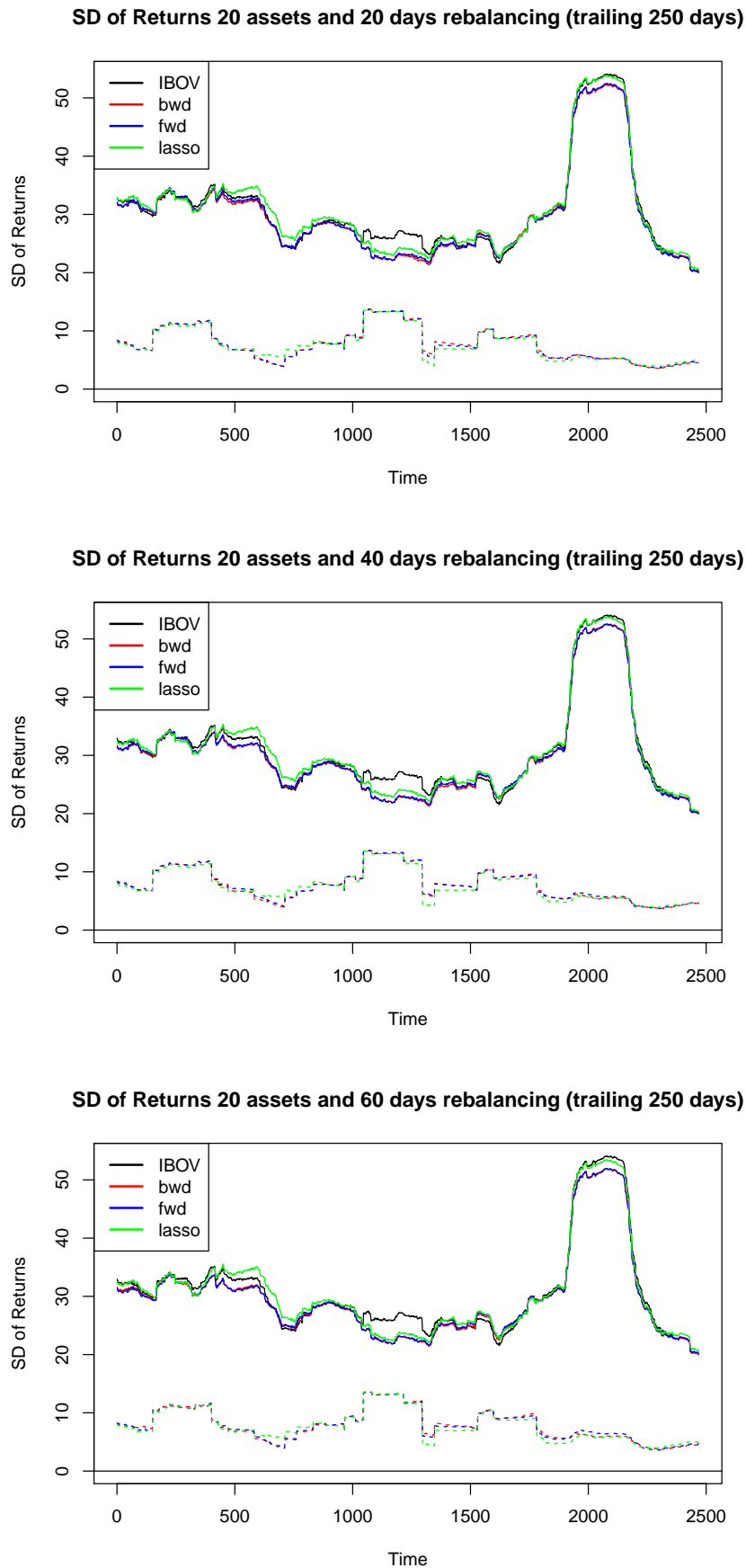
Source: The author.

