Journal of the Operational Research Society

ISSN: 0160-5682 (Print) 1476-9360 (Online) Journal homepage: www.tandfonline.com/journals/tjor20

Sparse and robust portfolio selection via semi-definite relaxation

Yongjae Lee, Min Jeong Kim, Jang Ho Kim, Ju Ri Jang & Woo Chang Kim

To cite this article: Yongjae Lee, Min Jeong Kim, Jang Ho Kim, Ju Ri Jang & Woo Chang Kim (2020) Sparse and robust portfolio selection via semi-definite relaxation, Journal of the Operational Research Society, 71:5, 687-699, DOI: 10.1080/01605682.2019.1581408

To link to this article: https://doi.org/10.1080/01605682.2019.1581408

Published online: 14 Jul 2019.

JOURNAL OF THE OPERATIONAL RESEARCH SOCIETY
2020, VOL. 71, NO. 5, 687–699
https://doi.org/10.1080/01605682.2019.1581408

ORIGINAL ARTICLE

Sparse and robust portfolio selection via semi-definite relaxation

Yongjae Lee[a], Min Jeong Kim[b], Jang Ho Kim[c], Ju Ri Jang[d] and Woo Chang Kim[e]

[a]School of Management Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan, South Korea; [b]Investment Policy Division, National Pension Research Institute of Korea, South Korea; [c]Industrial and Management Systems Engineering, Kyung Hee University, Yongin-si, Gyeonggi-do, South Korea; [d]Financial Business Division, NICE Information Service, South Korea; [e]Industrial and Systems Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea

ABSTRACT
In investment management, especially for automated investment services, it is critical for portfolios to have a manageable number of assets and robust performance. First, portfolios should not contain too many assets in order to reduce the management fees, transaction costs, and taxes. Second, portfolios should be robust as investment environments change rapidly. In this study, therefore, we propose two convex portfolio selection models that provide portfolios that are sparse and robust. We first perform semi-definite relaxation to develop a sparse mean-variance portfolio selection model, and further extend the model by using L_2-norm regularization and worst-case optimization to formulate two sparse and robust portfolio selection models. Empirical analyses with historical stock returns demonstrate the effectiveness of the proposed models in forming sparse and robust portfolios.

ARTICLE HISTORY
Received 4 January 2018
Accepted 23 January 2019

KEYWORDS
Portfolio selection; sparse portfolio; L_2-norm regularization; robust optimization; semi-definite relaxation; robo-advisor

1. Introduction

Due to the remarkable development of information and communication technology, a large part of financial business is now being automated and portfolio management is no exception. In automated portfolio management, portfolios are constructed and rebalanced by algorithms based on investor information, such as current savings, future income, and risk appetite.[1] Portfolios constructed by such automated platforms should exhibit the following two characteristics. First, it should be robust, as there is no manager to instantly reallocate the portfolio when the market environment changes. This is particularly important as the financial markets are highly volatile. Second, the portfolio should be sparse in order to reduce the transaction costs of the initial construction and future reallocations. Controlling the cardinality of a portfolio (i.e., the number of assets invested) is especially important since it is known to be effective at lower total cost of investments that are not taken into account by mean-variance portfolios (Wilding, 2003).

Unfortunately, the classical mean-variance model developed by Markowitz (1952) often produces portfolios that are not suitable for automated portfolio management. Most importantly, the mean-variance model is known to be highly sensitive to the input parameters such as expected returns and covariances (Best & Grauer, 1991; Chopra & Ziemba, 1993). Furthermore, even though the portfolio weights obtained from the mean-variance model are, in general, highly concentrated on a few assets (Garlappi, Uppal, & Wang, 2007), weights on other assets are still non-zero. In other words, traditional mean-variance portfolios are definitely not sparse, and moreover, the cardinality of mean-variance portfolios are almost the same as the number of candidate assets, as will be demonstrated in Section 4.

Numerous variations of the Markowitz's mean-variance model have been developed to reduce its shortcomings. In terms of improving portfolio robustness, some studies have sought ways to control the sensitivity of portfolios to estimation errors by improving the parameter estimation process itself through shrinkage methods or Bayesian approaches (Black & Litterman, 1991, 1992; Jorion, 1986; Ledoit & Wolf, 2003); the resulting portfolios are more robust because the input estimations are more robust. One of the most prominent methodologies to reduce the sensitivity of the mean-variance model is robust portfolio optimization (Goldfarb & Iyengar, 2003; Halldórsson & Tütüncü, 2003; Kim, Kim, Ahn, & Fabozzi, 2013; Kim, Kim, & Fabozzi, 2014; Lobo & Boyd, 2000; Tütüncü & Koenig, 2004). Robust portfolio optimization forms a portfolio by defining an uncertainty set of all possible realizations of uncertain input parameters and reformulates the problem to optimize the worst-case. In addition, there have been attempts to reduce estimation errors of the Markowitz model by introducing additional constraints on portfolio weights (Chopra, 1993; DeMiguel, Garlappi, Nogales, & Uppal, 2009;

