



HAL
open science

Mean-Variance Efficient Large Portfolios: A Simple Machine Learning Heuristic Technique based on the Two-Fund Separation Theorem

Michele Costola, Bertrand Maillet, Zhining Yuan, Xiang Zhang

► **To cite this version:**

Michele Costola, Bertrand Maillet, Zhining Yuan, Xiang Zhang. Mean-Variance Efficient Large Portfolios: A Simple Machine Learning Heuristic Technique based on the Two-Fund Separation Theorem. *Annals of Operations Research*, 2024, 334 (1-3), 133-155 p. hal-04514343

HAL Id: hal-04514343

<https://hal.science/hal-04514343>

Submitted on 21 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Mean-Variance Efficient Large Portfolios: A Simple Machine Learning Heuristic Technique based on the Two-Fund Separation Theorem

Michele Costola¹ · Bertrand Maillet² · Zhining Yuan³ · Xiang Zhang⁴

(Accepted Version – July 2022)

Abstract.

We revisit in this article the Two-Fund Separation Theorem as a simple technique for the Mean-Variance optimization of large portfolios. The proposed approach is fast and scalable and provides equivalent results of commonly used ML techniques but, with computing time differences counted in hours (1 minute *versus* several hours). In the empirical application, we consider three geographic areas (China, US, and French stock markets) and show that the Two-Fund Separation Theorem holds exactly when no constraints are imposed and is approximately true with (realistic) positive constraints on weights. This technique is shown to be of interest to both scholars and practitioners involved in portfolio optimization tasks.

Keywords. Mean-Variance Efficient Portfolios, Two-Fund Separation Theorem, Machine Learning, Robust Portfolios, High-dimensional Portfolios.

¹ Ca' Foscari University of Venice, Venice, Italy; Leibniz Institute for Financial Research SAFE, Frankfurt, Germany – michele.costola@unive.it

² emlyon business school (AIM QUANT Research Center), Paris, France; Univ. La Reunion, Saint-Denis, France; Variances, Paris, France – bmaillet@em-lyon.com

³ emlyon business school (AIM QUANT Research Center), Paris, France; Coactis, Univ. of Lyon, Lyon, France – zhining.yuan@edu.em-lyon.com

⁴ School of Finance/Institute of Big Data Research, Southwestern University of Finance and Economics (SWUFE), Chengdu, China – xiangzhang@swufe.edu.cn

1. Introduction

Mutual fund separation is a key concept in portfolio selection and asset allocation. It implies that all rational investors' optimal portfolio choices can be built as a linear combination of a set of mutual funds whatever the initial wealth level is, where a mutual fund can be any portfolio of tradable assets in the market. Thus, under mutual fund separation, any investor can achieve the same level of utility from the individual assets, as if they were only offered a set of mutual funds. As noted by Dybvig and Liu (2018), the term "separation" comes from the fact that any investor can separate her portfolio choice in two steps. First, she chooses a small set of funds that spans optimal portfolios of all wealth levels. Secondly, the investor determines the optimal mixture of the separating funds based on her current wealth level.

The aim of this article is to revisit the classical fund separation theorem as a tool to simplify the computations of optimal Mean-Variance portfolios (Markowitz, 1952), for a large number of assets. This long-time established theory (see Markowitz, 1999) has experienced in recent decades a resurgence of research activity and the topic of robust portfolio optimization has become quite substantial, as shown for instance by Markowitz (2014) and Kolm *et al.* (2014), who review the 60-year course of portfolio optimization. However, the existing practitioner applied finance literature mainly uses the so-called Generalized Reduced Gradient (GRC) algorithm (*e.g.*, Abadie, 1969), available in the MS Excel Solver function for instance, to optimize more-than-two-asset portfolios (see also Laws, 2003; Grover and Lavin, 2007; Bai *et al.*, 2011). The brute force

results obtained using a solver are accurate, which possibly explains the lack of contemporary studies within the Mean-Variance framework. In this article, we propose a simple alternative to improve on the existing mainstream framework. The implemented methodology requires less computational power, is automatized and, shows the interrelation between assets' returns and the logic of the portfolio optimization process. Furthermore, it can ultimately be run on a simple spreadsheet application.

Differently from the existing approaches mentioned above, our proposal implies a different road and takes advantage of some well-known results from the Two-Fund Separation Theorem⁵ (*Cf.* Tobin, 1958). which is already largely mentioned in financial decision theory reference textbooks, such as in Ingersoll (1987) and in Huang and Litzenberger (1988)⁶. Specifically, we show that the findings for the two-asset case can be easily generalized to an N -asset portfolio, with additional unrevealed interesting properties that allow building optimal portfolios directly.

Several studies show applications of regression-type approaches to obtain optimal Mean-Variance portfolios (*e.g.*, Broadie, 1993; Stevens, 1998; Britten-Jones, 1999; Kempf and Memmel, 2006). Using penalized regressions, recent literature

⁵ See also Markowitz (1959); Sharpe (1964); Mossin (1966); Lintner, (1965a) and (1965b); Black (1972), as well as Dybvig and Liu (2018), and further adaptations of the original Separation Theorem when agents do not exhibit Mean-Variance preferences or when the densities of return are not Gaussian, as in Pye (1967), Hakansson (1969), Cass and Stiglitz (1970), Harvey and Siddique (2000), Jurczenko and Maillet (2006), Dahlquist *et al.* (2017), or in a continuous-time stochastic setting, as in Samuelson (1967), Merton (1973), Ross (1978), and more recently in Cairns *et al.* (2006), Deguest *et al.* (2018) and Bernard *et al.* (2021).

⁶ We will adopt hereafter the presentation and (very similar) notations used in both of these books.

provides benefits in terms of the sparsity of the portfolio (indirectly associated with diversification/concentration, positive weights and turnover) and elaborates on the properties of optimal portfolios in various market conditions (*e.g.*, Brodie *et al.*, 2009; DeMiguel *et al.*, 2009a; Fan *et al.*, 2012; Yen and Yen, 2014; Fastrich *et al.*, 2015; Bonaccolto *et al.*, 2018; Bonaccolto and Paterlini, 2020; Bonaccolto *et al.*, 2021). In this article, we show that the Two-Fund Separation Theorem approximately holds when supplementary realistic constraints are considered in the one-period Markowitz' (1952) seminal setting. In this regard, we apply the proposed approach by considering three continents (Chinese, US, and French stock markets), large databases (with more than 100 assets), and sub-samples of dates and stocks, trying to deal with uncertainties of the results in the presence of model risks.

The contribution of our article can be summarized as follows: 1) we provide a general structure of the weight of every optimal portfolio; 2) we provide a Mahalanobis-type transformation for every desirable (and applicable) weight to become optimal; 3) we show that the Two-Fund Separation Theorem holds under several realistic constraints and thus, it provides insightful results for financial practitioners.

The rest of the article is organized as follows. Section 2 briefly presents the optimization in the N -asset case. Section 3 recalls Mean-Variance portfolios in the regression-type framework. Section 4 provides a discussion on the optimal portfolio weights resulting from applications in a penalized regression framework. Section 5 focuses on the results provided by our approach and includes robustness checks. Finally, Section 6 concludes. A companion Web Appendix

provides sketches of the demonstration of results presented within the *corpus* of the article, the descriptive statistics of the database we use in the empirical studies, as well as some complementary robustness tests and illustrations.

2. The Traditional Portfolio Optimization in the N -asset Case

Let us now consider N assets. Matrix notations show that the expected return on portfolio p , denoted $E(R_p)$, writes:

$$E(R_p) = \mathbf{w}'_p \mathbf{E}, \quad (1)$$

where \mathbf{w}'_p is the $(N \times 1)$ transposed vector of weights \mathbf{w}_p , in which, on each line, an invested share of wealth is allocated in each asset, and \mathbf{E} is the $(N \times 1)$ vector of the expected returns on the N assets.

The variance of the portfolio returns, denoted $\sigma^2(R_p)$, will be:

$$\sigma^2(R_p) = \mathbf{w}'_p \mathbf{\Omega} \mathbf{w}_p, \quad (2)$$

where the $(N \times N)$ Variance-CoVariance matrix of returns on the assets is denoted $\mathbf{\Omega}$.

Finally, the budget constraint when saturated, reads:

$$\mathbf{w}'_p \mathbf{1} = 1, \quad (3)$$

with $\mathbf{1}$ the $(N \times 1)$ unit vector⁷.

As *per* definition, a portfolio is a Minimum-Variance Portfolio on the Efficient Frontier, if it has the smallest variance amongst all feasible portfolios that have the same targeted expected rate of

⁷ For the sake of simplicity and clarification, each time we have a vector of 1 with dimensions $(N \times 1)$ with N being the number of the assets, we use: $\mathbf{1} := \mathbf{1}_{(N \times 1)}$ as a notation, whilst for a vector of 1 with dimensions $(T \times 1)$, with T being the number of dates of the time-series of returns on an asset, we denote it as: $\mathbf{1}_{(T \times 1)}$. When we will need a $(N \times N)$ matrix of 1, we will further denote it: $\mathbf{1}_{(N \times N)}$.

return. The weights of any Minimum-Variance Frontier Portfolio are thus the solution of the following program:

$$\begin{aligned} \mathbf{w}_p^* &= \arg \min_{\mathbf{w}_p \in \mathbb{R}^N} \left\{ \frac{1}{2} \mathbf{w}_p' \boldsymbol{\Omega} \mathbf{w}_p \right\} \\ & \text{s.t.} \quad \begin{cases} \mathbf{w}_p' \mathbf{E} = \bar{\mu}_p \\ \mathbf{w}_p' \mathbf{1} = 1, \end{cases} \end{aligned} \quad (4)$$

where $\bar{\mu}_p$ is the targeted expected return for portfolio p such as: $\bar{\mu}_p = E(R_p)$, that can be approximated in the naive approach⁸ by the sample mean $\hat{\mu}_p$.

Let us first define the Mahalanobis-type distance⁹ rescaled transform of any vector: $\mathbf{Y} \rightarrow \mathbf{Y}^\circ \rightarrow \check{\mathbf{Y}}$, as such:

$$\mathbf{Y}^\circ = \boldsymbol{\Omega}^{-1} \mathbf{Y}, \quad (5)$$

and:

$$\check{\mathbf{Y}} = \frac{\mathbf{Y}^\circ}{\mathbf{1}' \mathbf{Y}^\circ} = \frac{\boldsymbol{\Omega}^{-1} \mathbf{Y}}{\mathbf{1}' \boldsymbol{\Omega}^{-1} \mathbf{Y}}. \quad (6)$$

This transform thus uses a vector of asset weights denoted \mathbf{Y} , first multiplied by $\boldsymbol{\Omega}^{-1}$, the inverse of the Variance-CoVariance matrix, that becomes \mathbf{Y}° , which is the Mahalanobis-type norm of the original vector \mathbf{Y} (*i.e.*, $\mathbf{Y}^\circ = \|\mathbf{Y}\|_{\mathbf{M}^\circ}$, where $\|\cdot\|_{\mathbf{M}^\circ}$ is the Mahalanobis-type norm operator). Then, secondly, this vector \mathbf{Y}° is rescaled in order to obtain a vector of weights denoted $\check{\mathbf{Y}}$, summing at unity (such as: $\mathbf{1}' \check{\mathbf{Y}} = 1$). And so we have finally: $\check{\mathbf{Y}} = \|\mathbf{Y}\|_{\check{\mathbf{M}}}$, where the rescaled Mahalanobis-type transform operator $\|\mathbf{Y}\|_{\check{\mathbf{M}}}$,

⁸ See, for instance, Black and Litterman (1992) for an alternative approach with deformed expectations according to views in a Bayesian context.