Figure 4.5: Volatility of returns with 15 assets (IBOV)



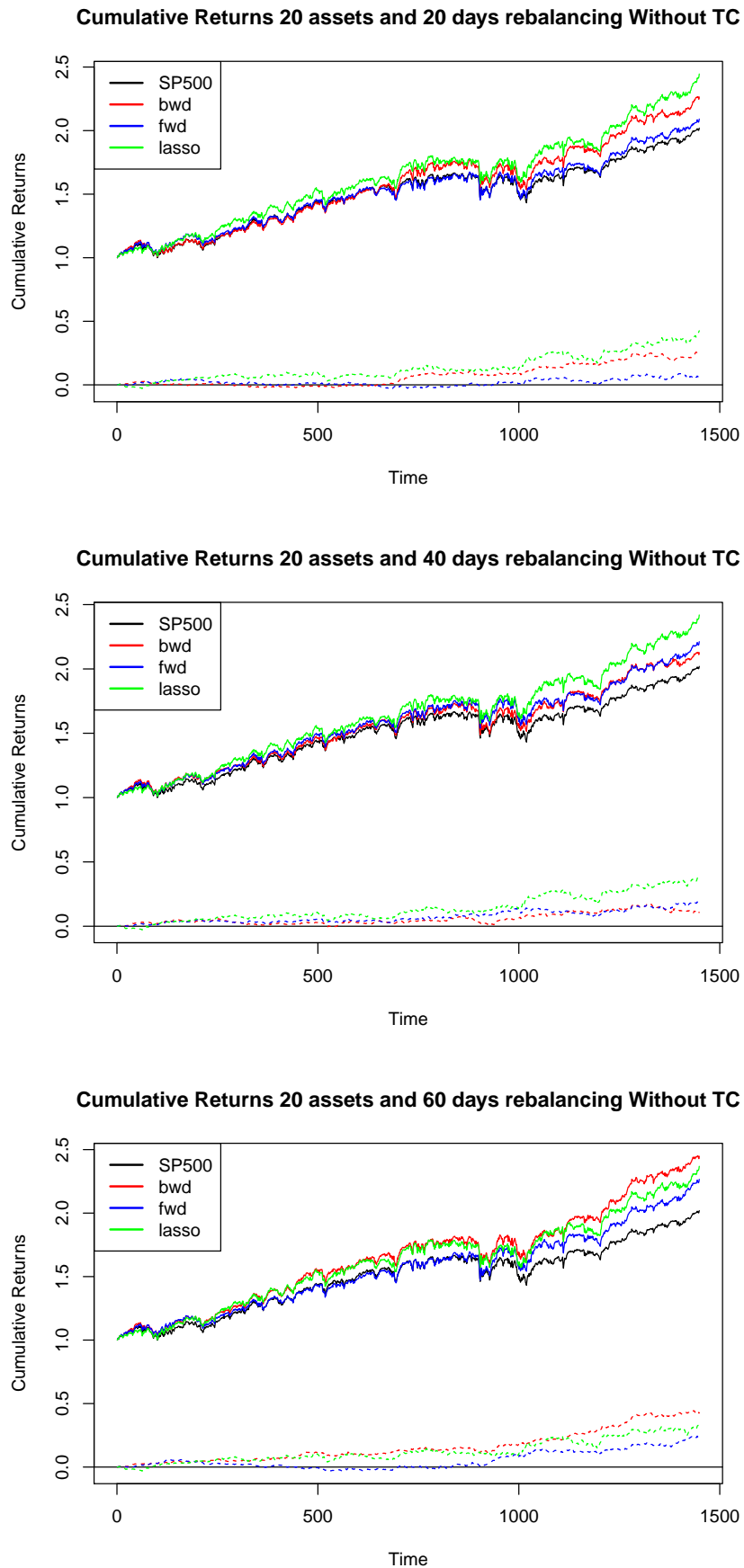
Source: The author.

Figure 4.6: Volatility of returns with 20 assets (IBOV)



Source: The author.