CONTACT Jang Ho Kim janghokim@khu.ac.kr Industrial and Management Systems Engineering, Kyung Hee University, Yongin-si, Gyeonggi-do, South Korea

Frost & Savarino, 1988). For example, DeMiguel et al. (2009) constrained the L_2-norm of the portfolio weight to be smaller than a given threshold to make portfolios more stable. For multi-stage portfolio problems, stochastic programming is often utilized to increase robustness (Birge & Louveaux, 2011; Dantzig & Infanger, 1993). More recently, stochastic dominance approaches are also shown to improve portfolio robustness (Bruni, Cesarone, Scozzari, & Tardella, 2017; Roman, Mitra, & Zverovich, 2013).

As for limiting portfolio cardinality, most sparse portfolio methods can be categorized into two streams. The first group of studies show that the cardinality of portfolios can be reduced indirectly by including L_1-norm, or $L_{p,0<p<1}$-norm regularization on portfolio weights (Brodie, Daubechies, De Mol, Giannone, & Loris, 2009; Chen, Li, Tolman, Wang, & Ye, 2013).[2] Measuring portfolio risk using absolute deviation, or L_1-norm, can be also used to form portfolios with reduced number of assets (Konno & Yamazaki, 1991; Konno & Yamamoto, 2005). The second stream of studies for achieving sparsity follows a direct approach by adding cardinality constraints to portfolio formulations. However, the cardinality-constraint makes the problem computationally intractable since the cardinality-constrained portfolio problem is a special case of general cardinality-constrained quadratic programming, which is known to be NP-hard (Bienstock, 1996; Gao & Li, 2013a, 2013b; Garey & Johnson, 1979), and thus, most researchers have focused on developing efficient algorithms to solve the problem (Bertsimas & Shioda, 2009; Bienstock, 1996; Borchers & Mitchell, 1994, 1997; Chang, Meade, Beasley, & Sharaiha, 2000; Gao & Li, 2011, 2013a, 2013b; Li, Sun, & Wang, 2006; Maringer & Kellerer, 2003). On the other hand, d'Aspremont, El Ghaoui, Jordan, and Lanckriet (2007) developed a method that converts a cardinality-constrained principal component analysis (PCA) problem into a convex optimization problem via semi-definite relaxation. Kim, Lee, Kim, and Kim (2016) applied the semi-definite relaxation method to the sparse portfolio selection problem, particularly, a Sharpe ratio maximization problem.

As summarized above, robust portfolio selection and sparse portfolio selection are extensively studied subjects, but they have mostly been studied separately. In this work, therefore, we propose two portfolio selection methods that construct sparse and robust portfolios efficiently. More specifically, we apply the sparse portfolio selection method of Kim et al. (2016), which took the semi-definite relaxation approach to a cardinality-constrained Sharpe ratio maximization problem, to two robust portfolio selection methods: robust portfolio optimization and L_2-norm regularization (DeMiguel et al., 2009). In order to do so, we first apply the semi-definite relaxation method of Kim et al. (2016) to the traditional mean-variance model with a cardinality constraint, and we further extend it to robust portfolio optimization and L_2-norm regularization.

Although portfolio sparsity and robustness are often analyzed separately, a holistic understanding is critical for managing portfolios. There is a tradeoff because reduced number of assets will form concentrated portfolios that are highly sensitive to the invested assets whereas a diversified portfolio will be robust but include many assets. Our objective is to propose formulations that can find a balance between the two. As mentioned above, it is particularly interesting to analyze norm-constrained portfolio optimization and robust optimization in conjunction because some norm-constraints are also observed to reduce portfolio sensitivity. Xu, Caramanis, and Mannor (2009) discuss norm constraints and robustness, and Gotoh and Takeda (2011) expand the analysis to portfolio optimization. In Gotoh and Takeda, they derive how norm-constrained formulations can be interpreted as robust counterparts of value-at-risk or conditional value-at-risk minimizations. Olivares-Nadal and DeMiguel (2018) use transaction costs expressed as norms to express the relationship between mean-variance problems with norm-constraints and robust mean-variance problems. In this work, we derive formulations that explicitly control the number of assets and portfolio robustness by adding cardinality constraints to robust optimization and norm-constrained portfolio problems. The proposed model allows controlling both sparsity and robustness and also analyzing the tradeoff between the two.