⁹ See Mahalanobis (1927, 1936); see also Mardia *et al.* (1979), Meucci (2009), Breuer *et al.* (2009), Breuer and Csiszár (2013), Geyer *et al.* (2014), Flood and Korenko (2015), McNeil *et al.* (2015) and De Meo (2021) for other uses of the Mahalanobis norm in financial applications.

changes a non-linear relation into an equivalent linear one, as we will see below. The transformed vector $\tilde{\mathbf{Y}}$ will be thus a vector of asset weights, used as a “pivotal vector”, whose relative structure is defined by the original vector \mathbf{Y} , whilst saturating the budget constraint.

For simplifying the notations, define now the following classical scalars¹⁰:

$$\begin{cases} a = \mathbf{1}'\boldsymbol{\Omega}^{-1}\mathbf{E} \\ b = \mathbf{E}'\boldsymbol{\Omega}^{-1}\mathbf{E} \\ c = \mathbf{1}'\boldsymbol{\Omega}^{-1}\mathbf{1} \\ d = (bc - a^2), \end{cases} \quad (7)$$

where $\mathbf{1}$ is the $(N \times 1)$ vector of 1 (with N equals to the number of the assets) and $\mathbf{1}_N$ is a $(N \times N)$ matrix with all elements equal to 1.

We note that a , b and c can be expressed as inner products on \mathbb{R}^n , where *per* definition the inner product of two vectors \mathbf{X} and \mathbf{Y} is defined here as: $\langle \mathbf{X}, \mathbf{Y} \rangle_{\boldsymbol{\Omega}} = \mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{Y}$. Also note that, from the Cauchy-Schwarz inequality (leading to $a^2 \leq bc$), the parameter d is nonnegative¹¹.

Using next the quadratic relationship between risk and return entails the following propositions, that will lead themselves to the definition of the set of the Feasible Portfolios, in the so-called opportunity set, whose envelop will be denoted \mathcal{F} , when the

¹⁰ Here we follow Huang and Litzenberger (1988) p.64 with $A = a$, $B = b$, $C = c$, $D = d$, whilst Ingersoll (1987) p.54 uses $B = a$, $C = b$, $A = c$, and $\Delta = d$.

¹¹ Furthermore, it is not equal to zero, and positive in our financial applications – see the Appendix, which allows us to use the term d in some denominators below, with no risk of indetermination.

restricted set of Efficient Portfolios is denoted herein $\overline{\mathcal{F}}$ (and the lower part of the set of dominated portfolios will be $\underline{\mathcal{F}}$)¹².

Forming now the Lagrangian of the program in Eq. (4) and using notations of the scalars in Eq. (7), will lead after some straightforward calculations to the following property, which links, on the one hand, the weights of any Efficient Portfolios and, on the other hand, those of some special ones.

Proposition 1.

The weight of any Efficient Portfolio \mathbf{p} , belonging to $\overline{\mathcal{F}}$, can be expressed as:

$$\mathbf{w}_p^* = \Delta_{a^*R_p|b^*} \check{\mathbf{E}} + (1 - \Delta_{a^*R_p|b^*}) \check{\mathbf{I}}, \quad (8)$$

*where $\Delta_{a^*R_p|b^*} = a^*E(R_p) - b^*$, with $a^* = ac/d$ and $b^* = a^2/d$, is a transform of the first Lagrangian coefficient of the global program in Eq. (4), which can be recognized as a Rescaled Fitted Performance Measure¹³, comparing the targeted expected return of the portfolio to b^* , whilst $\check{\mathbf{E}}$ and $\check{\mathbf{I}}$ are the Mahalanobis-type norm defined in Eq. (7).*

Proof. See the Appendix.

Thus, these general expressions of Frontier Portfolios weights (on \mathcal{F}) only depend upon some exogenous market variables denoted a , b , c and d , and on a subjective variable: the targeted portfolio

¹² We will use later on, similarly, the notations \mathcal{F}^* , $\overline{\mathcal{F}}^*$ and $\underline{\mathcal{F}}^*$ for the same defined sets in a constrained context.

¹³ See also Gouriéroux and Jouneau (1999) for a discussion about fitted performance measures.

expected return, which depends on the investor's preferences and needs.

For better explanations in the rest of this article, we define some Pivotal Portfolios, before further utilizing with necessities. Here we start with the so-called Global Minimum-Variance Portfolio (GMVP, denoted (V) for *Vertex* hereafter), which *per* definition produces the minimum variance by a combination of all feasible assets.

Moreover, let us further introduce two other special portfolios, starting first with the so-called Tangent Portfolio (T). It is the Efficient Portfolio with the highest possible Sharpe ratio¹⁴, thus belonging to both the Efficient Frontier and the tangent line to the Efficient Frontier starting from the risk-free asset, called the Capital Market Line (CML).

Then, let us consider some other generic portfolios denoted by Z_p . In general, any arbitrary Frontier Portfolio p (except the GMVP) has a unique uncorrelated Frontier Portfolio (or zero-correlation / zero-*beta* portfolio), denoted here Z_p . Indeed, from the Mean-CoVariance relation between any couple of Frontier Portfolios p and q belonging to \mathcal{F} , as such¹⁵:

$$\begin{aligned} Cov(R_p, R_q) &= \mathbf{w}_p' \mathbf{\Omega} \mathbf{w}_q & (9) \\ &= [cE(R_p)E(R_q) - a[E(R_p) + E(R_q)] + b]/d, \end{aligned}$$

and since by definition for the two portfolios p and Z_p , we have:

$$Cov(R_p, R_{Z_p}) = 0, \quad (10)$$

so that the following equality holds:

¹⁴ such as: $S_p = [E(R_p) - R_f]/\sigma(R_p)$, with R_f the risk-free rate.

¹⁵ Cf. Ingersoll (1987), p.56 and Huang and Litzenberger (1988), p.66.

$$E(R_p) = \frac{[aE(R_{Z_p}) - b]}{[cE(R_{Z_p}) - a]}. \quad (11)$$

Among all of these portfolios Z_p , let us highlight the specific zero-mean Frontier Portfolio denoted G (see below for more details on this portfolio), as being the zero-mean Frontier Portfolio of a particular Frontier Portfolio N , called the Null Index Portfolio (*i.e.*, N is a special Z_G , or equivalently, in other words, G is a specific Z_N).

From setting $E(R_{Z_N}) = E(R_G) = 0$ in the Eq. (11), we have:

$$E(R_N) = \frac{[aE(R_{Z_G}) - b]}{[cE(R_{Z_G}) - a]} = \frac{b}{a}, \quad (12)$$

and when we set the target expected return as in Eq. (12) in the expression of all optimal portfolios in Proposition 1, we then have:

$$\begin{aligned} \mathbf{w}_N &= \Delta_{a^*R_N|b^*} \check{\mathbf{E}} + (1 - \Delta_{a^*R_N|b^*}) \check{\mathbf{I}} \\ &= \check{\mathbf{E}}, \end{aligned} \quad (13)$$

where $\Delta_{a^*R_N|b^*} = a^*E(R_N) - b^* = 1$ with $a^* = ac/d$ and $b^* = a^2/d$.

Finally, the covariance between the returns on Portfolio N and on Portfolio G is (as *per* definition of portfolios N and G): $Cov(R_N, R_G) = 0$, where G is a special (zero-correlation) Frontier Portfolio (generally on the below part $\underline{\mathcal{F}}$ of \mathcal{F} in normal circumstances with a positive return on the GMVP) having an expected return of 0, whilst the portfolio N is an Efficient Portfolio (on the above part $\overline{\mathcal{F}}$ of \mathcal{F}) as illustrated in the next Figure 1 below.

The following proposition basically tells us that the weights of any of the Efficient Portfolios denoted \mathbf{w}_p^* (for instance \mathbf{w}_V , \mathbf{w}_N and \mathbf{w}_T , respectively) share the same structure, which is composed by

the weight of a special vector \mathbf{Y} ($\mathbf{1}$, \mathbf{E} , \mathbf{E}_x , respectively), expressed in a Mahalanobis-type norm (divided by $\mathbf{\Omega}^{-1}$), becoming thus \mathbf{Y}° ($\mathbf{1}^\circ$, \mathbf{E}° , \mathbf{E}_x° , respectively), then rescaled to respect the budget constraint (*i.e.* divided by the unscaled sum of weights of the special vector into consideration), becoming ultimately $\check{\mathbf{Y}}$ ($\check{\mathbf{1}}$, $\check{\mathbf{E}}$, $\check{\mathbf{E}}_x$, respectively).

We will now make use of the Two-Fund Separation Theorem (*e.g.*, Tobin, 1958; Markowitz, 1959; Sharpe, 1964; Mossin, 1966; Lintner, 1965a and 1965b; Black, 1972). Basically, this theorem says that any Mean-Variance Efficient Portfolio can be generated by any two distinct Mean-Variance Frontier Portfolios.

In Proposition 2, we show hereafter that all Efficient Portfolios share the same structure of weights and are based on an affine transform of a linear combination of two special notorious portfolios (\mathbf{E} and $\mathbf{1}$).¹⁶

Proposition 2.

All Efficient Portfolios in $\overline{\mathcal{F}}$ share the same structure of weights \mathbf{w}_p^ that reads:*

$$\mathbf{w}_p^* = \check{\mathbf{Y}}, \tag{14}$$

with:

$$\mathbf{Y} = f\{\delta_0 \mathbf{E} + \delta_1 \mathbf{1}\},$$

¹⁶ See also Corollary 2.1 in the appendix A-4, when inverting the problem and setting first the targeted expected return (for Asset-Liability Management reasons for instance), we deduce the optimal related Efficient Portfolio and its linked implicit risk and just *vice-versa* in the Corollary 2.2 (see the appendix A-5). Corollary 2.3 (see in the appendix A-6) also interestingly illustrates the proper structure of some notorious Portfolios according to Proposition 2, with explicit parameters.

where $f(\cdot)$ is an affine transform of the underlying vector and the $\delta_i(\cdot)$, for $i = \{0,1\}$, are functions of expected return, $E(R_p) = \bar{\mathbf{Y}}' \mathbf{E}$ that ultimately goes with a risk that reads: $\sigma^2(R_p) = \bar{\mathbf{Y}}' \mathbf{\Omega} \bar{\mathbf{Y}}$.

Proof. See the Appendix.

With the help of Proposition 2 (and also with the related Corollary 2.1, 2.2 and 2.3 of Proposition 2 reported in the appendix A-4, A-5 and A-6), users can easily self-construct their satisfying portfolio considering the different expectations of return and the endurance of risk.