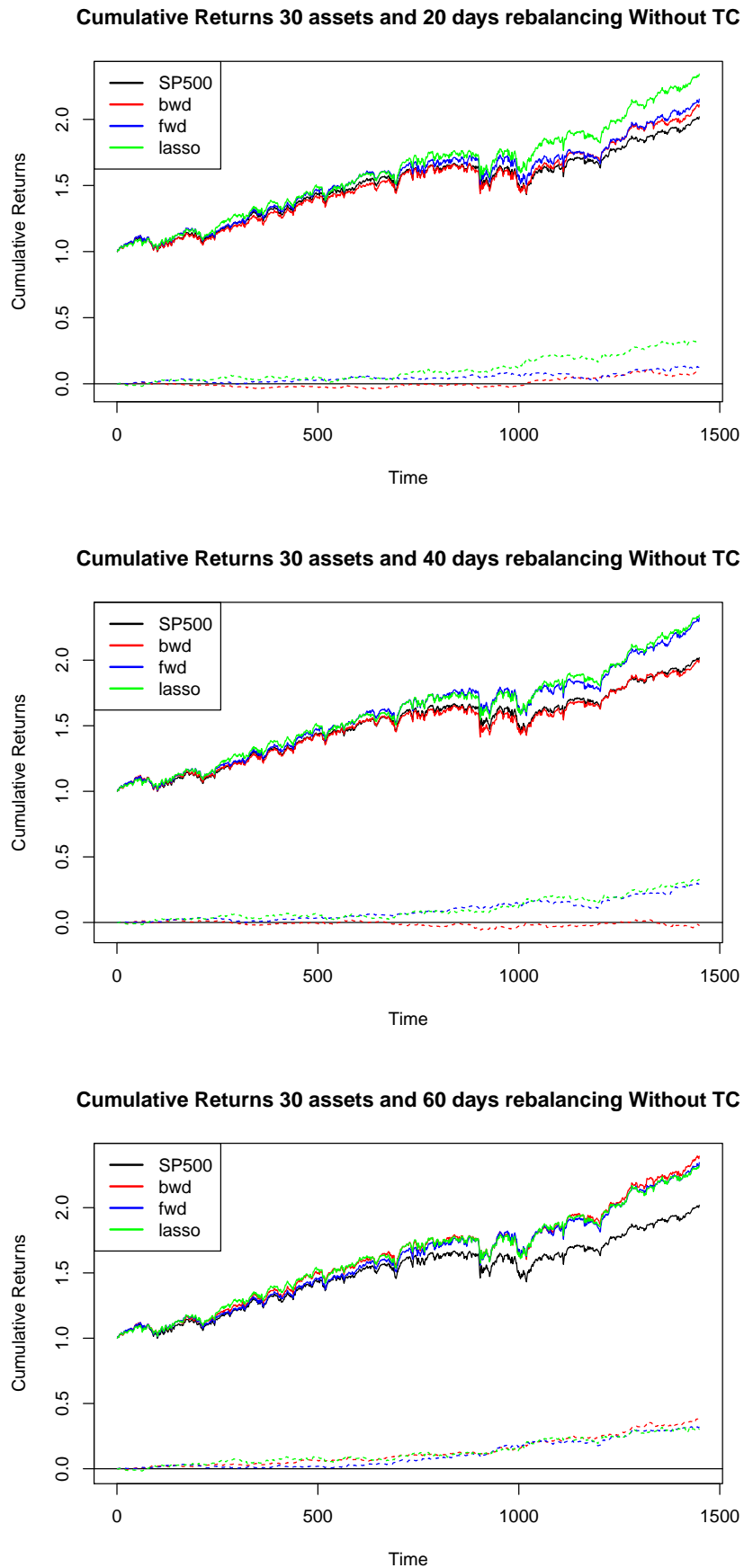
Figure 4.7: Cumulative returns with 20 assets (SP500)



Source: The author.