The rest of the article is organized as follows. Section 2 develops the sparse mean-variance portfolio selection model based on the semi-definite relaxation method of Kim et al. (2016). In Section 3, we extend the model to two sparse and robust portfolio selection models by applying L_2-norm regularization and robust portfolio optimization, respectively. In Section 4, our numerical experiments confirm that our models successfully construct portfolios that are both sparse and robust. Finally, Section 5 concludes the study.

2. Sparse Mean-Variance portfolio selection model

In this section, we develop a sparse mean-variance portfolio selection model by applying the semi-definite relaxation technique of d'Aspremont et al. (2007) and Kim et al. (2016) to the traditional mean-variance model with a cardinality constraint.

Note that d'Aspremont et al. (2007) and Kim et al. (2016) have applied the semi-definite relaxation technique to PCA and Sharpe ratio maximization problems, respectively. While Kim et al. (2016) propose a convex formulation for computing the tangent portfolio (with maximum Sharpe ratio) with cardinality control, we derive a convex relaxation of a cardinality-constrained mean-variance portfolio problem that finds efficient portfolios for various risk preferences.

Consider n risky assets with the expected returns $\mu \in \mathbb{R}^n$, and the covariance of asset returns $\Sigma \in \mathbb{R}^{n\times n}$, and a portfolio weight vector $w \in \mathbb{R}^n$. In addition, there is a constant $\lambda \in \mathbb{R}$, which represents the risk preference of an investor, and a vector $\imath \in \mathbb{R}^n$, which is a vector of ones. Then, the traditional mean-variance model can be represented as follows.

$$\begin{aligned} \min_{w} \quad & w^T\Sigma w - \lambda \mu^T w \\ \text{s.t.} \quad & \imath^T w = 1 \end{aligned} \tag{MV.1}$$

We first add a cardinality constraint, which restricts the number of invested assets to be less than or equal to a positive integer k, to the mean-variance model.

$$\begin{aligned} \min_{w} \quad & w^T\Sigma w - \lambda \mu^T w \\ \text{s.t.} \quad & \imath^T w = 1 \\ & \mathbf{Card}(w) \le k \end{aligned} \tag{MV.2}$$

Here, $\mathbf{Card}(x)$ is the cardinality of any vector $x \in \mathbb{R}^n$. As (MV.2) becomes an NP-Hard problem (Garey & Johnson, 1979), we relax the problem into a semi-definite programming (SDP) problem, which is convex. Before applying the relaxation, we reformulate (MV.2) into an equivalent problem with a positive semi-definite variable W in order to apply the semi-definite relaxation technique.

Proposition 1. The optimization problem (MV.2) is equivalent to the following problem (MV.3) with $w^* {w^*}^T = W^*$, where w^* and W^* are optimal solutions of (MV.2) and (MV.3), respectively.

$$\begin{aligned} \min_{W} \quad & \mathbf{Tr}(\Sigma W) - \lambda\left(\imath^T W \mu\right) \\ \text{s.t.} \quad & \mathbf{Tr}(\mathbb{I} W) = 1 \\ & \mathbf{Card}(W) \le k^2 \\ & \mathbf{Rank}(W) = 1 \\ & W \in \mathbf{S}_n^+ \end{aligned} \tag{MV.3}$$

Here, $\mathbf{S}_n^+$ indicates the set of $n \times n$-dimensional positive semi-definite matrices, and $\mathbb{I} = \imath\imath^T$. Also, $\mathbf{Tr}(X)$ and $\mathbf{Rank}(W)$ refer to the trace and the rank of a matrix X, respectively.

Proof. A feasible solution $W \in \mathbf{S}_n^+$ can be represented as $W = ww^T$ because a positive semi-definite matrix with rank equal to one has a unique eigenvector. Then, $\mathbf{Tr}(\mathbb{I}W) = 1$ and $\mathbf{Card}(W) \le k^2$ are equivalent to $\imath^T w = 1$ and $\mathbf{Card}(w) \le k$, respectively. Finally, note that $\mathbf{Tr}(\Sigma W) = w^T\Sigma w$ and $\imath^T W\mu = \imath^T ww^T \mu = \left(\imath^T w\right)\left(w^T\mu\right) = 1 \times \left(w^T \mu\right) = w^T\mu$. □

