Article

Mean Squared Variance Portfolio: A Mixed-Integer Linear Programming Formulation

Francisco Fernández-Navarro [1,*] **, Luisa Martínez-Nieto** [2] **, Mariano Carbonero-Ruz** [1] **and Teresa Montero-Romero** [2]

1   Department of Quantitative Methods, Universidad Loyola Andalucía, 41704 Sevilla, Spain; mariano@uloyola.es
2   Department of Financial Economics and Accounting, Universidad Loyola Andalucía, 41704 Sevilla, Spain; mlmartineznieto@al.uloyola.es (L.M.-N.); tmontero@uloyola.es (T.M.-R.)
*   Correspondence: fafernandez@uloyola.es

Abstract: The mean-variance (MV) portfolio is typically formulated as a quadratic programming (QP) problem that linearly combines the conflicting objectives of minimizing the risk and maximizing the expected return through a risk aversion profile parameter. In this formulation, the two objectives are expressed in different units, an issue that could definitely hamper obtaining a more competitive set of portfolio weights. For example, a modification in the scale in which returns are expressed (by one or percent) in the MV portfolio, implies a modification in the solution of the problem. Motivated by this issue, a novel mean squared variance (MSV) portfolio is proposed in this paper. The associated optimization problem of the proposed strategy is very similar to the Markowitz optimization, with the exception of the portfolio mean, which is presented in squared form in our formulation. The resulting portfolio model is a non-convex QP problem, which has been reformulated as a mixed-integer linear programming (MILP) problem. The reformulation of the initial non-convex QP problem into an MILP allows for future researchers and practitioners to obtain the global solution of the problem via the use of current state-of-the-art MILP solvers. Additionally, a novel purely data-driven method for determining the optimal value of the hyper-parameter that is associated with the MV and MSV approaches is also proposed in this paper. The MSV portfolio has been empirically tested on eight portfolio time series problems with three different estimation windows (composing a total of 24 datasets), showing very competitive performance in most of the problems.

Keywords: mean-variance portfolio; portfolio diversification; non-convex quadratic programming; mixed-integer linear programming

1. Introduction

Markowitz (1952) [1,2] proposed the well-known mean-variance (MV) portfolio model under the assumption that a rational investor aims at maximizing returns and minimizing risks. Mathematically speaking, the MV portfolio framework is a bi-objective optimization problem, where an efficient frontier is composed by all combination assets that are not dominated by any other in expected return and risk simultaneously. The two goals that are associated with the framework can be addressed separately as two independent optimization problems giving rise to the global maximum return (GMR) [3] and global minimum variance (GMV) [4–6] portfolios.

The MV portfolio has been widely implemented in the scientific community and accepted by professionals [7–9]. The main advantages of the approach are its ease of use, since the theory presents the concept of return and risk in a in a straightforward fashion and the ease in which the optimization problem is formulated. Unfortunately, the MV portfolio has yet to be proven competitive in out-of-sample validation [10–12]. Thus, the main limitations of the theory are as follows:

- High concentration: MV-inspired portfolios are highly concentrated on a few securities with the "best: features [13,14]. Assets with either high expected returns or low expected variance will be overweighted, in this way losing the power of diversification that the theory is supposed to ensure [15].
- Instability: MV portfolios tend to drastically re-allocate resources when the asset features change slightly, regardless of transaction costs or data inaccuracy [16,17]. This mainly occurs because MV portfolios do not take estimation inaccuracy into account and concentrate on assets with "good" features.
- Sensitivity to input errors: because the MV portfolio gives excessive weight to assets with high expected returns, the resulting portfolios are highly sensitive to errors in the input data [18–20].

The optimization problem that is associated with the MV portfolio can be formulated in several ways, depending on how the risk aversion parameter is incorporated into the equations. The main formulations for the optimization problem can be summarized, as follows:

1. Maximize the expected return for a specified risk: the first possible formulation includes the maximization of the portfolio mean in the objective function of the problem and the maximum level of risk that an investor is able to assume as a constraint of the problem.
2. Minimize the risk for a pre-determined expected return: the second alternative tries to minimize the risk and introduces the minimum level of the mean return as a constraint of the optimization.
3. Minimize the risk and maximize the expected return combining both of the objectives through a user-defined risk aversion parameter.

The above problems could have linear or non-linear constraints, equality, and inequality constraints. Furthermore, the three described optimization problems are mathematically equivalent and the solutions are called mean-variance (MV) efficient. Hence, the efficient points in the return-risk plot are called the efficient frontier.

One of the potential limitations of the third formulation of the problem is that it combines two objectives that are measured in different units. In fact, the variance is a second-order moment statistic and the mean is a first-order moment one. Consequently, a modification in the scale in which returns are expressed (both by one or percent, for example), differently affects them, accordingly changing the solution of the problem. Furthermore, the most widely used performance measure in the quantitative finance literature to compare portfolio strategies is the Sharpe ratio, which computes the quotient between the return of the portfolio with respect to the standard deviation of the portfolio's excess return [21–23]. As can be seen, the Sharpe ratio assesses the two objectives in the same unit level, unlike the third formulation of the optimization problem, which evaluates the objectives in different units [24].

Motivated by this fact, a novel portfolio strategy, which is denoted as mean squared variance (MSV), which calculates the two objectives of the minimization function of the problem in the same unit is proposed in this paper. Specifically, the mean return of the portfolio is expressed in squared form in the proposed strategy, aiming to provide a straightforward manner of comparing the two objectives in the same unit. The resulting portfolio model is a non-convex quadratic programming (QP) problem, which has been reformulated as a mixed-integer linear programming (MILP) problem to reach, in this way, the global solution of the problem.

The reformulation of the baseline QP problem as an MILP is done via the adaptation of the QP problem as a linear complementary problem, the use of binary variables and big-M constraints to model the complementary constraints. The estimation of the upper bounds on the dual variables was borrowed from the work of [25] for the case of a standard QP problem. The reformulation of the initial non-convex QP problem into an MILP has the advantage of ensuring that the global solution is obtained via the use of current state-of-the-art MILP solvers.

Finally, an innovative purely data-driven method for determining the optimal value of the hyper-parameter that is associated with the MV and MSV strategies (the risk aversion parameter, λ) is also proposed in this paper. The hyper-parameter optimization is based on a Bayesian approach for the global optimization of λ [26]. The approach will allow the manuscript authors to show that the proposed formulation of the MV problem helps the global optimization procedure achieve a more competitive set of portfolio weights (in terms of the Sharpe ratio).