Furthermore, if a risk-free asset exists, the Efficient Frontier $\bar{\mathcal{F}}$ is the Capital Market Line joining the risk-free rate and the Tangent Portfolio (T). In this case, the following Proposition 3 can be stated.

Proposition 3.

The weight of any Efficient Portfolio p , belonging to $\bar{\mathcal{F}}$, can be formed on the basis of two pivotal Efficient Portfolio weights, i.e.:

$$\mathbf{w}_p^* = \Delta_{a^*R_p|b^*} \mathbf{w}_N + (1 - \Delta_{a^*R_p|b^*}) \mathbf{w}_V, \quad (15)$$

where $\Delta_{a^*R_p|b^*} = a^*E(R_p) - b^*$, with $a^* = ac/d$ and $b^* = a^2/d$, is a transform of the first Lagrangian coefficient of the global program in Eq. (4), which can be recognized as a Rescaled Fitted Performance Measure, comparing the targeted expected return of the portfolio to b^* , whilst V is the Global Minimum-Variance Portfolio and N the so-called Null Index Portfolio, their weights being respectively: $\mathbf{w}_N = \mathbf{E}^\circ/a = \check{\mathbf{E}}$ and $\mathbf{w}_V = \mathbf{1}^\circ/c = \check{\mathbf{1}}$.

Proof. See the Appendix.

In other words, Propositions 2 and 3 show as announced that there is a linear relation between the weight of any Frontier Portfolio (any feasible portfolio belonging to \mathcal{F} or $\overline{\mathcal{F}}$) and those of two pivotal vectors (here N and V), where relative weights in the two pivotal portfolios are comparisons with transforms of the targeted expected return to some level defined *per* scalars as Rescaled Fitted Performance Measure. Thus, we here invest a sum (equals to $\Delta_{a^*R_p|b^*}$) in the first pivotal portfolio and the complementary sum ($1 - \Delta_{a^*R_p|b^*}$) in the second one (respecting the budget constraint), and this depends upon the target expected return of portfolio p : $\bar{\mu}_p := E(R_p) = \hat{\mu}_p$, and some technical parameters (a, b, c, d and their transforms a^* and b^*) that define the market technology characteristics. The higher (the lower) the targeted return, the higher the sum to be put in the first portfolio (the higher the sum to be invested in the second portfolio, respectively) and *vice-versa*.

In the following, we will characterize all Efficient Portfolios thanks to some special notorious portfolios, as represented in the following Figure 1, which illustrates the Unconstrained Efficient Frontier, the Capital Market Line, the asymptotes of the Frontier (on $\overline{\mathcal{F}}$ and $\underline{\mathcal{F}}$, respectively) and some notorious portfolios such as, namely: the Global Minimum-Variance Portfolio (V), the Null Index Portfolio (N), the Tangent Portfolio (T), the Centre Portfolio¹⁷ (C),

¹⁷ with $E(R_C) = a/c$ and $\sigma(R_C) = 0$, and the two asymptotes follow a relation as such: $E(R_p) = a/c \pm \sqrt{d/c} \sigma(R_p)$ - see A-9 for details in the Appendix.

and the two special Portfolios¹⁸: (G) , which is, generally¹⁹ on $\underline{\mathcal{F}}$, as previously defined with a nil expected return, and (G') , which is a special Portfolio on the Efficient set $\overline{\mathcal{F}}$ with a 100% expected return.

Proposition 4.

The Mean-Variance Efficient Frontier is characterized by the set $\overline{\mathcal{F}}$ of the Minimum-Variance Efficient Portfolios \mathbf{p} , whose weights satisfy the following relation:

$$\mathbf{w}_p^* = \Delta_{a^*R_p|d^*}\mathbf{w}_N + \left(1 - \Delta_{a^*R_p|d^*}\right)\mathbf{w}_V, \quad (16)$$

and that have expected rates of return higher than of those of the Global Minimum-Variance Portfolio, denoted V , i.e.: $E(R_p) > E(R_V)$ where $E(R_V) = a/c$.

Proof. By definition of the Efficient Frontier.

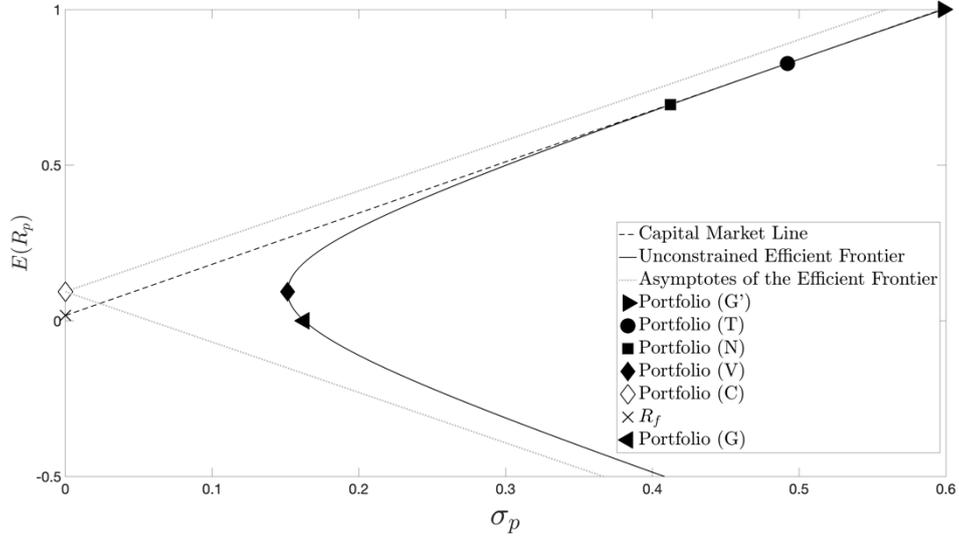
Indeed, the part of the Minimum-Variance Portfolio frontier that lies above the Global Minimum-Variance Portfolio is the Efficient Frontier (the upper limb of \mathcal{F} , denoted $\overline{\mathcal{F}}$). Portfolios that are on the lower part of the Minimum-Variance Portfolio frontier (the lower limb of \mathcal{F} , denoted $\underline{\mathcal{F}}$), or inside the bullet drawn by the Minimum-Variance frontier in the N -asset case, are called inefficient (or dominated) portfolios. For each inefficient portfolio, there are corresponding efficient ones, having the same variance (expected

¹⁸ Similarly, Black (1972) proves that every efficient portfolio can be generated by two arbitrary portfolios u and v , with $\beta_u = 1$ and $\beta_v = 0$, where β_u and β_v are measuring the volatility of an individual stock compared to the systematic risk

¹⁹ Or, exceptionally, on $\overline{\mathcal{F}}$, if the GMVP expected return is negative.

return, respectively), but a higher expected rate of return by expansion (or lower variance by contraction, respectively), positioned directly above it, in the north direction (or on its left, in the west direction, respectively).

Figure 1: An Illustration of Mean-Variance Efficient Portfolios and Some Notorious Portfolios



Source: Wind; daily quotes of the 40 largest companies that are part of the CSI300, from 08/18/2008 to 12/31/2019. Plotted are: the Unconstrained Efficient Frontier, the Capital Market Line (CML) and some notorious portfolios (V , N , T , C , G and G'). Computations by the authors.

3. The Efficient Frontier within a Regression-type Approach

Let us focus now on “regression-type approaches”, first by following the developments of Britten-Jones (1999)²⁰, and then by

²⁰ See also Broadie (1993), Kempf and Memmel (2006) and the Appendix for their approach, and Stevens (1998) who proves that elements of $\mathbf{\Omega}^{-1}$ are:

$$\mathbf{\Omega}^{-1}_{ij} = (-1)^o \times [\sigma_{ii}(1 - R_i^2)]^{-1} \beta_{ij},$$

where $\mathbf{\Omega}^{-1}_{ij}$ is the element on the i^{th} row and j^{th} column of $\mathbf{\Omega}^{-1}$, coefficient o is equal to 0 if $j = i$ and 1 otherwise, σ_{ii} the variance of returns on the i^{th} asset, β_{ij} the sensitivities (*i.e.* regression coefficients) for the regression of the return of the i^{th} asset on those of all other

recalling the principle of penalized regressions, before adapting them in the following section to the context of determining Efficient Portfolios with penalized regressions in Section 4.

Britten-Jones (1999) indeed propose a regression approach to find the weights of the Tangent Portfolio (T). He first defines a “highly desirable” portfolio - denoted (A), with a return of 1 (100%) and a volatility of 0 (0%). This is an Arbitrage Portfolio since it has zero risk (which is, nevertheless, unattainable in reality). The regression approach for portfolio selection is built on targeting the Portfolio (A), ultimately finding the closest Portfolio to this very advantageous Arbitrage Portfolio. It is therefore a question of minimizing the squared error terms between the expected excess returns over the risk-free rate of a constructed Efficient Portfolio and those of the Arbitrage Portfolio.

This minimization problem may be solved by Ordinary Least Squares (OLS) under some classical assumptions. Thus, we can reformulate the Markowitz problem in Eq. (4) in the form of a classical regression as such:

$$\mathbf{w}_p^* = \arg \min_{\mathbf{w}_p \in \mathbb{R}^N} \left\{ \left\| \mathbf{1}\bar{\mu}_p - \mathbf{X}\mathbf{w}_p \right\|_2^2 \right\}, \quad (17)$$

with $\|\cdot\|_2^2$ the squared ℓ_2 -norm²¹, \mathbf{X} the matrix of excess returns of assets: $\mathbf{X} = \mathbf{R} - R_f \mathbf{1}$, with \mathbf{R} being the ($T \times N$) asset returns matrix,

$N-1$ assets and R_t^2 the multiple regression coefficient (*i.e.* coefficient of determination) for the same regression. As expected, we find in the hereafter empirical studies that direct estimates of the inverse Variance-CoVariance matrix, denoted $\widehat{\mathbf{\Omega}}^{-1}$ (both in the unconstrained and constrained cases), correspond to the one defined according to the regression method by Stevens (1998) – See A-9 in the Appendix.

²¹ where, following the notations by Hastie *et al.* (2015), the ℓ_2 -norm is such as: $\left\| \mathbf{1}\bar{\mu}_p - \mathbf{X}\mathbf{w}_p \right\|_2 = \|\mathbf{u}\|_2 = \left(\sum_{t=[1, \dots, T]} |u_t|^2 \right)^{1/2}$, where u_t is the t -th element of the vector

R_f the risk-free rate and $\bar{\mu}_p$ the portfolio target return (feasible or not), subject to that the sum of the standardized weights is equal to one (same structure as the original portfolio obtained by OLS but with standardized weights with a unitary sum).

Such a previous regression in Eq. (17) gives the weights of an Efficient Portfolio based on a sample of returns. In the case of the Tangent Portfolio, for a target expected return corresponding to: $\bar{\mu}_p := E(R_T) = \hat{\mu}_T$, the sample mean, we can state the following proposition.

Proposition 5.