Since (MV.2) is still non-convex because of the cardinality constraint, we apply the semi-definite relaxation method used in d'Aspremont et al. (2007) and Kim et al. (2016). For any vector $w \in \mathbb{R}^n$, the inequality $||w||_1 \le \sqrt{n}||w||_2$ holds, and it can be tightened as $||w||_1 \le \sqrt{k}||w||_2$ if $\mathbf{Card}(w) \le k$. Therefore, the cardinality constraint $\mathbf{Card}(W) \le k^2$ can be relaxed into a weaker but convex constraint $\imath^T|W|\imath \le k\mathbf{Tr}(W)$, where $|\cdot|$ is the element-wise absolute value operator. Thus, by relaxing the cardinality constraint and dropping the rank constraint, we arrive at the following convex problem.

$$\begin{aligned} \min_{W} \quad & \mathbf{Tr}(\Sigma W) - \lambda\left(\imath^T W \mu\right) \\ \text{s.t.} \quad & \mathbf{Tr}(\mathbb{I} W) = 1 \\ & \imath^T|W|\imath \le k\mathbf{Tr}(W) \\ & W \in \mathbf{S}_n^+ \end{aligned} \tag{MV.4}$$

The optimal solution of the relaxed problem (MV.4) should be identical to the optimal solution of (MV.3), if the former has a rank equal to one. Otherwise, we should approximate a solution by rescaling the eigenvector, which corresponds to the largest eigenvalue, of the optimal solution of (MV.4). Furthermore, as the problem is relaxed, a positive number k does not mean the maximum cardinality but it rather can be regarded as a sparsity control parameter in the sense that smaller k's lead to more sparse portfolios.

3. Sparse and robust portfolio selection models

Section 2, as well as Kim et al. (2016), address applying cardinality constraints to mean-variance efficient portfolios. In Section 3, we extend the derivation to applying cardinality control to robust portfolio formulations. We develop two versions of sparse and robust portfolio selection models based on the semi-definite relaxation method. More specifically, we extend the sparse mean-variance portfolio selection model developed in the previous section to construct portfolios that are more robust by taking either 1) L_2-norm regularization approach or 2) worst-case (or robust) optimization approach.

3.1. L_2-norm regularization approach

DeMiguel et al. (2009) showed that a L_2-norm regularized global minimum variance portfolio exhibits better out-of-sample performance than that of a mean-variance optimal portfolio in the presence of estimation errors. Their L_2-norm regularized Markowitz model (L2.1) that bounds the L_2-norm ($||\cdot||_2$) of an optimal portfolio weight to be less than or equal to $\delta \in \mathbb{R}^+$ can be formulated as follows.

$$\begin{aligned} \min_w \quad & w^T \Sigma w - \lambda \mu^T w \qquad \text{(L2.1)} \\ \text{s.t.} \quad & \imath^T w = 1 \\ & ||w||_2 \le \delta \end{aligned}$$

Now we restrict the maximum number of assets under investment to be less than or equal to a positive integer k.

$$\begin{aligned} \min_w \quad & w^T \Sigma w - \lambda \mu^T w \qquad \text{(L2.2)} \\ \text{s.t.} \quad & \imath^T w = 1 \\ & ||w||_2 \le \delta \\ & \mathbf{Card}(w) \le k \end{aligned}$$

Again, the problem becomes NP-hard, and thus, we relax the problem into a convex optimization problem via the semi-definite relaxation method. First, Proposition 1 provides an equivalent formulation of (L2.2) that has a positive semi-definite variable W as a decision variable, but the equivalent problem is still non-convex at this stage.

Proposition 2. The optimization problem (L2.2) is equivalent to the following problem (L2.3) with $w^* w^{*T} = W^*$, where w^* and W^* are optimal solutions of (L2.2) and (L2.3), respectively.

$$\begin{aligned} \min W \quad & \mathbf{Tr}(\Sigma W) - \lambda (\imath^T W \mu) \qquad \text{(L2.3)} \\ \text{s.t.} \quad & \mathbf{Tr}(\mathbb{I} W) = 1 \\ & \mathbf{Tr}(W) \le \delta^2 \\ & \mathbf{Card}(W) \le k^2 \\ & \mathbf{Rank}(W) = 1 \end{aligned}$$

Here, $\mathbf{S}_n^+$ indicates the set of $n \times n$-dimensional positive semi-definite matrices, and $\mathbb{I} = \imath\imath^T$.