The rest of the paper is organized, as follows. Section 2 describes the foundations of the proposed mean squared variance (MSV) strategy and its reformulation as a mixed-integer linear programming (MILP) optimization problem. Section 3 details the experimental framework that was developed for the empirical comparison of the state-of-the-art portfolio strategies implemented. Section 4 reports the statistical results obtained by the different strategies and Section 5 presents the conclusions that were reached in the study.

2. The Proposed Method

The aim of this section is to describe the mathematical formulation (Section 2.1) and foundations (Section 2.2) of the proposed method, named the mean squared variance (MSV) portfolio, as well as to provide the necessary details to reformulate the model as a mixed-integer linear programming (MILP) optimization problem (Section 2.3).

2.1. Mathematical Formulation of the Model

The proposed method determines, for a portfolio consisting of N assets ($n = 1, \ldots, N$), the optimal weights of the portfolio's value invested in each asset, $w_1, \ldots, w_N \geq 0$ ($\sum_{n=1}^{N} w_n = 1$), while using, as inputs of the problem, the expected returns (μ_n), risks (σ_n) and covariances between assets (σ_{nm}). As previously mentioned, the associated optimization problem of the model is very similar to the Markowitz optimization, except in the part of the portfolio mean, which is presented in squared form in our formulation. Thus, the MSV portfolio is defined as:

$$\begin{aligned}
\min_{w_1, \ldots, w_N} \quad & \lambda \sum_{n,m=1}^{N} w_n w_m \sigma_{nm} - (1-\lambda) \left(\sum_{n=1}^{N} w_n \mu_n \right)^2. \\
\text{s. t.} \quad & \sum_{n=1}^{N} w_n = 1. \\
& w_1, \ldots, w_N \geq 0,
\end{aligned} \tag{1}$$

where the term $\sum_{n=1}^{N} w_n \mu_n$ is the portfolio mean, $\sum_{n,m=1}^{N} w_n w_m \sigma_{nm}$, ($\sigma_{nn} = \sigma_n^2$) is the portfolio variance and $\lambda \in [0,1]$ is a hyper-parameter of the problem that weights the relative importance of the risk with respect to the mean squared return. The model can also be formulated in matrix form as follows:

$$\begin{aligned}
\min_{\mathbf{w} \in \mathbb{R}^N} \quad & \lambda \mathbf{w}^T \mathbf{\Sigma} \mathbf{w} - (1-\lambda) \mathbf{w}^T \boldsymbol{\mu} \boldsymbol{\mu}^T \mathbf{w}. \\
\text{s. t.} \quad & \mathbf{1}^T \mathbf{w} = 1. \\
& \mathbf{w} \geq \mathbf{0},
\end{aligned} \tag{2}$$

where $\mathbf{w} = (w_1 \ldots w_N)^T \in \mathbb{R}^N$ is the vector with the optimal weights of the portfolio, $\mathbf{\Sigma} \in \mathbb{R}^{N \times N}$ is the covariance matrix of the asset's returns, $\boldsymbol{\mu} = (\mu_1 \ldots \mu_N)^T \in \mathbb{R}^N$ is the vector with the expected returns, and $\mathbf{0}$ and $\mathbf{1}$ are N-dimensional vectors with zeros and ones in all the rows, respectively. For ease of reference, the proposed MSV portfolio is expressed as:

$$\min_{\mathbf{w} \in \mathbb{R}^N} \quad \frac{1}{2} \mathbf{w}^T \mathbf{H} \mathbf{w}.$$
$$\text{s. t.} \quad \mathbf{1}^T \mathbf{w} = 1. \tag{3}$$
$$\mathbf{w} \geq \mathbf{0},$$

with $\mathbf{H} = \lambda \boldsymbol{\Sigma} - (1 - \lambda)\mathbf{R} \in \mathbb{R}^{N \times N}$, and $\mathbf{R} = \boldsymbol{\mu}\boldsymbol{\mu}^T$ and the constant $1/2$ has been incorporated in the objective function in order to facilitate its derivation.

2.2. Main Foundations of the Model

In this section, the optimal portfolios of the MSV and MV models are analytically compared. The GMV and GMR portfolios are also included in the study for reference. The four optimization problems share the same feasible region and, therefore, for their analysis, it will suffice to compare their objective functions:

$$F^{\text{MSV}}(\mathbf{w}|\lambda) = \lambda \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w} - (1 - \lambda)\mathbf{w}^T \boldsymbol{\mu}\boldsymbol{\mu}^T \mathbf{w}.$$
$$F^{\text{MV}}(\mathbf{w}|\lambda) = \lambda \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w} - (1 - \lambda)\mathbf{w}^T \boldsymbol{\mu}.$$
$$F^{\text{GMV}}(\mathbf{w}) = \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w}.$$
$$F^{\text{GMR}}(\mathbf{w}) = -\mathbf{w}^T \boldsymbol{\mu}.$$

It is straightforward to realize that, when only the expected return is taken into account ($\lambda = 0$), the MSV, MV, and GMR models provide the same optimal solution. Additionally, when only risk is taken into account ($\lambda = 1$), MSV, MV, and GMV achieve the same portfolio. The analysis in intermediate situations, when both return and risk ($0 < \lambda < 1$) are considered, is not so immediate. This is pursued assuming that $0 < \mathbf{w}^T \boldsymbol{\mu} < 1$ and, therefore, the expected return for any portfolio is positive and lower that 100%. Under these conditions, $\mathbf{w}^T \boldsymbol{\mu}\boldsymbol{\mu}^T \mathbf{w} = (\mathbf{w}^T \boldsymbol{\mu})^2 < \mathbf{w}^T \boldsymbol{\mu}$, so $F^{\text{MV}}(\mathbf{w}|\lambda) \leq F^{\text{MSV}}(\mathbf{w}|\lambda)$.

In any case, Section 4 shows how the mathematical formulation of the MSV problem helps the model to achieve significantly better out-of-sample Sharpe ratios than the ones that were reported by the traditional MV method.

2.3. Mixed-Integer Linear Programming Reformulation

The **H** matrix in the QP problem that is associated to the model (see Equation (3)) is the difference between two positively weighted matrices ($\boldsymbol{\Sigma}$ and $\mathbf{R}$). In this context, positivity of H can not be assured and the QP problem associated may be a non-convex QP problem. Due to this fact, the authors detail below how the model could be reformulated as an MILP optimization problem. The reformulation will allow for us to yield the global optimum in each portfolio problem using the current state-of-the-art MILP solvers.