Defining the matrix $(T \times N)$ of the assets' excess returns \mathbf{X} , reading:

$$\mathbf{X} = (\mathbf{R}_1 - R_f \mathbf{1}_{(T \times 1)} \quad \mathbf{R}_2 - R_f \mathbf{1}_{(T \times 1)} \quad \dots \quad \mathbf{R}_j - R_f \mathbf{1}_{(T \times 1)} \quad \dots \quad \mathbf{R}_N - R_f \mathbf{1}_{(T \times 1)}),$$

where the \mathbf{R}_i , with $i = [1, \dots, N]$, are the $(T \times 1)$ vectors of return on each asset i , corresponding to i -th column of \mathbf{R} , the $(T \times N)$ asset returns matrix, the theoretical OLS regression (without constant) of a unit vector $\mathbf{1}_{(T \times 1)}$ on the matrix $(T \times N)$ of the assets' excess returns \mathbf{X} , is:

$$\mathbf{1}_{(T \times 1)} = \underset{(T \times N)(N \times 1)}{\mathbf{X}\boldsymbol{\beta}} + \underset{(T \times 1)}{\mathbf{u}}, \quad (18)$$

which allows us to obtain a vector of theoretical coefficients such as:

$$\boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{1}_{(T \times 1)}, \quad (19)$$

which leads to the vector of weights of the Tangent Portfolio (T):

$$\mathbf{w}_T = \frac{\boldsymbol{\beta}}{\mathbf{1}'\boldsymbol{\beta}} = \frac{\boldsymbol{\Omega}^{-1}\mathbf{E}_X}{\mathbf{1}'\boldsymbol{\Omega}^{-1}\mathbf{E}_X} = \check{\mathbf{E}}_X, \quad (20)$$

of residuals denoted $\mathbf{u} = (\mathbf{1}\bar{\mu}_p - \mathbf{X}\mathbf{w}_p)$, and with $\|\cdot\|_2^2 = (\|\cdot\|_2)^2 = \sum_{t=1, \dots, T} |\cdot|^2$ being the squared ℓ_2 -norm.

where $\mathbf{E}_X = \mathbf{X}'\mathbf{1}_{(T \times 1)}/T$ is the $(N \times 1)$ vector of means, $\mathbf{\Omega} = (\mathbf{X} - \mathbf{1}_{(T \times 1)}\mathbf{E}_X)'\mathbf{X} - \mathbf{1}_{(T \times 1)}\mathbf{E}_X$ the Variance-CoVariance matrix and $\mathbf{1}_{(T \times 1)}$ is the $(T \times 1)$ vector²² with all elements equal to 1.

Proof. See Britten-Jones (1999) and the Appendix.

In other words, going from theoretical quantities to sample ones, we can estimate the normalized vector of standardized coefficients ($\underline{\hat{\beta}} = \hat{\beta}/\mathbf{1}'\hat{\beta}$, as it sums at unity) following the above relationship in Proposition 2 and Corollary 2.3, which is thus recognized as the weight vector of the Tangent Portfolio (T), as such:

$$\mathbf{w}_T = \frac{\hat{\mathbf{\Omega}}^{-1}\hat{\mathbf{E}}_X}{\mathbf{1}'\hat{\mathbf{\Omega}}^{-1}\hat{\mathbf{E}}_X} = \overline{(\hat{\mathbf{E}}_X)}, \quad (21)$$

which is obtained as the solution of a quadratic optimization program, where the estimated sample mean $\hat{\mathbf{E}}_X$ and the estimated sample Variance-CoVariance matrix $\hat{\mathbf{\Omega}}$ (by empirical counterparts, OLS or *quasi*-maximum likelihood) are used as parameters, and finally leads to the Mahalanobis-type transform of the estimated mean excess return denoted $\overline{(\hat{\mathbf{E}}_X)}$.

Figure 2 illustrates some notorious portfolios and specifically the Arbitrage Portfolio (A), the Tangent Portfolio (T) and the

²² For the sake of clarification, we use for vector of 1 with dimensions $(N \times 1)$ with N the number of the assets: $\mathbf{1} := \mathbf{1}_{(N \times 1)}$ as a notation directly, whilst for a vector of 1 with dimensions $(T \times 1)$, with T the number of dates of the time-series of returns on an asset, we denote it as: $\mathbf{1}_{(T \times 1)}$. Furthermore, we use the notations: $\mathbf{1}_{(T \times 1)}$ for a $(T \times 1)$ vector of 1 when observations (dates) are at stake (instead of $\mathbf{1}$ as used by Britten-Jones, 1999, page 658).

Portfolio (BJ) obtained by following the Britten-Jones (1999) method²³.

This kind of regression is quite “unusual”. There is no constant, the dependent variable is not stochastic and the residual vector \mathbf{u} is correlated with the regressors which are stochastic. However, the regression has a simple interpretation: the dependent variable $\mathbf{1}$ is the counterpart to arbitrage profits on the sample returns - a positive excess return with zero risk; the coefficients $\hat{\boldsymbol{\beta}}$ therefore are the weights on the risky assets of the portfolio, whilst the term $\mathbf{X}\hat{\boldsymbol{\beta}}$ represents the excess returns on this portfolio, and the residual vector \mathbf{u} shows deviations from $\mathbf{1}$ of the returns on this portfolio.

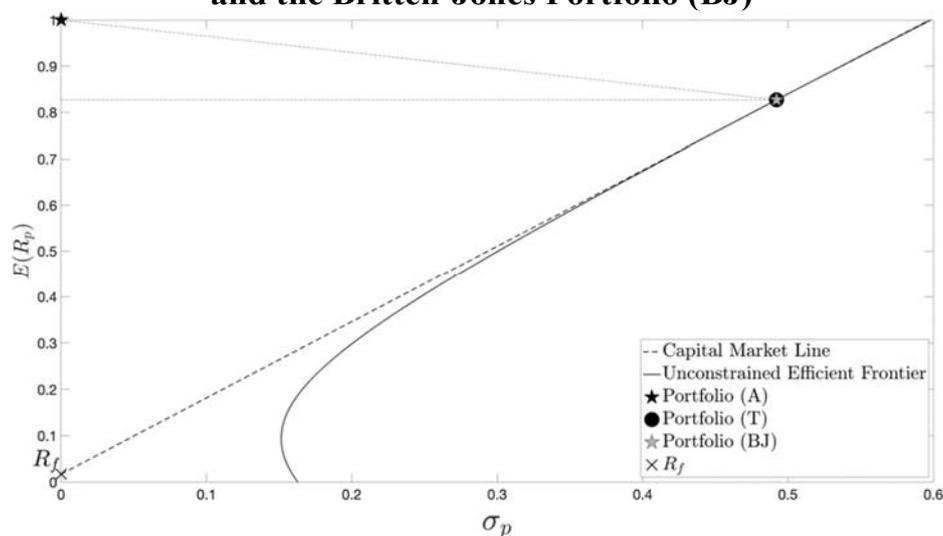
Nevertheless, the portfolio obtained by OLS does not respect the budget constraint. Only the standardized portfolio, denoted (BJ) Portfolio exactly confused with the Tangent Portfolio, with weights $\hat{\boldsymbol{\beta}}$, satisfies this constraint.

In brief, the estimated weights $\hat{\boldsymbol{\beta}}$ are linked to a vector of returns on the related portfolio, whilst also satisfying the budget constraint, which is the closest (in terms of OLS distance) to the one of the Arbitrage Portfolio (A).

This least square distance can be illustrated using the familiar mean- standard deviation scheme. The feasible set, constructed from the sample mean and the sample covariance, has a limit represented by the line (0, BJ) from the origin through the tangent portfolio (T) (Cf. Figure 2). The Arbitrage Portfolio (A) is at the coordinates (0, 1).

²³ Or, alternatively, by Stevens (1998) or Kempf and Memmel (2006) methods with same results.

**Figure 2: An Illustration of
Mean-Variance Efficient Portfolios
and the Britten-Jones Portfolio (BJ)**



Source: Wind; daily quotes of the 40 largest companies that are part of the CSI300, from 08/18/2008 to 12/31/2019. Plotted are: the Unconstrained Efficient Frontier, the Capital Market Line (CML) and some notorious portfolios (*A*, *T*, and *BJ*). Computations by the authors.

Therefore, the OLS approach²⁴, by minimizing the square root of the residuals, allows us to find a portfolio whose returns are located as close as possible to the point (0,1) in the plane (risk-return). Such a portfolio is illustrated in Figure 2 by a grey star indicating the point (*BJ*). Thus, by varying the return target (achievable or not as in the case of the Britten-Jones Arbitrage Portfolio), we can easily find all the Efficient Portfolios (with the notable exception of the Global Minimum-Variance Portfolio²⁵).

²⁴ See Appendix for more details on the OLS approach.

²⁵ See Appendix on the Kempf and Memmel (2006) approach in the specific case of the GMVP.

Let us examine next the introduction of constraints into the global program in Eq. (4).

4. From Normal Unconstrained Regressions to Penalized Regressions for Constraining the Efficient Frontier Weights

Many constraints have been added to the original programs in the literature. Among the first constraints on weights, we can cite as previously the budget constraint, but also the constraint of positivity on each of the weights (only positive investments in the assets), the constraint of a maximum weights on each of the weights (limiting the maximum investment on one specific asset), as well as the cardinality of the weights (purchase of an exact number of assets).

In general, penalized regressions have been proposed to solve several types of problems, the first being the “selection” of variables in a regression (*Cf.* Tibshirani, 1996; Hastie *et al.*, 2015). In our context, the addition of a penalty on the coefficients of the regression, according to the technique used, makes it possible to impose constraints on weights of the Efficient Portfolios - see Brodie *et al.* (2009), DeMiguel *et al.* (2009a) and (2009b), Candelon *et al.* (2012).

The global program can be rewritten as such (with previous notations):

$$\mathbf{w}_p^* = \arg \min_{\mathbf{w}_p \in \mathbb{R}^N} \left\{ \|\mathbf{1}\bar{\mu}_p - \mathbf{X}\mathbf{w}_p\|_2^2 + \gamma P(\mathbf{w}_p) \right\} \quad (22)$$

$$s.t. \begin{cases} \mathbf{A}_1 \mathbf{w}_p = \mathbf{b}_1 \\ \mathbf{A}_2 \mathbf{w}_p = \mathbf{b}_2 \\ \mathbf{C}_1 \mathbf{w}_p \leq \mathbf{d}_1 \\ \mathbf{C}_2 \mathbf{w}_p \leq \mathbf{d}_2, \end{cases}$$

where γ a fine-tuning constant and:

$$\begin{cases} \mathbf{A}_1 = \mathbf{E}' \text{ and } \mathbf{b}_1 = \bar{\mu} \\ \mathbf{A}_2 = \mathbf{1}' \text{ and } \mathbf{b}_2 = 1 \\ \mathbf{C}_1 = -1 \times \mathbf{I}_{(N \times N)} \\ \mathbf{C}_2 = \mathbf{I}_{(N \times N)}, \end{cases}$$

with $\mathbf{I}_{(N \times N)}$ the $(N \times N)$ Identity matrix, \mathbf{C}_1 a $(N \times N)$ matrix of coefficients for fixing some lower constraints on the weights of any single asset as expressed in the $(N \times 1)$ vector \mathbf{d}_1 , \mathbf{C}_2 an $(N \times N)$ matrix of coefficients for fixing some upper limits on the weights as defined in the $(N \times 1)$ vector \mathbf{d}_2 , and where the penalization is finally defined such as:

$$P(\mathbf{w}_p) \equiv \begin{cases} \|\mathbf{w}_p\|_1 = \mathbf{1}'|\mathbf{w}_p| = \sum_{i=1}^N |w_i| & \text{for LASSO} \\ \|\mathbf{w}_p\|_2^2 = \sum_{i=1}^N |w_i|^2 & \text{for RIDGE} \\ (1 - \alpha)\|\mathbf{w}_p\|_2^2 + \alpha\|\mathbf{w}_p\|_1 & \text{for Elastic Net,} \end{cases}$$

where the well-known penalized regressions mentioned in the literature are the Least Absolute Shrinkage and Selection Operator (LASSO) as expressed in Eq. (22), as well as the Regularization of Inverse Discrete Gradient Estimator (RIDGE) and the Elastic Net (see, respectively, Tibshirani, 1996, for LASSO, Hoerl and Kennard, 1970, for RIDGE and Zou and Hastie, 2005, for Elastic Net, and see also Hastie *et al.*, 2015)²⁶.