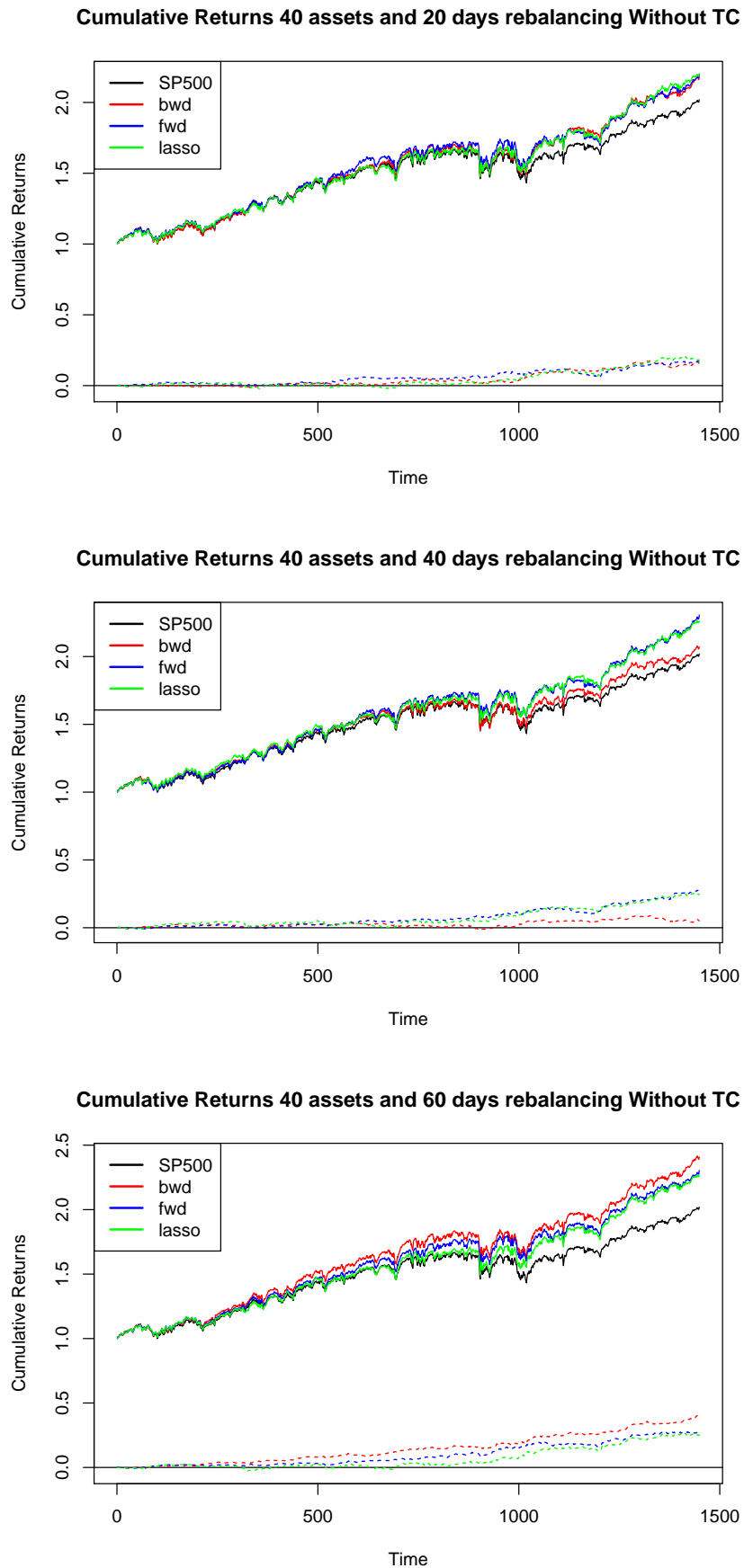


Figure 4.8: Cumulative returns with 30 assets (SP500)