The first step in the reformulation is the introduction of the Lagrange multipliers $\kappa \in \mathbb{R}$ for the equality constraint and $\boldsymbol{\delta} \in \mathbb{R}^N$ for the non-negativity constraints (no short sales restrictions). The KKT conditions for the problem after introducing those multipliers are given by:

$$\mathbf{H}\mathbf{w} + \mathbf{1}\kappa - \boldsymbol{\delta} = \mathbf{0}.$$
$$\mathbf{w}^T \boldsymbol{\delta} = 0.$$
$$\mathbf{1}^T \mathbf{w} = 1. \tag{4}$$
$$\mathbf{w}, \boldsymbol{\delta} \geq \mathbf{0}.$$

In this way, the KKT points of the optimization problem are denoted as:

$$\Omega_{\text{KKT}} = \{(\mathbf{w}, \kappa, \boldsymbol{\delta}) \in \mathbb{R}^{2N+1} : (\mathbf{w}, \kappa, \boldsymbol{\delta}) \text{ satisfy Equations in } (4)\} \tag{5}$$

In the proposed portfolio problem, the KKT conditions are first-order necessary conditions for the optimal solutions of the system, since the feasible set of QP is a polyhedron. Because of that, the KKT conditions can be incorporated into the optimization problem

as redundant constraints, in this way obtaining the following equivalent formulation of the problem:

$$\min_{\mathbf{w} \in \mathbb{R}^N} \quad (1/2)\mathbf{w}^T \mathbf{H} \mathbf{w}.$$
$$\text{s. t.} \quad \mathbf{H}\mathbf{w} + \mathbf{1}\kappa - \delta = \mathbf{0}.$$
$$\mathbf{w}^T \delta = 0. \tag{6}$$
$$\mathbf{1}^T \mathbf{w} = 1.$$
$$\mathbf{w}, \delta \geq \mathbf{0}.$$

Next, the KKT conditions are used in order to linearize the objective function, as suggested in [27]. First, the equality constraint $\mathbf{H}\mathbf{w} = \delta - \mathbf{1}\kappa$ is used to express the objective function as:

$$(1/2)\mathbf{w}^T \mathbf{H}\mathbf{w} = (1/2)\mathbf{w}^T(\delta - \mathbf{1}\kappa) = (1/2)\mathbf{w}^T \delta - (1/2)\mathbf{w}^T \mathbf{1}\kappa, \tag{7}$$

and, second, when considering that $\mathbf{w}^T \delta = 0$ and $\mathbf{1}^T \mathbf{w} = 1$, the final objective function, in linear form, is defined as $-(1/2)\kappa$, equivalent to $-\kappa$.

As a result of the previously mentioned transformations, the optimization problem (3) is equivalent to the following problem with a linear (instead of quadratic) objective:

$$\min_{\mathbf{w} \in \mathbb{R}^N} \quad -\kappa.$$
$$\text{s. t.} \quad \mathbf{H}\mathbf{w} + \mathbf{1}\kappa - \delta = \mathbf{0}.$$
$$\mathbf{w}^T \delta = 0. \tag{8}$$
$$\mathbf{1}^T \mathbf{w} = 1.$$
$$\mathbf{w}, \delta \geq 0.$$

and, while taking into account that the δ and $\mathbf{w}$ variables are complementary to each other (complementary constraint), the problem can be formulated as:

$$\min_{\mathbf{w} \in \mathbb{R}^N} \quad -\kappa.$$
$$\text{s. t.} \quad \mathbf{H}\mathbf{w} + \mathbf{1}\kappa - \delta = \mathbf{0}_N.$$
$$\mathbf{1}^T \mathbf{w} = 1. \tag{9}$$
$$\mathbf{0} \leq \mathbf{w} \perp \delta \geq \mathbf{0},$$

where $\mathbf{w} \perp \delta$ denotes perpendicularity between vectors $\mathbf{w}$ and δ, and it impedes the optimization problem as a linear problem. Note that problem (9) is equivalent to 2^N linear programs obtained by making each possible assignment for the complementary variables: either $\delta_n = 0$ or $w_n = 0$ for each $n = 1, \ldots, N$; hence, it is not possible for the optimization problem to have a finite optimal solution that is not attained. Therefore, problem (9) can be reformulated as an MILP while assuming that the upper bounds for $\mathbf{w}$ and δ exist and they are defined as vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^N$ with entries u_n and v_n:

$$\min_{\mathbf{w} \in \mathbb{R}^N} \quad -\kappa.$$
$$\text{s. t.} \quad \mathbf{H}\mathbf{w} + \mathbf{1}\kappa - \delta = \mathbf{0}.$$
$$\mathbf{1}^T \mathbf{w} = 1.$$
$$0 \leq w_n \geq z_n u_n. \tag{10}$$
$$0 \leq \delta_n \geq (1 - z_n)v_n.$$
$$z_n \in \{0, 1\},$$

for $n = 1, \ldots, N$. The obvious drawback of this formulation is that in order to find u_n and v_n, the upper bounds of w_n and δ_n should be known a priori. The upper bound, $\mathbf{u} \in \mathbb{R}^N$, for the primal variables, $\mathbf{w} \in \mathbb{R}^N$, is $\mathbf{u} = \mathbf{1}$ and it can be computed (while assuming that the feasible set of QP is non-empty and bounded) by setting:

$$u_n = \max\{w_n : \mathbf{1}^T \mathbf{w} = 1, w_n \geq 0\}, \tag{11}$$

for every $n = 1, \ldots, N$. The calculation of the upper bounds, $\mathbf{v} \in \mathbb{R}^N$, on the dual variables, $\delta \in \mathbb{R}^N$, was borrowed from the research work of [25], which sets those bounds (for a standard quadratic program) as:

$$\mathbf{v} \geq s\mathbf{1}, \tag{12}$$

with

$$s = 2N(||\mathbf{H}||_{\infty,\infty} + ||\mathbf{f}||_\infty), \tag{13}$$

where $||\mathbf{f}||_\infty = 0$ in the proposed formulation as $\mathbf{f}$ represents the vector that is associated with the linear part of the optimization problem (non-existent in the MSV strategy).

3. Experimental Framework

The aim of this section is to fully define the datasets and methods used for comparison purposes, the performance metrics, and the hypothesis tests that were implemented in the present empirical study.

3.1. Out-Of-Sample Empirical Validation and Portfolio Problems Selected

The empirical validation implemented is inspired by the rolling window out-of-sample validation method [28]. Therefore, given a time series with T months of securities returns, an estimation window length, M, is selected and, in each month t, starting from $t = M + 1$, the data in the previous M months are employed to compute the optimal weights that are associated with each strategy. The return in month t is computed while using the previously computed weights.

The optimization process is repeated Q times by adding, in each iteration, the return of the next period in the time series and removing from the estimation window the earliest return. In this way, the outcome of this empirical validation, for each strategy and time series tested, is a set of Q out-of-sample returns, $r_1, \ldots, r_Q$. In the experimental framework, the estimation window length, M, is set to 60, 120, and 240, as recommended in [10], $M \in \{60, 120, 240\}$. Additionally, the validation process is repeated 36 times, $Q = 36$.