²⁶ The ℓ_2 -norm penalization term is used in the Regularization of Inverse Discrete Gradient Estimator (RIDGE), such as (see Hoerl and Kennard, 1970, page 56): $P(\mathbf{w}_p) = \|\mathbf{w}_p\|_2^2 = \sum_{i=1}^N |w_i|^2$ uses the squared ℓ_2 -norm, and both the ℓ_1 -norm and the squared ℓ_2 -norm are used in Elastic Net regularization such as (see Zou and Hastie, 2005, page 304): $P(\mathbf{w}_p) = (1 - \alpha)\|\mathbf{w}_p\|_2^2 + \alpha\|\mathbf{w}_p\|_1$ where $\alpha \in [0,1]$. See Hastie *et al.* (2015), p.22 and p.57 and also Kremer *et al.* (2020), p.8 for geometric illustrations of the effects of such norms on coefficients.

If we take the previous approach by now using the simplest penalized regression (*Cf., e.g., Brodie et al., 2009*), the program resumes to:

$$\mathbf{w}_p^* = \arg \min_{\mathbf{w}_p \in \mathbb{R}^N} \left\{ \|\mathbf{1}\bar{\mu}_p - \mathbf{X}\mathbf{w}_p\|_2^2 + \gamma P(\mathbf{w}_p) \right\}, \quad (23)$$

with $\|\cdot\|_2^2$ the squared ℓ_2 -norm²⁷, \mathbf{X} the matrix of excess returns of an asset: $\mathbf{X} = \mathbf{R} - R_f \mathbf{1}$ and $\bar{\mu}_p$ the portfolio target return (feasible or not), γ a constant, $P(\cdot)$ the penalization on the weights (which will take different forms according to the technique used), written as:

$$P(\mathbf{w}_p) \equiv \|\mathbf{w}_p\|_1 = \mathbf{1}'|\mathbf{w}_p|, \quad (24)$$

with²⁸ $\|\cdot\|_1$ the ℓ_1 -norm.

Using some variants of the same approach, the optimization program can again be rewritten under different kinds of constraints according to the relevant program. For example, if we follow Jagannathan and Ma (2003), we may be willing for various reasons to restrict the maximum weight of each asset that should not exceed some limits²⁹ (let us say inferior to a limit l_2 hereafter for instance). If we further add the positivity constraint on the weights, we then can write the previous program with both constraints (for maximum and positive weight), fixed as such:

²⁷ with $\|\cdot\|_q^k$ denoting the power k of the ℓ_q -norm operator, such as: $\|x_i\|_q^k = (\|x_i\|_q)^k = (\sum_{i \in \mathbb{N}} |x_i|^q)^{k/q}$. As mentioned before, we use in this equation the squared ℓ_2 -norm as: $\|\mathbf{1}\bar{\mu} - \mathbf{X}\mathbf{w}_p\|_2^2 = \|\mathbf{u}\|_2^2 = (\sum_{t=[1, \dots, T]} |u_t|^2)$, where u_t is the t -th element of the vector of residuals denoted $\mathbf{u} = (\mathbf{1}\bar{\mu} - \mathbf{X}\mathbf{w}_p)$ – see Eq. (22) for instance.

²⁸ As mentioned previously, the ℓ_1 -norm penalization term in LASSO regression writes in Eq. (22) as $P(\mathbf{w}_p) = \|\mathbf{w}_p\|_1 = \sum_{i=1}^N |w_i|$ as in Brodie *et al.* (2009), p.12,268.

²⁹ In order to avoid concentrated and time-varying portfolios, or to deal with liquidity problem as in Vieira and Filomena (2020 – see Section 2.2.2, page 1060), or limit transaction costs (*e.g., Lobo et al., 2007*) or impose cardinality constraints (*e.g., Chang et al., 2000; Anagnostopoulos and Mamanis, 2011; Woodside-Oriakhi et al., 2011*).

$$\begin{cases} \mathbf{d}_1 = \mathbf{0} \\ \mathbf{d}_2 = l_2 \times \mathbf{1}, \end{cases}$$

with $\mathbf{0}$ the $(N \times 1)$ null matrix (for a unique lower bound), $\mathbf{1}$ a $(N \times 1)$ vector of 1, and as scalar l_2 is the unique upper limit on weights of all single assets.

Due to the addition of a penalty in the objective function of regressions in Eq. (22) and (23), the vector of coefficients of these regressions cannot be estimated by a conventional approach like OLS because of the resulting bias (on the estimators) inherent in the regression, due to the addition of this constraint. Since the objective function of a penalized regression is generally not differentiable (for example, the objective function of the LASSO is not differentiable at any point, because the ℓ_1 -norm on the coefficients of the regression is not differentiable in 0, unlike the objective function using the ℓ_2 -norm in the penalty of the RIDGE regression), different algorithms have been developed in order to find solutions.

Among these algorithms, one notably finds the Least-Angle Regression algorithm (LARS) allowing for the computation of the LASSO (see Zou and Hastie, 2005) or its extension called Least-Angle Regression Elastic Net (LARSEN), which also permits us to solve other penalties such as the RIDGE or Elastic Net type (see Hastie *et al.*, 2015). Gaines *et al.* (2018) compare different algorithms and provide practical improvements to utilize Constrained LASSO regression on various sizes of databases, which also helps to improve our regression program.

In our case, this will be mobilized to impose a positivity constraint on optimal portfolio weights. Thus, the use of the penalized regressions makes it possible, in our case, to be as close as

possible to the optimal portfolio without the constraint of positivity, by determining a portfolio composed only of long positions, and taking into account a positivity constraint on the weights.

Figure 3 below shows the following five frontiers:

- 1) The unconstrained one: i) without any other constraint than the budget constraint, obtained from the combination of the two optimal Efficient Portfolios (N) and (V); ii) an identical frontier as we have shown to be the one obtained by numerical application of the Markowitz quadratic program resolution (*e.g.* Turlach and Wright, 2015) from the use of the numerical resolution of the quadratic program; iii) identical also to the one obtained by applying the regression approach without penalization proposed by Britten-Jones (1999) and completed by Kempf and Memmel (2006);
- 2) A constrained one: iv) obtained by the LASSO penalized regression under usual budget constraints by varying the targets of expected return, with a light penalization (*i.e.*, corresponding to a low coefficient³⁰ of penalization γ in Eq. (22));
- 3) Another constrained one: v) obtained also, as in previous case iv) by the LASSO penalized regression with a stronger penalization (*i.e.*, corresponding to a higher coefficient³¹ of penalization γ in Eq. (22), to be compared to the previous efficient frontier in iv);
- 4) A positively constrained one: vi) obtained by the LASSO penalized regression under budgetary constraint and the

³⁰ equals here to 4,000 in our restricted Chinese illustration sample.

³¹ equals here to 7,800 in our illustrative sample.

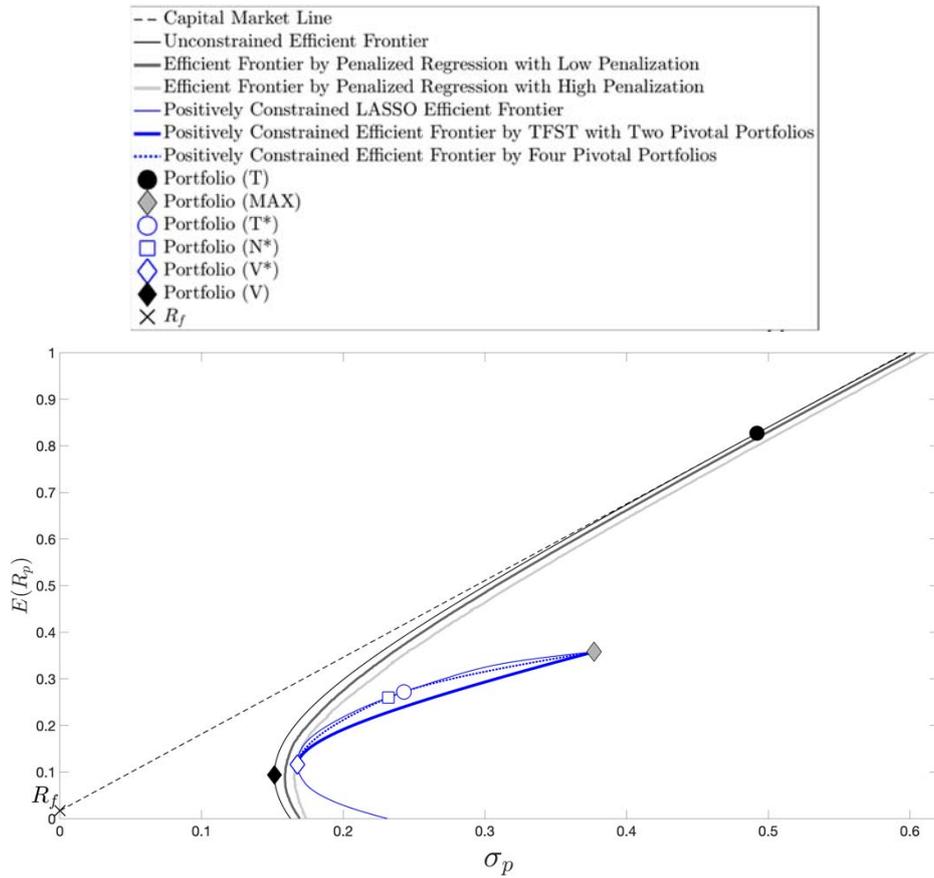
constraint of positivity on weights, by varying the targets of expected return;

- 5) Another positively constrained: vii) obtained by mixing different proportions of the Positively Constrained GMVP weights and the vector of weight having 100% on the maximum return security (0% in any other asset), utilizing the Two-Fund Separation Theorem; viii) obtained by mixing with different proportions of the constrained GMVP, the Null Index Portfolio, the Tangent Portfolio and the vector of weights having 100% on the maximum return security (and 0% in any other asset), utilizing a generalization to four fund of the Two-Fund Separation Theorem.