Proof. Note that $\mathbf{Tr}(W) = \mathbf{Tr}(ww^T) = ||w||_2^2$. Thus, the constraint $\mathbf{Tr}(W) \le \delta^2$ in (L2.3) is equivalent to the constraint $||w||_2 \le \delta$ in (L2.2). For the objective functions and other constraints, see the proof of Proposition 1. □

Finally, we have the following convex problem (L2.4) by relaxing the cardinality constraint and dropping the rank constraint as we did in Section 2.

$$\begin{aligned} \min_W \quad & \mathbf{Tr}(\Sigma W) - \lambda (\imath^T W \mu) \qquad \text{(L2.4)} \\ \text{s.t.} \quad & \mathbf{Tr}(\mathbb{I} W) = 1 \\ & \mathbf{Tr}(W) \le \delta^2 \\ & \imath^T |W| \imath \le k \mathbf{Tr}(W) \\ & W \in \mathbf{S}_n^+ \end{aligned}$$

3.2. Worst-case optimization approach

The worst-case optimization (or robust optimization) approach mitigates the effect of parameter uncertainties by considering the worst-case scenario. In this work, we focus on robust portfolio optimization under the assumption that uncertainties exist only in expected returns, because Chopra and Ziemba (1993) demonstrated that the mean-variance model is much more sensitive to expected returns than variances and covariances. While there are many ways to model uncertainty sets (or sets of possible values of uncertain parameters), we choose to employ ellipsoidal uncertainty sets that have been extensively studied in the robust optimization domain. An ellipsoidal uncertainty set for an expected return vector μ can be written as

$$\mathrm{U}_\gamma(\hat{\mu}) = \{\mu \in \mathbb{R}^n | \ (\mu - \hat{\mu})^{\mathrm{T}} \Sigma_\mu^{-1} (\mu - \hat{\mu}) \ \le \gamma^2\}$$

where $\hat{\mu} \in \mathbb{R}^n$ is the center of the ellipsoid, which can be normally set as an empirical expected returns vector, $\gamma \in \mathbb{R}$ determines the size of the ellipsoid, and $\Sigma_\mu \in \mathbb{R}^{n \times n}$ is the covariance matrix of estimation errors. A worst-case mean-variance portfolio selection model with an ellipsoidal uncertainty set can be written as follows.

$$\min_{\{w \in \mathbb{R}^n | \imath^T w = 1\}} \ \max_{\mu \in \mathrm{U}_\gamma(\hat{\mu})} \ w^T \Sigma w - \lambda \mu^T w$$

The above minimax problem can be transformed into an equivalent second-order cone programming (SOCP) problem (RO.1).

$$\begin{aligned} \min_w \quad & w^T \Sigma w - \lambda \left(\hat{\mu}^T w - \gamma \sqrt{w^T \Sigma_\mu w} \right) \qquad \text{(RO.1)} \\ \text{s.t.} \quad & \imath^T w = 1 \end{aligned}$$

We further transform the problem into an equivalent quadratically constrained quadratic programming (QCQP) problem, as we felt that it is easier to apply the semi-definite relaxation method to QCQP formulations than SOCP formulations. Note that the term $\lambda \gamma \sqrt{w^T \Sigma_\mu w}$ in the objective function can be dropped into a constraint $w^T \Sigma_\mu w \le \varepsilon$ for some $\varepsilon \in \mathbb{R}_+$. An equivalent QCQP formulation (RO.1') is given below.

$$\begin{aligned} \min_w \quad & w^T \Sigma w - \lambda (\hat{\mu}^T w) \qquad \text{(RO.1)} \\ \text{s.t.} \quad & \imath^T w = 1 \\ & w^T \Sigma_\mu w \le \varepsilon \end{aligned}$$

A cardinality-constrained version of (RO.1') is as follows.

$$\begin{aligned} \min_w \quad & w^T \Sigma w - \lambda (\hat{\mu}^T w) \qquad \text{(RO.2)} \\ \text{s.t.} \quad & \imath^T w = 1 \\ & w^T \Sigma_\mu w \le \varepsilon \\ & \mathbf{Card}(w) \le k \end{aligned}$$

Again, the problem (RO.2) becomes non-convex, and thus, we apply the semi-definite relaxation method to (RO.2) as in Sections 2 and 3.1. We first reformulate the problem (RO.2) with a positive semi-definite variable W.