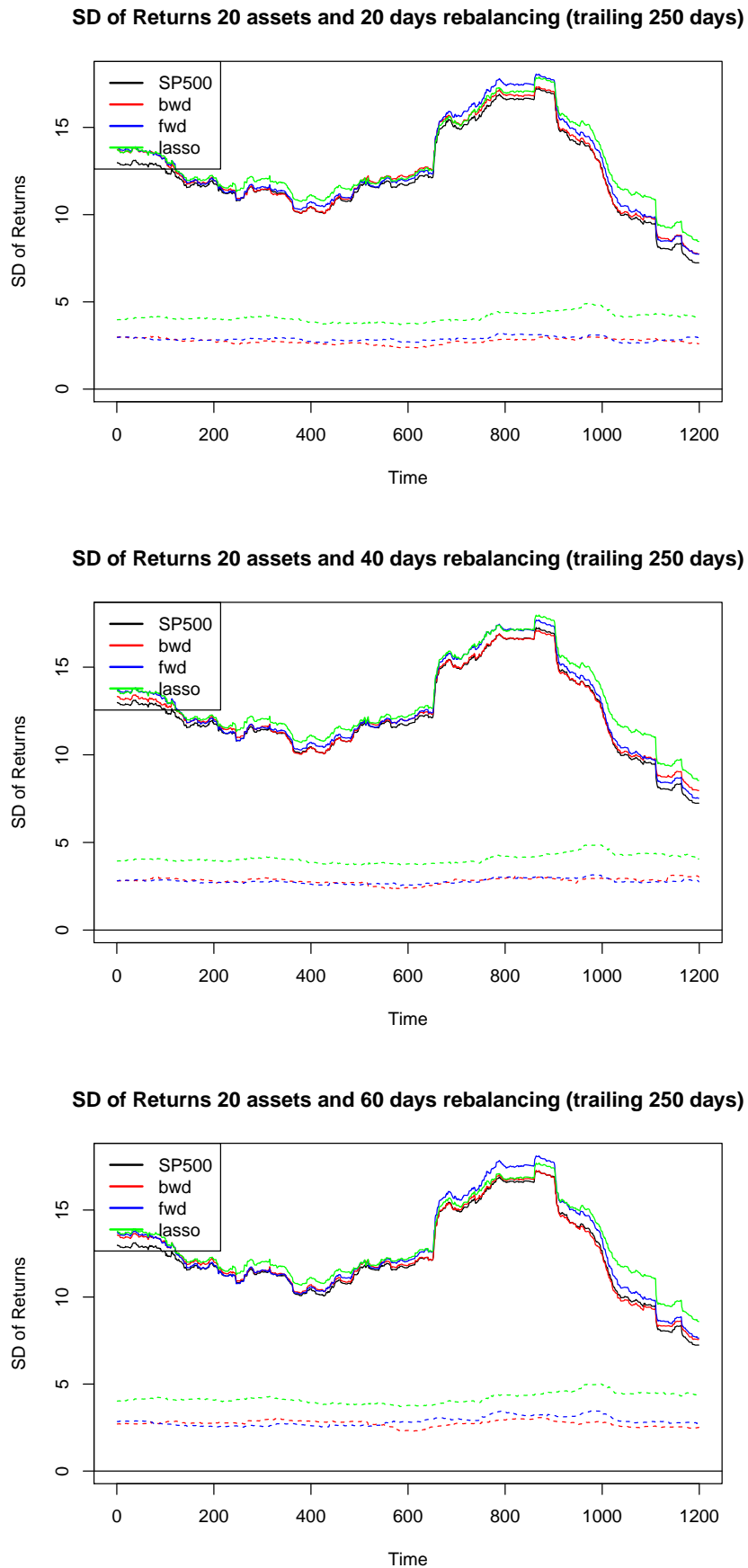
Source: The author.