As expected, although the general hyperbolic form is conserved, the effect of the positivity constraint on the weights has a significant “contracting” effect on the unconstrained frontier: it is no longer possible to obtain such high return levels, which is the price to pay for a more realistic representation of the efficient frontier.

Furthermore, we again here show that Two-Fund Separation Theorem perfectly works when there is no other constraint but the usual budget one. In other words, any Efficient Portfolio can be expressed as a linear combination of two other Efficient Portfolios with no other constraint.

Figure 3: An Illustration of Mean-Variance Efficient Portfolios with and without a Positive Constraint on the Weights



Source: Wind; daily quotes of the 40 largest companies that are part of the CSI300, from 08/18/2008 to 12/31/2019. Plotted are: the Unconstrained Efficient Frontier, the Efficient Frontier by Penalized Regression with low and high penalization, the Positively Constrained LASSO Efficient Frontier, the Positively Constrained Efficient Frontier by TFST, the Capital Market Line (CML) and some notorious portfolios (T , MAX , N and V), where a portfolio distinguished by a superscript “*” denotes the corresponding portfolio under positivity constraint. TFST corresponds to Two-Fund Separation Theorem, which is to build Efficient Frontiers by the combination of two pivotal portfolios: the Positively Constrained Global Minimum-Variance Portfolio (V) and the Asset with the Highest Expected Return or four pivotal portfolios: the Positively Constrained Global Minimum-Variance Portfolio, the Positively Constrained Null Index Portfolio, the Positively Constrained Tangent Portfolio and the Asset with the Highest Expected Return. Computations by the authors.

Now, if we compare the Positively Constrained Efficient Frontier and the set of linear combinations of the constrained GMVP weights and the asset with the maximum return as shown by efficient

frontier vi) and vii), it is clear in Figure 3 that the two frontiers are quite close and similar, whilst not identical. Compared with the traditional Two-Fund Separation Theorem, which works perfectly with no further constraint, such a theorem under positivity constraint holds only approximately, when using machine learning techniques to impose some realistic limitations on weights (*e.g.*, sparsity and positivity).

Let us now proceed in the next section to some robustness checks of previously highlighted results and resume them.

5. Main Results and Robustness Checks

To investigate the difference between the empirical Efficient Frontier we recover from the Positively Constrained LASSO regression and the Efficient Frontier by the Two-Fund Separation Theorem (TFST for short), we next produce some robustness checks.

Previously displayed results were presented on a restricted database of Chinese stocks. Hereafter, we first check whether we can exhibit the same type of qualitative results with large databases (for large-scale portfolios), across three continents (the Chinese, US and French stock markets) to evaluate if our results still hold when these hypotheses differ from the previous simplified ones.

Figure 4 illustrates the Unconstrained Efficient Frontier, the Positively Constrained Efficient Frontier by TFST with two Pivotal Portfolios mentioned above (V and MAX) and another one generated with four Pivotal Portfolios (V , N , T and MAX), as well as the whole gamut of notorious portfolios explored in this article.

It appears that the same results are obtained: the approximation obtained with the TFST Efficient Frontier, compared to the

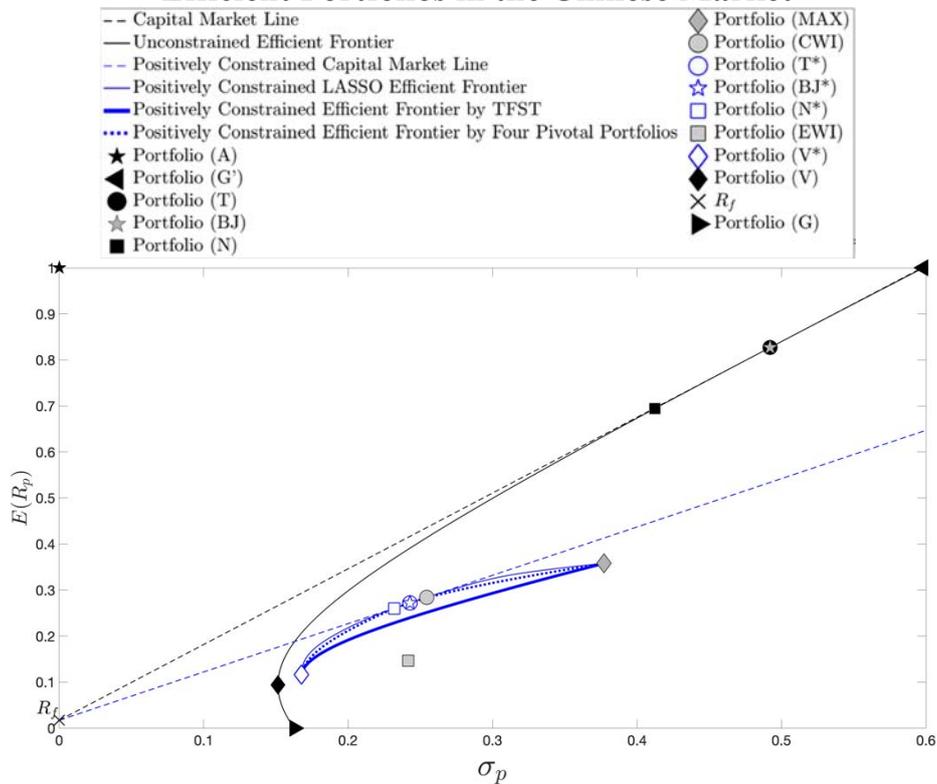
Constrained LASSO one, is rather accurate, in the Chinese market (Figure 4 - as well as in the Chinese, US market and the French market with restricted samples and large datasets – see Appendix). Furthermore, under the basic assumptions of the CAPM, the Tangent Portfolio should be the Market Portfolio (which is the Capitalization-Weighted Index) according to theory (see, among others, Sharpe, 1964). This model of course may not hold exactly for many various reasons (see, *e.g.*, Roll, 1977). However, we precisely note that, across three continents (Chinese, US and French stock markets), estimations illustrate that CAPM theory does not meet the stylized facts, and that positivism differs from deductivism, since at *equilibrium*: 1) the Market Portfolio, as represented either by the Capitalization-Weighted Index (CWI: resp. CSI300, S&P500 and SBF120 Index) or by the Equally-Weighted Index (EWI), is not on the (LASSO) Efficient Frontier; 2) these indices are not close to the Tangent Portfolio (T) on all the three markets. We can also note that the Britten-Jones Portfolio (BJ) corresponds to the Tangent Portfolio (T), both for the unconstrained programs of course, but also for positively constrained ones. The same for the GMVP: the Kempf and Memmel (2006) Portfolios and those of Stevens (1998) correspond to the LASSO-estimated GMVP, both for the unconstrained programs for sure, but also for positively constrained ones.

We may also want to know if the various portfolios from the Constrained LASSO Efficient Frontiers and the TFST implied ones, for each level of volatility, exhibit different characteristics in terms of expected returns, estimated risks and reward-to-risk ratios, or not.

One way of doing this is to compare their Sharpe ratios (see Jobson and Korkie, 1981; Memmel, 2003; Ledoit and Wolf, 2008;

Ledoit and Wolf, 2011). Following the methodology proposed by Ledoit and Wolf (2011), we compute Heteroskedasticity and Autocorrelation Consistent (HAC) robust standard error estimators for the difference of the estimated variances by the method of Andrews (1991) and Andrews and Monahan (1992), when computing the various Sharpe ratios (see also Israelsen, 2003 and 2005; Lo, 2002; Oppdyke, 2007; Vinod and Morey, 1999a and 1999b).

Figure 4: An Illustration of Empirical Mean-Variance Efficient Portfolios in the Chinese Market



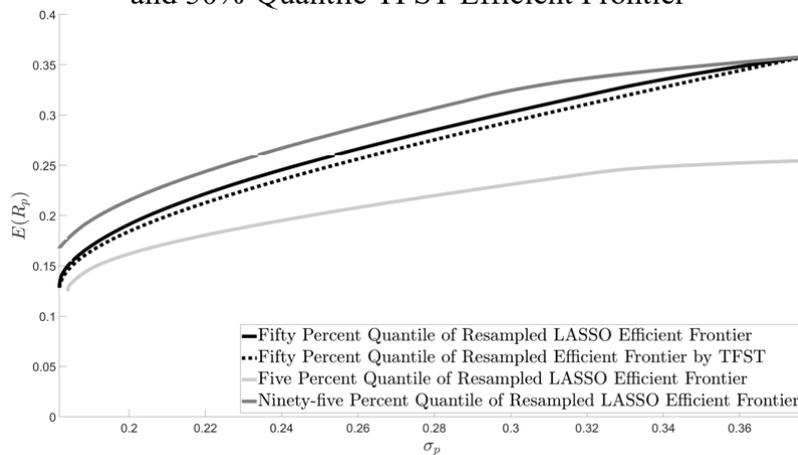
Source: Wind; daily quotes of the 40 largest companies that are part of the CSI300, from 08/18/2008 to 12/31/2019. Plotted are: the Capital Market Line (CML), the Unconstrained Efficient Frontier, the Positively Constrained LASSO Efficient Frontier, the Positively Constrained Efficient Frontier by TFST with Two Pivotal Portfolios, the Positively Constrained Efficient Frontier with Four Pivotal Portfolios and some notorious portfolios (A , MAX , V , T , G ,

G' , N , BJ , EWI , and CWI), where EWI is an Equally-Weighted Index Portfolio and CWI is a Capitalization-Weighted Index Portfolio. A portfolio distinguished by a superscript “*” denotes the corresponding portfolio under positivity constraint. TFST corresponds to Two-Fund Separation Theorem, which is to build Efficient Frontiers by the combination of two pivotal portfolios: the Positively Constrained Global Minimum-Variance Portfolio and the Asset with the Highest Expected Return (in the blue bold line) or four pivotal portfolios (in the blue dotted bold line): the Positively Constrained Global Minimum-Variance Portfolio, the Positively Constrained Null Index Portfolio, the Positively Constrained Tangent Portfolio and the Asset with the Highest Expected Return. Computations by the authors.

We reach the same main result: there are no significant differences in Sharpe ratios between portfolios induced by Positively Constrained LASSO and TFST, at all levels of volatility.

Furthermore, when trying to consider model risk, using Michaud (1989) and a Michaud and Michaud (2008) type of method as well as the Ledoit and Wolf (2008, 2011) methodology, we confirm no significant difference in Sharpe ratios (see also the Appendix - Figure C-6, C-7 and C-8).

Figure 5: An Illustration of Double-resampled (Dates and Stock Sub-samples) Constrained Mean-Variance Efficient Portfolios in the Chinese Market
The 5%-, 50%- and 95%-Quantile of Double Block-bootstrapped Estimated LASSO Efficient Frontier and 50%-Quantile TFST Efficient Frontier



Source: Wind; daily quotes of the 40 largest companies that are part of the CSI300, from 08/18/2008 to 12/31/2019. Plotted are: the 5%-, 50%- and 95%-quantile of the double block-bootstrapped resampled LASSO Efficient Frontier and the corresponding 50%-quantile of the Efficient Frontier by TFST. Resampled Efficient Frontiers are computed by using a double-block bootstrap method (see Michaud, 1989; Michaud and Michaud, 2008 and Ledoit and Wolf, 2008 and 2011) both on dates and sub-samples of stocks. TFST corresponds to Two-Fund Separation Theorem, which is used to build the Efficient Frontier by combinations of two pivotal portfolios, namely the Global Minimum-Variance Portfolio and the Asset with the Highest Expected Return. Computations by the authors.