Table 1 details the properties of the selected time series portfolio problems, for each portfolio problem, its ID, name (Dataset), the estimation window length (M), the number of total months considered (T) and the number of assets (N). The datasets considered were obtained from Kenneth French's web site at Dartmouth University (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/). It is important to mention that these portfolio problems have been widely used in the community for validating novel portfolio strategies [10,29,30]. The date of the last monthly excess returns considered for all the problems is 1 August 2020.

Each dataset contained monthly returns in excess of the one-month Treasury bill rate. All portfolios are constructed on June of each year with NYSE, AMEX, and NASDAQ stocks (real-world portfolio datasets). The industry portfolio holds one portfolio of each of five or ten different groups of sectors that formed by the SIC code (consumer durable and non-durables, manufacturing, oil and gas, business equipment, transmission, retail, healthcare, utilities, and others, etc.).

Table 1. Characteristics of the benchmark datasets.

ID.	Dataset	M	T	N
	Industry Portfolios			
$\{1,2,3\}$	5 Industry Portfolios	$\{60, 120, 240\}$	$\{96, 156, 276\}$	5
$\{4,5,6\}$	10 Industry Portfolios	$\{60, 120, 240\}$	$\{96, 156, 276\}$	10
	Emerging Market Factors			
$\{7,8,9\}$	6 Emerging Market Portfolios Formed on BM and OP	$\{60, 120, 240\}$	$\{96, 156, 276\}$	6
$\{10, 11, 12\}$	6 Emerging Market Portfolios Formed on Size and BM	$\{60, 120, 240\}$	$\{96, 156, 276\}$	6
$\{13, 14, 15\}$	6 Emerging Market Portfolios Formed on Size and OP	$\{60, 120, 240\}$	$\{96, 156, 276\}$	6
	Bivariate sorts on Size, BM and I			
$\{16, 17, 18\}$	6 Portfolios Formed on Size and BM	$\{60, 120, 240\}$	$\{96, 156, 276\}$	6
$\{19, 20, 21\}$	6 Portfolios Formed on Size and I	$\{60, 120, 240\}$	$\{96, 156, 276\}$	6
	Bivariate sorts on Size, BM and I			
$\{22, 23, 24\}$	25 European Portfolios Formed on Size and BM	$\{60, 120, 240\}$	$\{96, 156, 276\}$	25

BM: Book-to-Market; OP: Operating Profitability; I: Investment.

3.2. Strategies Implemented

The proposed strategy (MSV) is compared, in the experimental study, to the baseline MV strategy and its corresponding parts: the GMR and GMV strategies. The proposed strategy, MSV, and the well-known MV both include an additional hyper-parameter, λ, which defines the importance of the variance term with respect to the mean. The range of values of the λ hyper-parameter has been set to $\lambda \in [0, 1]$.

3.3. Performance Measures

In the experimentation, two well-known performance measures are computed for each strategy in each time series portfolio problem. Specifically:

1. The out-of-sample mean returns (MR):

$$\text{MR} = (1/Q) \sum_{q=1}^{Q} r_q. \tag{14}$$

2. The out-of-sample Sharpe ratio (SR), defined as the sample mean of out-of-sample excess returns, MR, divided by their corresponding sample standard deviation:

$$\text{SR} = \frac{\text{MR}}{\sqrt{(1/Q) \sum_{q=1}^{Q} (r_q - \text{MR})^2}}. \tag{15}$$

3.4. Hyper-Parameter Optimization

In this research, the hyper-parameters of each strategy have been globally optimized while using the data available in the estimation window part in order to adapt the model to the behavior of the market in each dataset. Thus, the last year of the estimation window data was used in an internal validation process to evaluate the performance, in terms of out-of-sample Sharpe ratio, which is associated with the specific value of the hyper-parameter being explored, whereas the previous years were employed to compute the optimal weights for this λ value.

The hyper-parameter optimization was implemented through a Bayesian process. Bayesian optimization is an algorithm employed in global optimization to minimize a certain objective function, treated as a black box, by varying the value of its independent variables. The algorithm itself relies on an internal Gaussian process that approximates the objective function and it is trained by subsequent evaluations of the true objective. The

approximated model is used for optimization to reduce computational costs and for its robust nature in addressing stochastic noise in function evaluations.

3.5. Statistical Hypothesis Testing

Statistical hypothesis testing was used in order to provide statistical support in the evaluation of the reported results. It is important to clarify that a statistical analysis through parametric tests could lead to mistaken conclusions in this study, since a previous evaluation of the MR and SR values that were provided by the strategies implemented resulted in rejecting the normality and equality of the variance hypothesis. Furthermore, as noted by Demšar [31], the independence condition is not truly verified in a rolling window out-of-sample validation. For these reasons, two non-parametric Friedman tests (with the ranking of MR and SR of the strategies as the test variables) were carried out to assess the statistical significance of the rank differences. Consequently, a non-parametric Holm post hoc test was used to ascertain which strategies were distinctive among the multiple comparisons performed [31].

4. Results

In this section, the different experimental analyses that were carried out with the proposed portfolio model are detailed. In particular, the goals of the section are as follows:

1. to compare the out-of-sample performance of the MSV portfolio with the performance provided by state-of-the-art MV-based strategies (Section 4.1); and,
2. to analyse the diversification levels produced by the proposed MSV portfolio and the MV portfolio in problems with different dimensions (Section 4.2).

4.1. Performance Analysis

In this subsection, the performance results of the MSV strategy are compared to those that are provided by the standard MV strategy and its corresponding individual components (the GMV and GMR strategies). Table 2 shows the MR and SR out-of-sample performance results for each time series and the implemented strategies. For the strategies requiring an internal hyper-parameter optimization (such as the MV or MSV), the validation process is repeated 10 times in order to evaluate the robustness of the stochastic optimization procedure (Section 3.4) and, consequently, the performance reported for those strategies in each dataset is the mean and standard deviation of the MR and SR performance achieved in the 10 repetitions considered for the time series being analyzed (Table 2). Based on the reported MR and SR for each technique and dataset, the ranking of each strategy in each dataset and measure (R = 1 for the best performing strategy and R = 4 for the worst one) is obtained. In addition, Table 2 presents the mean MR and SR over the 24 datasets that were considered for each strategy ($\overline{MR}$ and $\overline{SR}$) and their corresponding mean rankings ($\overline{R}_{MR}$ and $\overline{R}_{SR}$).