Proposition 3. The optimization problem (RO.2) is equivalent to the following problem (RO.3) with $w^* w^{*T} = W^*$, where w^* and W^* are optimal solutions of (RO.2) and (RO.3), respectively.

$$\begin{aligned} \min_{W} \quad & \mathbf{Tr}(\Sigma W) - \lambda(\iota^T W \hat{\mu}) \qquad \text{(RO.3)} \\ \text{s.t.} \quad & \mathbf{Tr}(\mathbb{I}W) = 1 \\ & \mathbf{Tr}(\Sigma_\mu W) \le \varepsilon \\ & \mathbf{Card}(W) \le k^2 \\ & \mathbf{Rank}(W) = 1 \\ & W \in \mathbf{S}_n^+ \end{aligned}$$

Proof. Note that $\mathbf{Tr}(\Sigma_\mu W) = w^T \Sigma_\mu w$, and thus $\mathbf{Tr}(\Sigma_\mu W) \le \varepsilon$ is equivalent to $w^T \Sigma_\mu w \le \varepsilon$. For the objective functions and other constraints, see the proof of Proposition 1. □

We arrive at the following convex formulation (RO.4) by weakening the cardinality constraint and ignoring the rank constraint as in Sections 2 and 3.1.

$$\begin{aligned} \min_{W} \quad & \mathbf{Tr}(\Sigma W) - \lambda(\iota^T W \mu) \qquad \text{(RO.4)} \\ \text{s.t.} \quad & \mathbf{Tr}(\mathbb{I}W) = 1 \\ & \mathbf{Tr}(\Sigma_\mu W) \le \varepsilon \\ & \iota^T |W| \iota \le k \mathbf{Tr}(W) \\ & W \in \mathbf{S}_n^+ \end{aligned}$$

4. Numerical tests

This section empirically demonstrates the validity of the sparse and robust portfolio selection models developed in Section 3 using historical stock returns data. We first determine whether the portfolios constructed by the proposed sparse portfolio selection models (MV.4), (L2.4), and (RO.4) are indeed sparse. Next, we check whether the robust models (L2.4) and (RO.4) can produce portfolios that are more robust than portfolios constructed by (MV.4).

We use Kenneth French's 49 industry portfolios, which divide all stocks listed on NYSE, AMEX, and NASDAQ into 49 portfolios in terms of their industry classifications.[3] The use of industries is due to concerns on data accessibility (using a subset of all stocks would require showing whether the subset is a good representation of the entire market), effectiveness of analyses (e.g., the results will be more sensitive to data if we were to analyze a portfolio of cardinality 10 when there are thousands of candidate stocks), and computational cost. On the other hand, industry portfolios contain all stocks in the U.S. market and divides the stock market into representative industries.

Table 1. List of tested portfolios.

Name	Description
MV	(MV.4)
L2[1]	(L2.4) with $\delta = 0.9 \parallel w^*_{(MV.1)} \parallel_2$
L2[2]	(L2.4) with $\delta = 0.8 \parallel w^*_{(MV.1)} \parallel_2$
L2[3]	(L2.4) with $\delta = 0.7 \parallel w^*_{(MV.1)} \parallel_2$
RO[1]	(RO.4) with $\varepsilon = 0.9 * w^T \Sigma_\mu w$
RO[2]	(RO.4) with $\varepsilon = 0.8 * w^T \Sigma_\mu w$
RO[3]	(RO.4) with $\varepsilon = 0.7 * w^T \Sigma_\mu w$

The empirical tests are conducted on an annual basis for a 40-year long period from 1975 to 2014. More specifically, our empirical tests use the following procedures:

For each year from 1975 to 2014:

Step 1: Estimate annual expected returns μ and annual covariance matrix Σ from historical stock returns data
Step 2: Solve models without cardinality constraints (MV.1), (L2.1), and (RO.1')
Step 3: Solve cardinality-constrained models (MV.4), (L2.4), and (RO.4) with various values of k, λ, δ and ε
Step 4: Obtain approximate solutions for (MV.2), (L2.2), and (RO.2) by finding leading eigenvectors of solutions of (MV.4), (L2.4), and (RO.4) and rescaling them to have elements sums equal to one
Step 5: Calculate portfolio performances
Step 6: Repeat Steps 1 to 5 from year 1975 to 2014
Step 7: Calculate average values of portfolio performances over the test period

To explain Step 3 in more detail, we test seven portfolios with various values of sparsity control parameter k and risk preference parameter λ, and the tested portfolios are listed in Table 1 with detailed descriptions. Here, $\parallel \cdot \parallel_2$ indicates the L_2-norm operator, and $w^*_{(MV.1)}$ indicates the optimal weight vector obtained from the original mean-variance model (MV.1).