Figure 4.9: Cumulative returns with 40 assets (SP500)



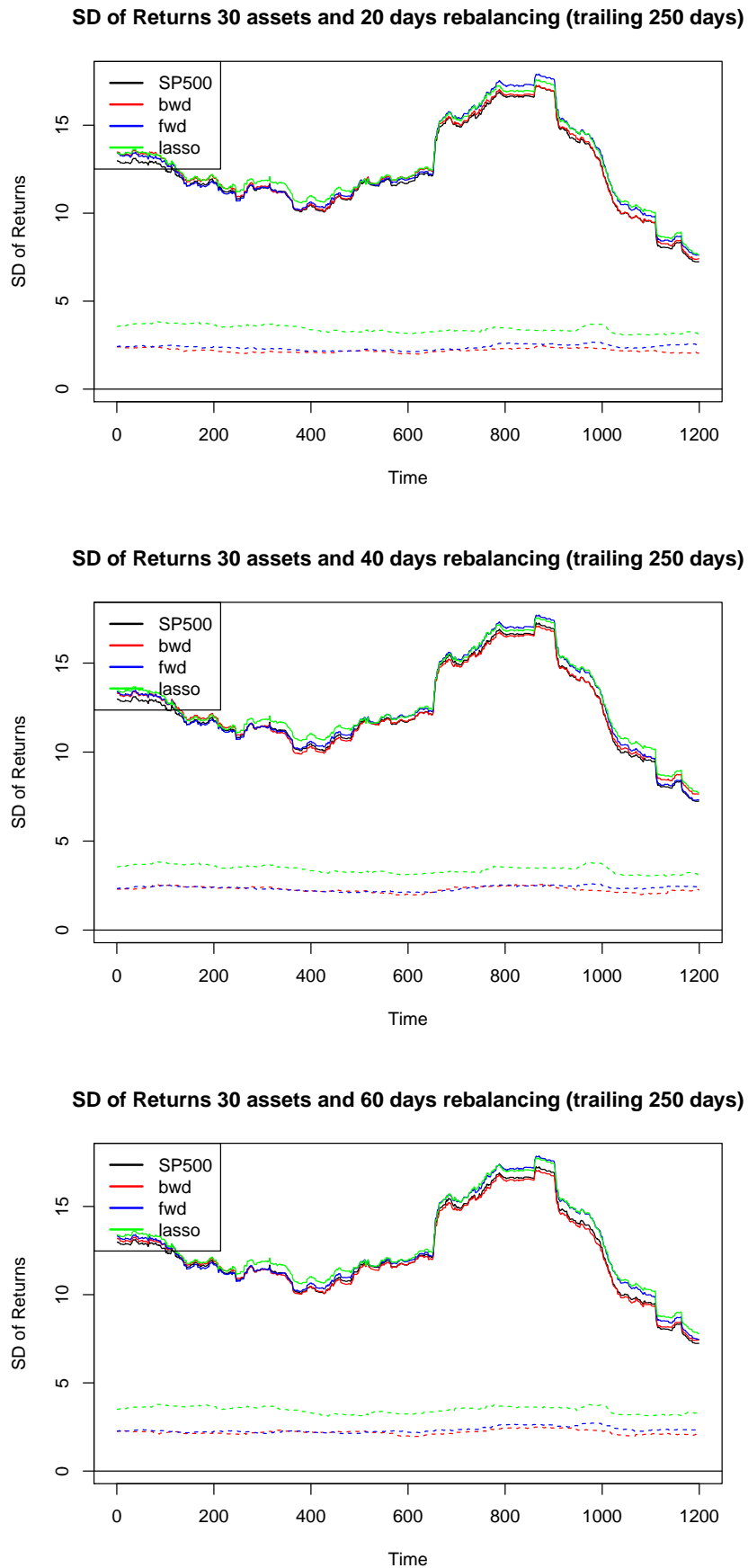
Source: The author.

Figure 4.10: Volatility of returns with 20 assets (SP500)



Source: The author.