From a purely descriptive point of view, it can be concluded that the MSV strategy obtained the best results in 12 datasets and the second-best results in nine problems when the MR measure is considered. With regard to the SR measure, the MSV portfolio achieved the best results in 10 time series and the second-best results in 12 datasets. Furthermore, the MSV portfolio yields the best mean ($\overline{MR} = 0.8731$, $\overline{SR} = 0.1607$) and ranking ($\overline{R}_{MR} = 1.7500$, $\overline{R}_{SR} = 1.7083$) in MR and SR, respectively. The second-best mean ranking is obtained by the baseline MV method in the two performance measures ($\overline{R}_{MR} = 2.4792$, $\overline{R}_{SR} = 2.5417$). The third most competitive method appears to be the GMR strategy that achieved the second-best mean MR ($\overline{MR} = 0.8426$) and the third-best mean MR and SR ranking ($\overline{R}_{MR} = 2.5208$, $\overline{R}_{SR} = 2.7292$). Finally, the worst performing portfolio was the GMV one, which reported the worst mean MR and SR ($\overline{MR} = 0.6934$, $\overline{SR} = 0.1369$) and worst mean MR and SR ranking ($\overline{R}_{MR} = 3.2500$, $\overline{R}_{SR} = 3.0208$).

The competitive performance of the proposed portfolio can also be graphically seen in Figure 1, in which the cumulative monthly returns of the four implemented strategies during the out-of-sample validation process for the five Industry Portfolio problems ($M = 60$, ID = 1) are reported. The MSV strategy seems to be an appealing approach for

portfolio optimization, in performance terms, as can be seen in Figure 1. Another aspect that is important to stress is that the MSV portfolio is less sensitive to the length of the estimation windows than its comparison strategies, reporting competitive results for different values of M (unlike, for example, the GMR strategy that usually only achieves competitive results for small values of M). In a similar fashion, the MSV strategy also achieves competitive results for different values of N (problems with a different number of assets).

Table 2. Statistical results of the strategies implemented: the mean returns (MR) and Sharpe ratio (SR) out-of-sample performance of each strategy and dataset (MR and SR), mean MR ($\overline{MR}$), mean MR ranking ($\overline{R}_{MR}$), mean SR ($\overline{SR}$) and mean SR ranking ($\overline{R}_{SR}$). The datasets are denoted by their IDs. The best results are in bold face and second-best results in italics.

ID	MR Metric				SR Metric				ID
	GMV	GMR	MV	MSV	GMV	GMR	MV	MSV	
1	0.9521	*1.6197*	$1.1717_{0.0065}$	$\mathbf{1.8161_{0.0000}}$	0.1330	*0.2243*	$0.1646_{0.0015}$	$\mathbf{0.2339_{0.0000}}$	1
2	1.4004	**1.8981**	$1.8977_{0.0003}$	$1.8981_{0.0000}$	0.2825	**0.3408**	$0.3391_{0.0008}$	$0.3408_{0.0000}$	2
3	1.3098	0.9897	$1.3146_{0.0013}$	$\mathbf{1.3478_{0.0034}}$	0.2698	0.1748	$0.2705_{0.0008}$	$\mathbf{0.2775_{0.0021}}$	3
4	0.6058	**2.1011**	$1.4175_{0.0008}$	$1.2533_{0.0338}$	0.1411	**0.3493**	$0.2426_{0.0005}$	$0.2965_{0.0098}$	4
5	0.6324	**1.9033**	$1.2299_{0.0012}$	$1.1355_{0.0004}$	0.1541	**0.3077**	$0.2174_{0.0013}$	$0.2062_{0.0003}$	5
6	0.7234	**1.2039**	$0.8252_{0.0021}$	$1.0470_{0.0866}$	0.1752	0.2079	$\mathbf{0.2093_{0.0007}}$	$0.2084_{0.0124}$	6
7	0.5404	0.5456	$0.5477_{0.0008}$	$\mathbf{0.5577_{0.0013}}$	*0.1016*	0.0960	$0.1016_{0.0003}$	$\mathbf{0.1028_{0.0005}}$	7
8	**0.5473**	0.4494	$0.5342_{0.0008}$	$0.4494_{0.0000}$	**0.1027**	0.0791	$0.0959_{0.0002}$	$0.0791_{0.0000}$	8
9	**0.5118**	-0.1094	$-0.1042_{0.0118}$	$0.0021_{0.0089}$	**0.0960**	-0.0175	$-0.0166_{0.0019}$	$0.0021_{0.0017}$	9
10	0.1898	0.6411	$0.6524_{0.0419}$	$0.6848_{0.0321}$	0.0324	0.1277	$0.1295_{0.0072}$	$\mathbf{0.1351_{0.0057}}$	10
11	0.4162	**0.6039**	$0.5962_{0.0051}$	$0.6005_{0.0041}$	0.0742	*0.1268*	$0.1249_{0.0010}$	$\mathbf{0.1270_{0.0007}}$	11
12	0.3150	*0.1772*	$0.1772_{0.0000}$	$\mathbf{0.1775_{0.0012}}$	*0.0561*	0.0306	$0.0306_{0.0000}$	$0.0309_{0.0002}$	12
13	0.1505	**0.4075**	$0.1259_{0.0031}$	$0.2530_{0.0810}$	0.0262	**0.0747**	$0.0220_{0.0010}$	$0.0456_{0.0266}$	13
14	0.2519	**0.3078**	$0.2679_{0.0014}$	$0.2756_{0.0294}$	0.0444	**0.0614**	$0.0474_{0.0008}$	$0.0515_{0.0091}$	14
15	0.1601	**0.3181**	$0.1640_{0.0321}$	$0.1919_{0.0705}$	0.0283	**0.0561**	$0.0291_{0.0087}$	$0.0339_{0.0124}$	15
16	1.1048	*1.7043*	$1.5803_{0.1945}$	$\mathbf{1.7876_{0.0000}}$	0.2017	*0.3252*	$0.2992_{0.0413}$	$\mathbf{0.3288_{0.0000}}$	16
17	1.4131	1.4092	$1.4394_{0.0187}$	$\mathbf{1.4656_{0.0021}}$	0.2534	0.2600	$\mathbf{0.2682_{0.0022}}$	$0.2639_{0.0007}$	17
18	1.0582	0.1845	$1.1500_{0.0018}$	$\mathbf{1.2829_{0.3904}}$	*0.2032*	0.0276	$0.2013_{0.0004}$	$\mathbf{0.2280_{0.0085}}$	18
19	1.0294	1.4941	$1.4982_{0.1002}$	$\mathbf{1.5056_{0.0196}}$	0.1948	0.2271	$0.2445_{0.0106}$	$\mathbf{0.2621_{0.0023}}$	19
20	1.0063	*1.0665*	$\mathbf{1.0790_{0.0008}}$	$1.0067_{0.0009}$	0.1908	0.1721	$\mathbf{0.1930_{0.0001}}$	$0.1912_{0.0003}$	20
21	*0.7662*	0.0917	$0.7658_{0.0009}$	$\mathbf{0.7664_{0.0007}}$	0.1488	0.0134	$\mathbf{0.1497_{0.0002}}$	$0.1490_{0.0002}$	21
22	0.4391	0.5806	$\mathbf{0.5843_{0.0091}}$	$0.5823_{0.0064}$	0.1011	*0.1019*	$0.0468_{0.0015}$	$\mathbf{0.1048_{0.0023}}$	22
23	0.7158	0.8431	$0.8613_{0.0148}$	$\mathbf{0.8641_{0.0217}}$	**0.1801**	0.1530	$0.1548_{0.0027}$	$0.1577_{0.0031}$	23
24	**0.4016**	-0.2092	$-0.1046_{0.0113}$	$0.0032_{0.0283}$	**0.0953**	-0.0347	$-0.0324_{0.0023}$	$0.0008_{0.0009}$	24
$\overline{MR}$	0.6934	*0.8426*	0.8196	**0.8731**	0.1369	0.1452	*0.1472*	**0.1607**	$\overline{SR}$
$\overline{R}_{MR}$	3.2500	2.5208	*2.4792*	**1.7500**	3.0208	2.7292	*2.5417*	**1.7083**	$\overline{R}_{SR}$