We construct these portfolios with various $k = \{10,\ 12.5,\ 15,\ 17.5,\ 20,\ 22.5,\ 25,\ \text{No card.}\}$ and $\lambda = \{0.001,\ 0.005,\ 0.01,\ 0.05\}$[4] where 'No card.' indicates that the corresponding portfolios are constructed by models without a cardinality constraint. In addition, for portfolios constructed from (RO.4), we let the covariance matrix of estimation errors $\Sigma_\mu = \frac{1}{T} * diag(\Sigma)$, where Σ is the covariance matrix of stock returns, T is the sample size, and $diag(A)$ denotes a function that returns a matrix that has the diagonal elements of matrix A and filled with zeros for the remaining off-diagonal elements. From now on, we let MV, L2, and RO denote all portfolios

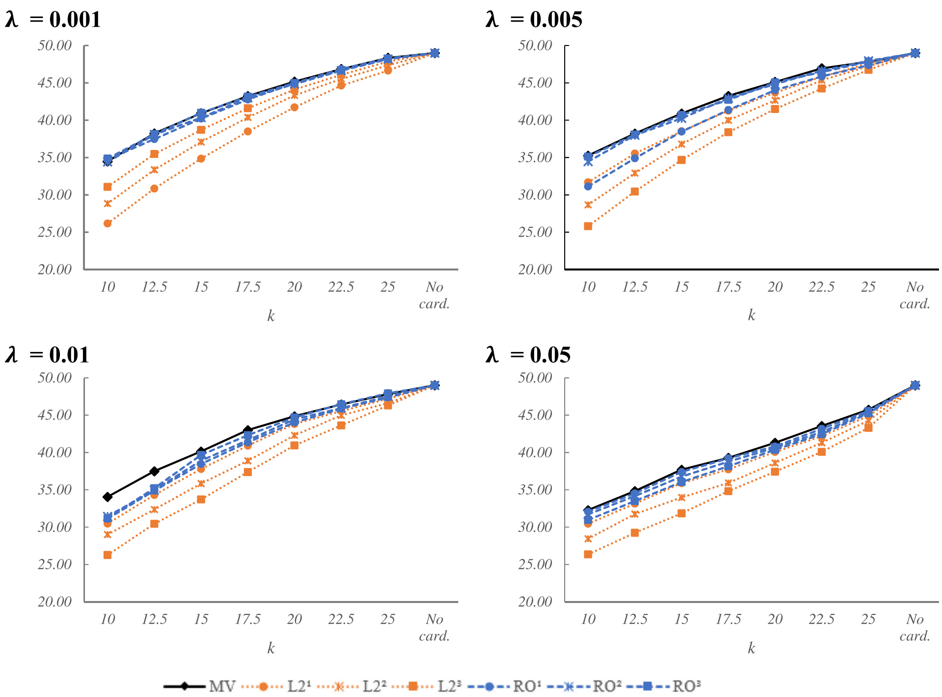

Figure 1. Average portfolio cardinalities of MV, L2, and RO.

constructed from traditional mean-variance models (with and without a cardinality constraint), L_2-norm regularized models (with and without a cardinality constraint), and worst-case optimization models (with and without cardinality constraint), respectively. We also note that for small values of k, some of the problems might become infeasible due to the small degree of freedom. Thus, such infeasible points are accordingly excluded.

4.1. Sparsity

Figure 1 clearly shows that all three models successfully construct sparse portfolios as all portfolios from cardinality-constrained models have smaller cardinalities than portfolios constructed without cardinality constraints. Note that we construct all portfolios considering 49 number of assets, and the portfolios without a cardinality constraint actually invest in almost all 49 assets. This justifies recent efforts to develop sparse portfolio selection models, including this study, as portfolios from sparse portfolio selection models invest in 5 to 25 fewer assets than portfolios from non-sparse models. As mentioned earlier, the parameter k's in the relaxed models do not indicate the maximum cardinalities of portfolios to be constructed, but rather they serve as sparsity control parameters for which smaller k's lead to smaller cardinalities.

4.2. Robustness

Now that we know the proposed models successfully construct sparse portfolios, we move on to examine if the robust models (L2.4) and (RO.4) can provide portfolios that are more robust than (MV.4). We investigate the robustness of portfolios in two ways. First, we check how risky the test portfolios are in terms of three of the most popular risk measures: return volatility, value-at-risk (VaR), and maximum drawdown (MDD). Second, we see if the portfolios continue to perform well when the investment environment changes. In other words, we compare the out-of-sample performances of the test portfolios in terms of various performance measures: rate of return, Sharpe ratio, and the above three risk measures (volatility, VaR, and MDD). Note that the out-of-sample performances of the portfolios constructed in year i are calculated using the returns of the portfolios in year $i+1$.