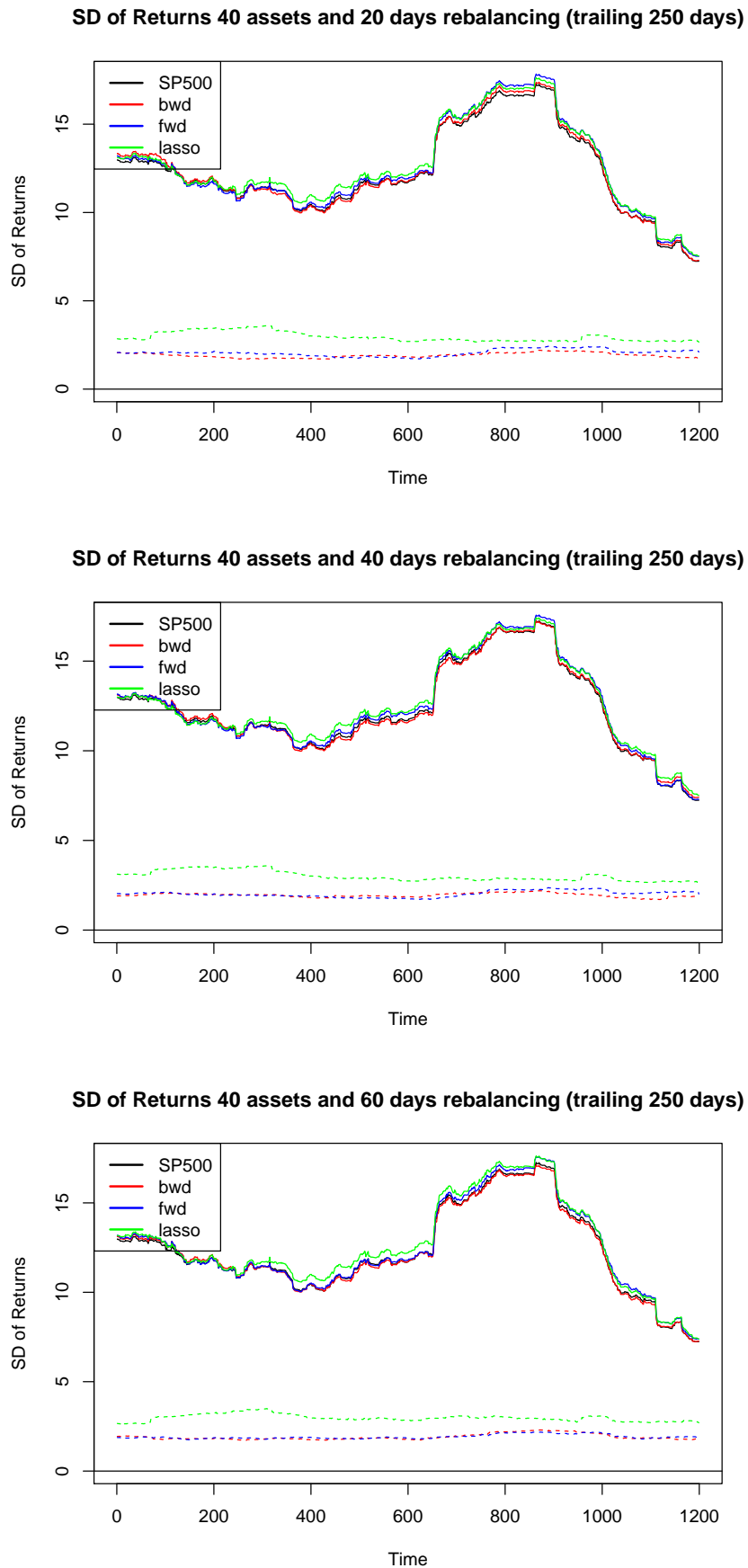
Figure 4.11: Volatility of returns with 30 assets (SP500)



Source: The author.

Source: The author.

Figure 4.12: Volatility of returns with 40 assets (SP500)



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).<sup>1</sup> 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 experiment 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 reape.