In order to determine the statistical significance of the rank differences reported for each strategy in the different problems, a non-parametric Friedman test [32] with the ranking of MR and SR of the techniques that were implemented as the test variables has been carried out. The test shows that the effect of the strategy used for portfolio optimization is statistically significant at a significance level of 5%, as the confidence interval is $C_0 = (0, F_{0.05} = 2.7374)$ and the F-distribution statistical values are $F^\star = 6.6840 \notin C_0$ for MR and $F^\star = 5.4109 \notin C_0$ for SR. Consequently, the null hypothesis stating that all strategies perform equally in mean ranking is rejected.

Based on this rejection, the non-parametric Holm test was implemented to compare all of the strategies to the proposed MSV method (which was used as the control method). Table 3 shows the results of the Holm test for the MR and SR performance measures. From the results of this test, it can be concluded that MSV obtains a significantly better MR ranking than all of the remaining strategies for $\alpha = 0.10$ and statistically outperforms the GMV strategy for $\alpha = 0.05$. While using the Sharpe ratio, SR, as the test variable, the MSV portfolio achieves significantly better results than the comparison methods for $\alpha = 0.10$ and $\alpha = 0.05$.
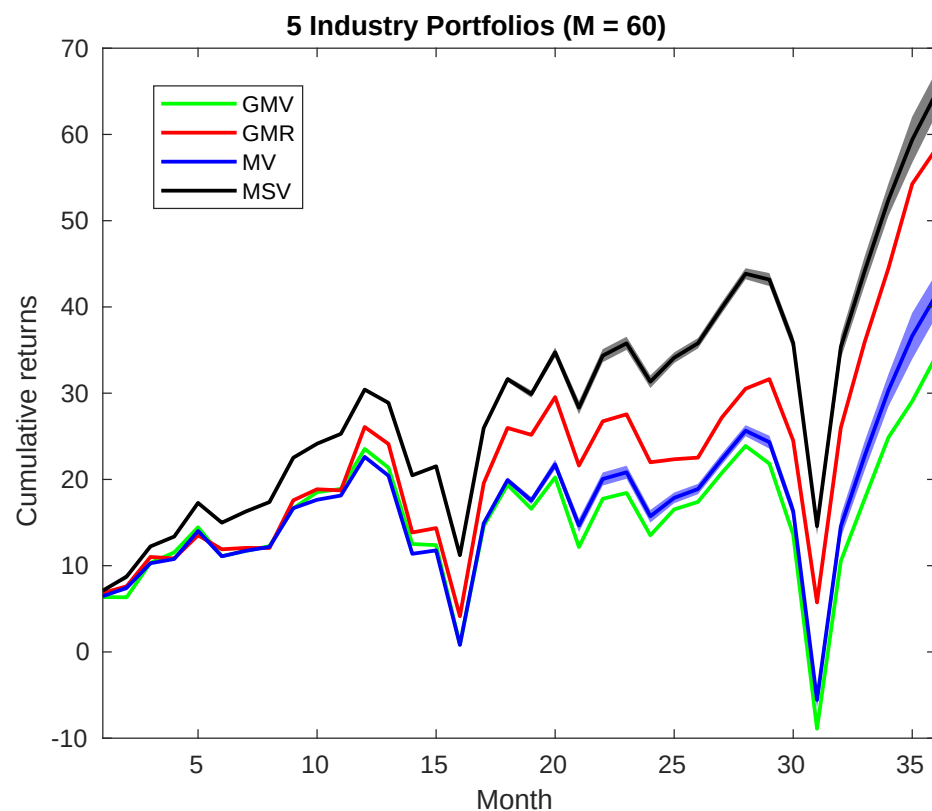
Figure 1. Cumulative monthly returns plot of the four portfolios implemented from 1 September 2017–1 August 2020.

As can be seen, the MSV model achieve significantly better out-of-sample results (in MR and SR) that its MV counterpart. Thus, it is assumed that the current formulation of the model helps in the tuning of the risk aversion profile parameter and, consequently, it helps in obtaining a final competitive out-of-sample performance.

Table 3. Statistical results for the Holm test for $\alpha = 0.10$ and $\alpha = 0.05$ using the mean squared variance (MSV) strategy as the control method: mean MR and SR ranking of the strategies implemented ($\overline{R}_{MR}$ and $\overline{R}_{SR}$), *z*-statistics and *p*-values of the Holm tests for the MR and SR analysis and adjusted α Holm values ($\alpha_{0.10}$ and $\alpha_{0.05}$). The best results are in bold face and second-best results in italics.

MR Analysis					
Method	$\overline{R}_{MR}$	*z*-statistic	*p*-value	$\alpha_{0.10}$	$\alpha_{0.05}$
GMV$_{\bullet,\circ}$	3.2500	4.0249	1×10^{-4}	0.0333	0.0167
GMR$_{\bullet}$	2.5208	2.0683	0.0386	0.0500	0.0250
MV$_{\bullet}$	2.4792	1.9566	0.0504	0.1000	0.0500
MSV	**1.7500**	-	-	-	-
SR Analysis					
Method	$\overline{R}_{SR}$	*z*-statistic	*p*-value	$\alpha_{0.10}$	$\alpha_{0.05}$
GMV$_{\bullet,\circ}$	3.0208	3.5218	4×10^{-4}	0.0333	0.0167
GMR$_{\bullet,\circ}$	2.7292	2.7394	0.0062	0.0500	0.0250
MV$_{\bullet,\circ}$	*2.5417*	2.2362	0.0253	0.1000	0.0500
MSV	**1.7083**	-	-	-	-

$\bullet$: Statistical difference with $\alpha = 0.10$; $\circ$: Statistical difference with $\alpha = 0.05$.