And so it happens that the TFST Efficient Portfolios all lie in between the 5%- and 95%-confidence bounds for all volatilities and there are no significant differences in Sharpe ratios between Portfolios in competition (Constrained LASSO *versus* TFST ones with the same level of volatility) for almost all levels of volatility, in the three markets (Chinese: see Figure 5 above, as well as in the Chinese, US, and French markets with limited samples of 40 stocks and with the big datasets of the total sample of available asset series over the entire period – see Appendix).

Finally, we hereafter resume with the main results obtained from the three sets of databases, trying to distinguish the main stylized facts.

First, when we compare outputs using the full datasets to those with the corresponding datasets but only including the assets with positive excess expected returns, we find that the Efficient Frontiers and Pivotal Portfolios are just only slightly different in terms of performance and risk (and, obviously, there is no difference when the positivity constraint is considered, since assets with negative excess returns are just discarded). Thus, this might offer us an opportunity to further reduce the computation-time of our approach, when only pre-selecting the “good assets”.

Secondly, when we focus on results obtained when there is no positivity constraint, it is clear that the pivotal Portfolio (V), Portfolio (N) and Portfolio (T) are always on the Efficient part of the envelope $\overline{\mathcal{F}}$. Moreover, Portfolio (T) can always be targeted by the Britten-Jones (1999) regression approach, *i.e.*, Portfolio (BJ) is always equivalent to Portfolio (T), whilst Stevens (1998) always helps to obtain the Efficient Frontiers, and Portfolio (V) can always be targeted by the regression approach by Kempf and Memmel (2006). Furthermore, Portfolio (N) is always close³² to Portfolio (T) and both Portfolios are always positioned above the risk-free asset (*i.e.*, with a higher expected return). The expected return on Portfolio (V) is in most cases positive (but not always), and sometimes rather high (*e.g.*, with high-dimensional datasets on the Chinese market, see Figure C-9 in the Appendix). Finally, we note that the volatility of Portfolio (G') is always higher than the one of Portfolio (G) as intuited.

Thirdly, when we now impose a positivity constraint, we find that the pivotal Portfolio (V), Portfolio (N) and Portfolio (T) are on the constrained version of efficient frontier $\overline{\mathcal{F}}^*$, as when there is no constraint. Portfolio (T) and (V) can also be, once again, computed through Stevens (1998), Britten-Jones (1999) and Kempf and Memmel (2006) approaches as in the previous non-constrained case. Portfolio (N) and (T) are still close and the expected returns on both portfolios are higher than R_f . The expected return of the constrained Portfolio (V) is as well always positive in every subset of all our tests.

³² This relies on the expected return of the risk-free asset, where we find that the distance between Portfolio (N) and (T) in the Chinese and US market (with a higher return on the risk-free asset) is significantly higher than that in the French market (with a very low expected return of the risk-free asset).

In the Chinese and US markets, the Capitalized-Weighted Index (CWI) is on $\overline{\mathcal{F}}^*$, that is, in other words, the CWI is optimal in the Chinese and US market³³. Furthermore, in our static framework, it appears unfortunately that the Equally-Weighted Index (EWI) is not optimal in every one of our tests (market and periods).

Finally, we note that the TFST, with or without constraint, gives very similar Efficient Frontiers as we intuited at the beginning of this work.

Conclusion

In this article, we revisit a simple technique for the Markowitz Mean-Variance portfolio optimization that promotes a higher level of computational efficiency and flexibility in comparison with predominantly used methods.

More precisely, we demonstrate that the weights of any Efficient Portfolio are a linear combination of the weights of two special pivotal Efficient Portfolios, which are non-linear transforms (*i.e.*, using rescaled Mahalanobis-type norms) of simple weight vectors.

Knowing that the Two-Fund Separation Theorem holds exactly when no constraints are considered, we show, using some recent machine learning techniques, that the Two-Fund Separation Theorem is only approximative but accurate, under specific realistic constraints. In this context, we show the Positive Constrained (LASSO) Efficient Frontier can be generated thanks to combinations of two pivotal portfolios. Furthermore, we show that the Positively

³³ whilst the same result does not hold in all markets: the Capitalized-Weighted Index in the French market is not optimal for instance.

Constrained (LASSO) Efficient Frontier can be fairly and better approximated (with almost no noticeable differences from a rough eye-ball analysis) by the combination of four pivotal portfolios, namely, the Positively Constrained Global Minimum-Variance Portfolio, the weight of the Positively Constrained Null Index, the weight of the Positively Constrained Tangent Portfolio, and a vector of scalars whose sum respects the non-negative budget constraint.

The proposed approach shows several advantages, which include: i) computational simplicity; ii) time-efficiency; and iii) explanatory power. Moreover, the implemented methodology does not require providing the initial weights in the portfolio optimization process. We believe that the presented approach can be of interest to researchers, academics and students involved in practically oriented portfolio optimization tasks since it represents a significant computing-time advantage.³⁴ Finally, it allows to include several realistic constraints for large-scale portfolios.

Several extensions of this article could be considered. One natural direction would be, following the works by Yang (2000, 2001, 2004), Duchin and Levy (2009), Tu and Zhou (2011), Schanbacher (2015), Bonaccolto *et al.* (2018) and Joo and Park (2021), to study furthermore the validity of the Two-Fund Separation Theorem (TFST) in a (conditional) dynamic portfolio management out-of-sample exercise, that might lead to a Generalized K-Fund Separation Theorem (as in Dybvig and Liu, 2018 and Deguest *et al.*,

³⁴ For instance, the computation of the whole Efficient Frontier on the high-dimensional data in the Chinese stock market with the Positively Constrained LASSO technique takes approximately 35 hours on a 64 cores machine, whilst it takes less than a minute with a one core-computer when using the extended 4-fund TFST (with a related efficiency ratio of approximately 130,000).

2018), relying as in Bonaccolto and Paterlini (2020) on an empirical multi-strategy that optimally combines various pivotal portfolios related to sophisticated strategies, *i.e.*: Markowitz (1952), Jorion (1986), MacKinlay and Pástor (2000); Kan and Zhou (2007) and Tu and Zhou (2011) dealing with the Equally-Weighted Portfolio (see also DeMiguel *et al.*, 2009b), ultimately targeting a risk reduction and an efficiency increase as in Zhao *et al.* (2020), in a large portfolio context following Belloni and Chernozhukov (2011). These main complementary studies are now put into our future research agenda.

Acknowledgments

We thank Philippe Bernard, François Quittard-Pinon for previous collaborations on this topic, as well as Jean-Charles Garibal for preliminary research assistance, and Jean-Luc Prigent and Thierry Roncalli for their positive suggestions when writing first drafts of this article. We also sincerely appreciate the constructive discussion by Dionisis Philippas during the FEM2021 conference (Paris, June 2021). We are also grateful to the participants of the XXXVth AFFI Conference, of the XXXVth JMA conference, as well as Referees of the IFC2020, the ISFBI2021 and the FEM2021 conferences, as well as the Editor of Annals of Operations Research in charge and the two anonymous Referees for their fair comments and suggestions. Resources linked to this article are available on: www.performance-metrics.eu. The usual disclaimer applies.

References

- Abadie, J. (1969). Generalization of the Wolfe reduced gradient method to the case of nonlinear constraints. *Optimization*, 37-47.
- Anagnostopoulos, K., & Mamanis, G. (2011). The mean-variance cardinality constrained portfolio optimization problem: An experimental evaluation of five multi-objective evolutionary algorithms. *Expert Systems with Applications*, 38(11), 14208–14217.

-
- Andrews, D. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59, 817-858.
- Andrews, D., & Monahan, J. (1992). An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica*, 60, 953-966.
- Bai, L., Newsom, P., & Zhang, J. (2011). Teaching utility theory with an application in modern portfolio optimization. *Decision Sciences Journal of Innovative Education*, 9(1), 107-112.
- Bawa, V., Brown, S., & Klein, R. (1979). Estimation risk and optimal portfolio choice. North Holland, Amsterdam.
- Belloni, A., & Chernozhukov, V. (2011). ℓ_1 -penalized quantile regression in high-dimensional sparse models. *The Annals of Statistics*, 39(1), 82-130.
- Bernard, C., De Vecchi, C., & Vanduffel, S. (2021). When do two-or three-fund separation theorems hold? *Quantitative Finance*, 1-15.
- Black, F. (1972). Capital market equilibrium with restricted borrowing. *The Journal of Business*, 45(3), 444-455.
- Black, F., & Litterman, R. (1992). Global portfolio optimization. *Financial Analysts Journal*, 48(5), 28-43.
- Bodie, Z., Kane, A., & Marcus, A. (2012). *Essentials of investments*, 9th Edition. McGraw-Hill.
- Bonaccolto G., Caporin, M., & Maillet, B. (2021). Large financial network *via* conditional expected shortfall. Mimeo. R&R in the *European Journal of Operational Research*.
- Bonaccolto, G., Caporin, M., & Paterlini, S. (2018). Asset allocation strategies based on penalized quantile regression. *Computational Management Science*, 15(1), 1-32.
- Bonaccolto, G., & Paterlini, S. (2020). Developing new portfolio strategies by aggregation. *Annals of Operations Research*, 292(2), 933-971.
- Breuer, T., & Csiszár, I. (2013). Systematic stress tests with entropic plausibility constraints. *Journal of Banking & Finance*, 37(5), 1552-1559.
- Breuer, T., Jandacka, M., Rheinberger, K., & Summer, M. (2009). How to find plausible, severe, and useful stress scenarios. *International Journal of Central Banking*, 5, 205-224.
- Britten-Jones, M. (1999). The sampling error in estimates of mean-variance efficient portfolio weights. *The Journal of Finance*, 54(2), 655-671.
- Broadie, M. (1993). Computing efficient frontiers using estimated parameters. *Annals of Operations Research*, 45(1), 21-58.