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 December 2010

---

<sup>1</sup>A preliminary version of this essay was accepted to the XIX Econtro Brasileiro de Finanças with the title “Tangency Portfolio: A Critique”

where we try to follow the Bovespa Index. The results we find is that for the brazilian case the algorithms deliver similar Tracking Error volatility, but the lasso delivers higher returns, which leads to a higher risk-adjusted performance. For the American case, the lasso also delivers better risk-adjusted performance, but with higher volatility than the other methods. This is undesirable in our case, so we favor the other methods for the American case. 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*, v. 55, n. 2, p. 773–806, 2000.
- BEST, M. *Portfolio Optimization*. [S.l.]: Taylor & Francis, 2010. (Chapman & Hall/CRC finance series).
- BEST, M. J.; GRAUER, R. R. Sensitivity Analysis for Mean-Variance Portfolio Problems. *Management Science*, v. 37, n. 8, p. 980–989, August 1991.
- BODNAR, T.; PAROLYA, N.; SCHMID, W. On the equivalence of quadratic optimization problems commonly used in portfolio theory. *European Journal of Operational Research*, 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: Tools and Techniques*. 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*, v. 54, n. 2, p. 655–671, 1999.
- BRODIE, J. et al. Sparse and stable markowitz portfolios. *Proceedings of the National Academy of Sciences*, 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*, 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*, 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. *The Journal of Portfolio Management*, v. 33, n. 1, p. 10–24, 2006.
- CONSTANTINIDES, G.; MALLIARIS, A. Portfolio theory. In: *Finance*. [S.l.]: 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*, 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*. [S.l.]: Wiley-Blackwell, 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*, 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*, 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*, v. 58, n. 4, p. 1651–1684, 08 2003.
- JAMES, G. et al. *An Introduction to Statistical Learning: With Applications in R*. [S.l.]: Springer Publishing Company, Incorporated, 2014.
- JANSEN, R.; DIJK, R. van. Optimal benchmark tracking with small portfolios. *The Journal of Portfolio Management*, 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*, v. 42, n. 3, p. 621–656, 2007.
- KEMPF, A.; MEMMEL, C. Estimating the global minimum variance portfolio. *Schmalenbach Business Review*, 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*, 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*, 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*, v. 66, n. 2, p. 31–39, 2010.
- LEDOIT, O.; WOLF, M. Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, v. 15, n. 5, p. 850–859, December 2008.
- LEDOIT, O.; WOLF, M. Robust performance hypothesis testing with the Variance. *Wilmott Magazine*, p. 86–89, September 2011.
- LIU, Q. On portfolio optimization: How and when do we benefit from high-frequency data? *Journal of Applied Econometrics*, v. 24, n. 4, p. 560–582, 2009.
- LUMLEY, T. *leaps: Regression Subset Selection*. [S.l.], 2017. R package version 3.0.
- MARKOWITZ, H. Portfolio Selection. *Journal of Finance*, v. 7, n. 1, p. 77–91, 03 1952.
- MARKOWITZ, H. *Portfolio Selection: Efficient Diversification of Investments*. [S.l.]: Yale University Press, 1959. (Cowles Commission for Research in Economics: Monographs). ISBN 9780300013726.
- MCWILLIAMS, B.; MONTANA, G. Sparse partial least squares regression for on-line variable selection with multivariate data streams. *Statistical Analysis and Data Mining*, 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*, v. 99, n. 2, p. nil, 2015.
- MERTON, R. C. An Analytic Derivation of the Efficient Portfolio Frontier. *Journal of Financial and Quantitative Analysis*, v. 7, n. 04, p. 1851–1872, September 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*, v. 134, n. 1, p. 235 – 256, 2006.
- PETERSEN, K. B.; PEDERSEN, M. S. *The Matrix Cookbook*. [S.l.]: Technical University of Denmark, 2012. Version 20121115.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2019.
- RAO, C. R. et al. *Linear Models and Generalizations*. 3rd. ed. [S.l.]: Springer, 2008. (Springer Series in Statistics).
- RENCHEER, A. C.; SCHAALJE, G. B. *Linear Models in Statistics*. 2nd. ed. [S.l.]: 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*, 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*, v. 227, p. 541 – 552, 2014.
- WU, L.; YANG, Y.; LIU, H. Nonnegative-lasso and application in index tracking. *Computational Statistics & Data Analysis*, v. 70, p. 116 – 126, 2014.