4.2. Diversification Analysis

An additional experimental study was performed in order to graphically illustrate the level of portfolio concentration of the MSV portfolio in contrast to the MV portfolio [4,13].

In this empirical study, the diversification levels of the MSV and MV portfolios in different types of problems were compared according to their $\mathbf{w}^T\mathbf{w}$ values, a mathematical proxy for assessing the degree of concentration in the portfolio weights [12,33–35]. Thus, Figure 2 shows the boxplots of diversification of the optimal weights of the MV and MSV portfolios for $\lambda \in \{0.2, 0.4, 0.6, 0.8\}$ in problems with different numbers of assets ($N \in \{7, 8, 9, 10\}$). For each value of N, the boxplot of each strategy is computed according to 30 randomly generated portfolio problems with the same number of assets and λ values. In this specific experiment, the randomly generated expected returns are within the range $[0.5, 1.0]$.
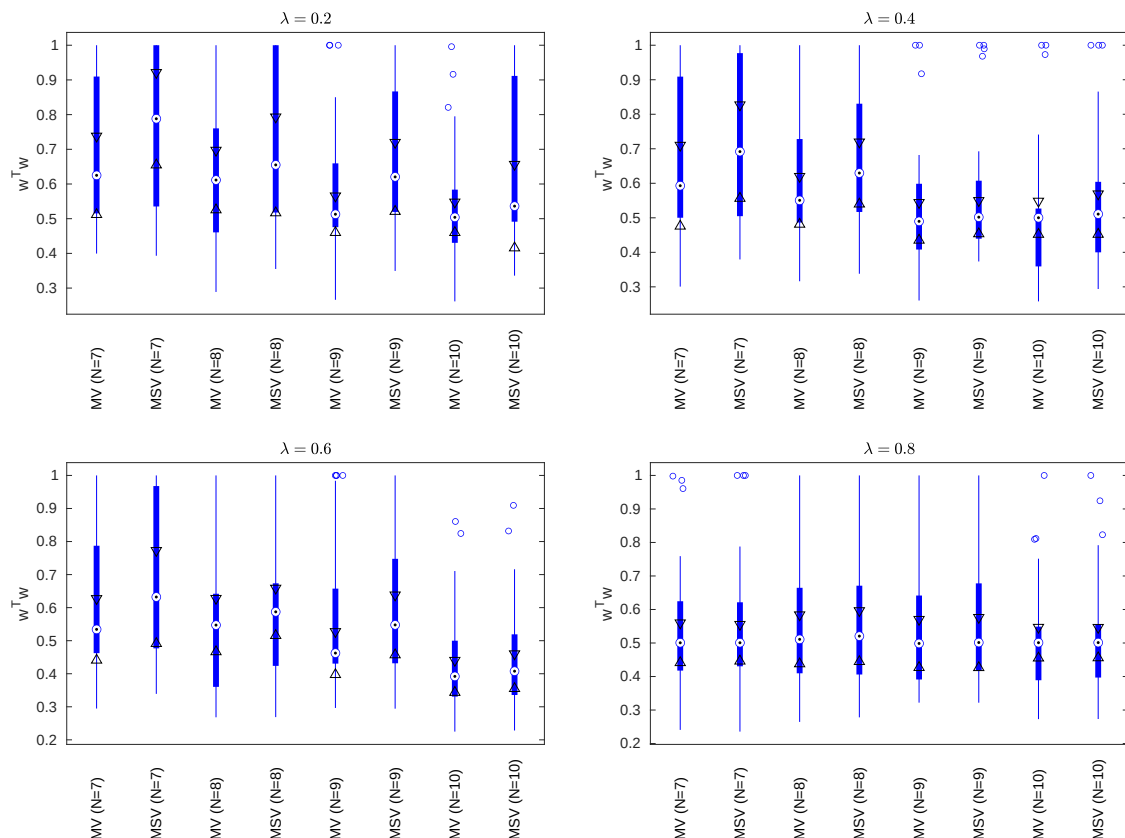

Figure 2. The boxplots of diversification for the MV and MSV strategies for $N \in \{7, 8, 9, 10\}$ and $\lambda \in \{0.2, 0.4, 0.6, 0.8\}$.

The MSV tends to produce more concentrated portfolios than its MV counterpart when the expected return of the portfolio is within the range $[0.5, 1.0]$, as can be seen in Figure 2. In cases where the randomly generated expected returns are within the range $[0.0, 0.5]$, the results (with respect to the diversification level) are just the opposite. As expected, the difference in concentrations is reduced when λ is close to one, since both of the portfolios converge to the GMV portfolio when $\lambda = 1$.

5. Conclusions

In this paper, an alternative to the traditional mean-variance (MV) strategy, named the mean squared variance (MSV) portfolio, is proposed. The novel proposed portfolio model overcomes the original limitation of the traditional model that expresses the two conflicting objectives in different units by including the expected return term in squared form. The proposed portfolio model is presented as a non-convex QP problem, which has been reformulated as a mixed-integer linear programming (MILP) problem to be globally solved by state-of-the-art MILP solvers.

Besides, a novel data-driven approach that is based on a Bayesian approach to estimate the optimal value of the hyper-parameter associated with the MV and MSV strategies is also proposed in this paper. The aim is to show that the formulation of the MSV problem

helps the global optimization procedure achieve a more competitive set of portfolio weights (in terms of the Sharpe ratio).

The proposed portfolio model was empirically tested on 8 portfolio problems with three different estimation windows (24 datasets were considered), reporting in most of the problems a very competitive performance of both the mean return and Sharpe ratio. In this research work, it is assumed that the current formulation of the model helps the tuning of the risk aversion profile parameter and, consequently, it helps in obtaining a final competitive performance.