-
- Brodie, J., Daubechies, I., De Mol, C., Giannone, D., & Loris, I. (2009). Sparse and stable Markowitz portfolios. *Proceedings of the National Academic Science of the USA*, 106(30), 12267-12272.
- Brown, S. (1976). Optimal portfolio choice under uncertainty: A Bayesian approach. Ph. D. Dissertation, Graduate School of Business, University of Chicago.
- Cairns, A. J., Blake, D., & Dowd, K. (2006). Stochastic lifestyle: Optimal dynamic asset allocation for defined contribution pension plans. *Journal of Economic Dynamics and Control*, 30(5), 843-877.
- Candelon, B., Hurlin, C., & Tokpavi, S. (2012). Sampling error and double shrinkage estimation of minimum variance portfolios. *Journal of Empirical Finance*, 19(4), 511-527.
- Cass, D., & Stiglitz, J. (1970). The structure of investor preferences and asset returns, and separability in portfolio allocation: A contribution to the pure theory of mutual funds. *Journal of Economic Theory*, 2(2), 122-160.
- Chang, T., Meade, N., Beasley, J., & Sharaiha, Y. (2000). Heuristics for cardinality constrained portfolio optimisation. *Computers & Operations Research*, 27(13), 1271-1302.
- Dahlquist, M., Farago, A., & Tédongap, R. (2017). Asymmetries and portfolio choice. *The Review of Financial Studies*, 30(2), 667-702.
- Deguest, R., Martellini, L., & Milhau, V. (2018). A reinterpretation of the optimal demand for risky assets in fund separation theorems. *Management Science*, 64(9), 4333-4347.
- De Meo, E. (2021). Scenario design for macro-financial stress testing. Available at SSRN 3493554.
- DeMiguel, V., Garlappi, L., Nogales, F., & Uppal, R. (2009a). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science*, 55(5), 798-812.
- DeMiguel, V., Garlappi, L., & Uppal, R. (2009b). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *The Review of Financial Studies*, 22(5), 1915-1953.
- DeMiguel, V., Martin-Utrera, A., & Nogales, F. (2013). Size matters: Optimal calibration of shrinkage estimators for portfolio selection. *Journal of Banking & Finance*, 37(8), 3018-3034.
- DeMiguel, V., Nogales, F., & Uppal, R. (2014). Stock return serial dependence and out-of-sample portfolio performance. *The Review of Financial Studies*, 27(4), 1031-1073.

-
- Duchin, R., & Levy, H. (2009). Markowitz versus the Talmudic portfolio diversification strategies. *The Journal of Portfolio Management*, 35(2), 71-74.
- Dybvig, P., & Liu, F. (2018). On investor preferences and mutual fund separation. *Journal of Economic Theory*, 174, 224-260.
- Fan, J., Zhang J., & Yu K. (2012). Vast portfolio selection with gross-exposure constraints. *Journal of the American Statistical Association* 107(498), 592–606.
- Fastrich, B., Paterlini S., & Winker P. (2015). Constructing optimal sparse portfolios using regularization methods. *Computational Management Science* 12(3), 417–434.
- Flood, M., & Korenko, G. (2015). Systematic scenario selection: stress testing and the nature of uncertainty. *Quantitative Finance*, 15(1), 43-59.
- Gaines, B., Kim, J., & Zhou, H. (2018). Algorithms for fitting the constrained lasso. *Journal of Computational and Graphical Statistics*, 27(4), 861-871.
- Geyer, A., Hanke, M., & Weissensteiner, A. (2014). No-arbitrage bounds for financial scenarios. *European Journal of Operational Research*, 236(2), 657-663.
- Gouriéroux, C., & Jouneau, F. (1999). Econometrics of efficient fitted portfolios. *Journal of Empirical Finance*, 6(1), 87-118.
- Grover, J., & Lavin, A. (2007). Modern portfolio optimization: A practical approach using an excel solver single-index model. *The Journal of Wealth Management*, 10(1), 60-72.
- Hakansson, N. H. (1969). Risk disposition and the separation property in portfolio selection. *Journal of Financial and Quantitative Analysis*, 401-416.
- Hastie, T., Tibshirani, R., & Wainwright, M. (2015). *Statistical learning with sparsity: The lasso and generalizations*. CRC Press.
- Harvey, C., & Siddique, A. (2000). Conditional skewness in asset pricing tests. *The Journal of Finance*, 55(3), 1263-1295.
- Hoerl, A., & Kennard, R. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55-67.
- Huang, C., & Litzenberger, R. (1988). *Foundations for financial economics*. North-Holland.
- Ingersoll, J. (1987). *Theory of financial decision making*. Rowman & Littlefield.
- Israelsen, C. (2003). Sharpening the Sharpe Ratio. *Financial Planning*, 33(1), 49-51.

-
- Israelsen, C. (2005). A refinement to the Sharpe ratio and information ratio. *Journal of Asset Management*, 5(6), 423-427.
- Jagannathan, R., & Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance*, 58(4), 1651-1683.
- Jobson, J., & Korkie, B. (1981). Performance hypothesis testing with the Sharpe and Treynor measures. *Journal of Finance*, 889-908.
- Joo, Y., & Park, S. (2021). Optimal portfolio selection using a simple double-shrinkage selection rule. *Finance Research Letters*, 102019.
- Jorion, P. (1986). Bayes–Stein estimation for portfolio analysis. *The Journal of Financial and Quantitative Analysis*, 21, 279–292.
- Jurczenko, E., & Maillet, B. (2006). The four-moment capital asset pricing model: between asset pricing and asset allocation. In *Multi-moment asset allocation and pricing models*, New York: John Wiley, Chapter 6, 113-163.
- Kan, R., & Zhou, G. (2007). Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis*, 621-656.
- Kempf, A., & Memmel, C. (2006). Estimating the global minimum variance portfolio. *Schmalenbach Business Review*, 58(4), 332-348.
- Kim, J., Lee, Y., Kim, W., & Fabozzi, F. (2021). Mean–variance optimization for asset allocation. *The Journal of Portfolio Management* 47(4), 24-40.
- Klein, R., & Bawa, V. (1976). The effect of estimation risk on optimal portfolio choice. *Journal of Financial Economics*, 3(3), 215-231.
- Kolm, P., Tütüncü, R., & Fabozzi, F. (2014). 60 Years of portfolio optimization: Practical challenges and current trends. *European Journal of Operational Research*, 234(2), 356-371.
- Kremer, P., Lee, S., Bogdan, M., & Paterlini, S. (2020). Sparse portfolio selection *via* the sorted ℓ_1 -norm. *Journal of Banking and Finance*, 110, 105687.
- Laws, J. (2003). Portfolio analysis using excel. *Applied quantitative methods for trading and investment*, 293.
- Ledoit, O., & Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, 15(5), 850-859.
- Ledoit, O., & Wolf, M. (2011). Robust performances hypothesis testing with the variance. *Wilmott*, 2011(55), 86-89.
- Lintner, J. (1965a). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, 47 (1), 13-37.

-
- Lintner, J. (1965b). Security prices, risk, and maximal gains from diversification. *The Journal of Finance*, 20(4), 587-615.
- Lo, A. W. (2002). The statistics of Sharpe ratios. *Financial Analysts Journal*, 58(4), 36-52.
- Lobo, M., Fazel, M., & Boyd, S. (2007). Portfolio optimization with linear and fixed transaction costs. *Annals of Operations Research*, 152(1), 341-365.
- MacKinlay, A., & Pástor, L. (2000). Asset pricing models: Implications for expected returns and portfolio selection. *The Review of Financial Studies*, 13(4), 883-916.
- Mahalanobis, P. (1927). Analysis of race-mixture in Bengal. *Journal of the Asiatic Society of Bengal*, 23, 301-333.
- Mahalanobis, P. (1936). On the generalized distance in statistics. National Institute of Science of India.
- Mardia, K., Kent, J., & Bibby, J. (1979). *Multivariate analysis*. Amsterdam: Academic Press.
- Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77-91.
- Markowitz, H. (1959). *Portfolio selection: Efficient diversification of investments*. New York: John Wiley.
- Markowitz, H. M. (1999). The early history of portfolio theory: 1600-1960. *Financial Analysts Journal*, 55(4), 5-16.
- Markowitz, H. (2014). Mean-variance approximations to expected utility. *European Journal of Operational Research*, 234(2), 346-355.
- McNeil, A., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: concepts, techniques and tools*. Princeton university press.
- Mommel, C. (2003). Performance hypothesis testing with the Sharpe ratio. *Finance Research Letters* 1, 21-23.
- Merton, R. (1973). An intertemporal capital asset pricing model. *Econometrica*, 41(5), 867-887.
- Merton, R. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, 8(4), 323-361.
- Meucci, A. (2009). *Risk and asset allocation*. Springer.
- Michaud, R. (1989). The Markowitz optimization enigma: Is 'optimized' optimal? *Financial Analysts Journal*, 45(1), 31-42.
- Michaud, R., & Michaud R. (2008). Estimation error and portfolio optimization: A resampling solution. *Journal of Investment Management*, 6(1), 8-28.

-
- Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica*, 768-783.
- Opdyke, J. (2007). Comparing Sharpe ratios: so where are the p-values? *Journal of Asset Management*, 8(5), 308-336.
- Pye, G. (1967). Portfolio selection and security prices. *The Review of Economics and Statistics*, 49(1), 111-115.
- Roll, R. (1977). A critique of the asset pricing theory's tests. *Journal of Financial Economics*, 4, 129-176.
- Ross, S. (1978). Mutual fund separation in financial theory - The separating distributions. *Journal of Economic Theory*, 17(2), 254-286.
- Samuelson, P. (1967). General proof that diversification pays. *Journal of Financial and Quantitative Analysis*, 2(1), 1-13.
- Schanbacher, P. (2015). Averaging across asset allocation models. *Jahrbücher für Nationalökonomie und Statistik*, 235(1), 61-81.
- Sharpe, W. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3), 425-442.
- Stevens, G. (1998). On the inverse of the covariance matrix in portfolio analysis. *The Journal of Finance*, 53(5), 1821-1827.
- Tobin, J. (1958). Liquidity preference as behavior towards risk. *The Review of Economic Studies*, 25(2), 65-86.
- Tibshirani, R. (1996). Regression shrinkage and selection *via* the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288.
- Tu, J., & Zhou, G. (2011). Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies. *Journal of Financial Economics*, 99(1), 204-215.
- Turlach, B., & Wright, S. (2015). Quadratic programming. *Wiley Interdisciplinary Reviews: Computational Statistics*, 7(2), 153-159.
- Vieira, E., & Filomena, T. (2020). Liquidity constraints for portfolio selection based on financial volume. *Computational Economics*, 56, 1055-1077.
- Vinod, H., & Morey, M. (1999). Confidence intervals and hypothesis testing for the Sharpe and Treynor performance measures: A bootstrap approach. *Computational Finance*, 99, 25-40.
- Woodside-Oriakhi, M., Lucas, C., & Beasley, J. (2011). Heuristic algorithms for the cardinality constrained efficient frontier. *European Journal of Operational Research*, 213(3), 538-550.

-
- Yang, Y. (2000). Combining different procedures for adaptive regression. *Journal of Multivariate Analysis*, 74(1), 135-161.
- Yang, Y. (2001). Adaptive regression by mixing. *Journal of the American Statistical Association*, 96(454), 574-588.
- Yang, Y. (2004). Aggregating regression procedures to improve performance. *Bernoulli*, 10(1), 25-47.
- Yen, Y., & Yen, T. (2014). Solving norm constrained portfolio optimization *via* coordinate-wise descent algorithms. *Computational Statistics & Data Analysis*, 76, 737-759.
- Zhao, Z., Ledoit, O., & Jiang, H. (2020). Risk reduction and efficiency increase in large portfolios: leverage and shrinkage. University of Zurich, Department of Economics, Working Paper, No.328, 34 pages.
- Zou, H., & Hastie, T. (2005). Regularization and variable selection *via* the elastic net. *Journal of the Royal Statistical Society: Series B (statistical methodology)*, 67(2), 301-320.