Author Contributions: Conceptualization, F.F.-N. and M.C.-R.; Methodology, F.F.-N. and M.C.-R.; Software, F.F.-N.; Validation, F.F.-N.; Formal analysis, F.F.-N. and L.M.-N.; Data curation, L.M.-N. and T.M.-R.; Writing—original draft preparation, F.F.-N.; Writing—review and editing, L.M.-N., T.M.-R. and M.C.-R. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

BM	Book-to-market
GMR	Global maximum return
GMV	Global minimum variance
I	Investment
MILP	Mixed-integer linear programming
MR	Mean return
MSV	Mean squared variance
MV	Mean-variance
OP	Operating profitability
QP	Quadratic programming
SR	Sharpe ratio

References

1. Markowitz, H. Portfolio selection. *J. Financ.* **1952**, *7*, 77–91.
2. Markowitz, H. Mean–variance approximations to expected utility. *Eur. J. Oper. Res.* **2014**, *234*, 346–355. [CrossRef]
3. Zhou, R.; Palomar, D.P. Understanding the Quintile Portfolio. *IEEE Trans. Signal Process.* **2020**, *68*, 4030–4040. [CrossRef]
4. Coqueret, G. Diversified minimum-variance portfolios. *Ann. Financ.* **2015**, *11*, 221–241. [CrossRef]
5. Maillet, B.; Tokpavi, S.; Vaucher, B. Global minimum variance portfolio optimisation under some model risk: A robust regression-based approach. *Eur. J. Oper. Res.* **2015**, *244*, 289–299. [CrossRef]
6. Bodnar, T.; Parolya, N.; Schmid, W. Estimation of the global minimum variance portfolio in high dimensions. *Eur. J. Oper. Res.* **2018**, *266*, 371–390. [CrossRef]
7. Lim, A.E.; Zhou, X.Y. Mean-variance portfolio selection with random parameters in a complete market. *Math. Oper. Res.* **2002**, *27*, 101–120. [CrossRef]
8. Yin, G.; Zhou, X.Y. Markowitz's mean-variance portfolio selection with regime switching: From discrete-time models to their continuous-time limits. *IEEE Trans. Autom. Control* **2004**, *49*, 349–360. [CrossRef]
9. Usta, I.; Kantar, Y.M. Mean-variance-skewness-entropy measures: A multi-objective approach for portfolio selection. *Entropy* **2011**, *13*, 117–133. [CrossRef]
10. DeMiguel, V.; Garlappi, L.; Uppal, R. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Rev. Financ. Stud.* **2009**, *22*, 1915–1953. [CrossRef]
11. Tu, J.; Zhou, G. Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies. *J. Financ. Econ.* **2011**, *99*, 204–215. [CrossRef]
12. Li, H.; Huang, Q.; Wu, B. Improving the naive diversification: An enhanced indexation approach. *Financ. Res. Lett.* **2020**, 101661. [CrossRef]
13. Schmidt, A.B. Managing portfolio diversity within the mean variance theory. *Ann. Oper. Res.* **2019**, *282*, 315–329. [CrossRef]

14. Cai, H.; Schmidt, A.B. Comparing mean–variance portfolios and equal-weight portfolios for major US equity indexes. *J. Asset Manag.* **2020**, *21*, 326–332. [CrossRef]

15. Bird, R.; Tippett, M. Note—naive diversification and portfolio risk—A note. *Manag. Sci.* **1986**, *32*, 244–251. [CrossRef]

16. Schreiner, J. Portfolio revision: A turnover-constrained approach. *Financ. Manag.* **1980**, *9*, 67–75. [CrossRef]

17. Kourtis, A. A Stability Approach to Mean-Variance Optimization. *Financ. Rev.* **2015**, *50*, 301–330. [CrossRef]

18. Best, M.J.; Grauer, R.R. On the sensitivity of mean-variance-efficient portfolios to changes in asset means: Some analytical and computational results. *Rev. Financ. Stud.* **1991**, *4*, 315–342. [CrossRef]

19. Chopra, V.K.; Ziemba, W.T. The effect of errors in means, variances, and covariances on optimal portfolio choice. In *Handbook of the Fundamentals of Financial Decision Making: Part I*; World Scientific: London, UK, 2013; pp. 365–373.

20. Palczewski, A.; Palczewski, J. Theoretical and empirical estimates of mean–variance portfolio sensitivity. *Eur. J. Oper. Res.* **2014**, *234*, 402–410. [CrossRef]

21. Sharpe, W.F. Capital asset prices: A theory of market equilibrium under conditions of risk. *J. Financ.* **1964**, *19*, 425–442.

22. Sharpe, W.F. *Portfolio Theory and Capital Markets*; McGraw-Hill College: New York, NY, USA, 1970.

23. Bailey, D.H.; Lopez de Prado, M. The Sharpe ratio efficient frontier. *J. Risk* **2012**, *15*, 3–44. [CrossRef]

24. Sharpe, W.F. The sharpe ratio. *J. Portf. Manag.* **1994**, *21*, 49–58. [CrossRef]

25. Xia, W.; Vera, J.C.; Zuluaga, L.F. Globally solving nonconvex quadratic programs via linear integer programming techniques. *INFORMS J. Comput.* **2020**, *32*, 40–56. [CrossRef]

26. Mockus, J. *Bayesian Approach to Global Optimization: Theory and Applications*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012; Volume 37.

27. Giannessi, F.; Tomasin, E. Nonconvex quadratic programs, linear complementarity problems, and integer linear programs. In Proceedings of the IFIP Technical Conference on Optimization Techniques, Rome, Italy, 7–11 May 1973; Springer: Berlin/Heidelberg, Germany, 1973; pp. 437–449.

28. Bergmeir, C.; Benítez, J.M. On the use of cross-validation for time series predictor evaluation. *Inf. Sci.* **2012**, *191*, 192–213. [CrossRef]

29. Black, A.J.; McMillan, D.G. Non-linear predictability of value and growth stocks and economic activity. *J. Bus. Financ. Account.* **2004**, *31*, 439–474. [CrossRef]

30. Swinkels, L.; Tjong-A-Tjoe, L. Can mutual funds time investment styles? *J. Asset Manag.* **2007**, *8*, 123–132. [CrossRef]

31. Demšar, J. Statistical comparisons of classifiers over multiple data sets. *J. Mach. Learn. Res.* **2006**, *7*, 1–30.

32. Friedman, M. A comparison of alternative tests of significance for the problem of m rankings. *Ann. Math. Stat.* **1940**, *11*, 86–92. [CrossRef]

33. Bender, J.; Briand, R.; Nielsen, F.; Stefek, D. Portfolio of risk premia: A new approach to diversification. *J. Portf. Manag.* **2010**, *36*, 17–25. [CrossRef]

34. Liesiö, J.; Xu, P.; Kuosmanen, T. Portfolio Diversification based on Stochastic Dominance under Incomplete Probability Information. *Eur. J. Oper. Res.* **2020**, *286*, 755–768. [CrossRef]

35. Cesarone, F.; Scozzari, A.; Tardella, F. An optimization–diversification approach to portfolio selection. *J. Glob. Optim.* **2020**, *76*, 245–265. [CrossRef]