Before we present the results, it is important to note that portfolio constraints such as minimum required returns or maximum volatilities are not directly imposed when constructing the portfolios. Our models have risk preference parameter λ's instead of fixing minimum returns or maximum volatilities of portfolios. Thus, while all formulations express the objective using λ, the same value of λ does not result in portfolios with the same level of return or risk. Instead of focusing on the results for specific values of λ, a holistic interpretation and analysis over a range of return and risk levels are appropriate. Furthermore, measures of mean-variance efficiency such as the Sharpe ratio that account for the relationship between return and risk also provide valuable comparisons.

4.2.1. Sharpe ratio

Figure 2 shows the Sharpe ratios of the test portfolios. The Sharpe ratio would arguably be the most

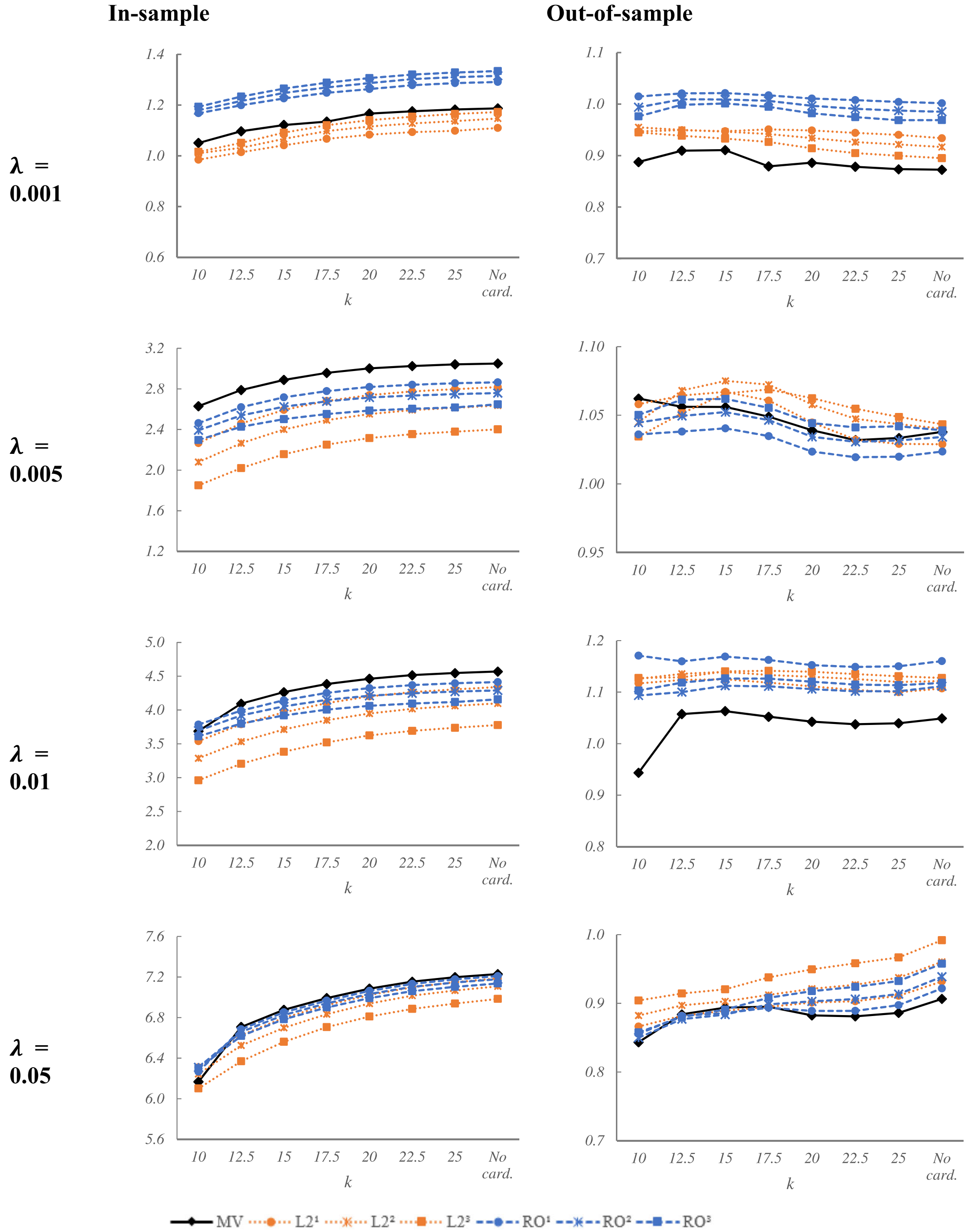

Figure 2. Average portfolio Sharpe ratios of MV, L2, and RO.

representative performance measure in mean-variance portfolio theory, as this single measure can show the mean-variance efficiency of a portfolio. In most cases, MV portfolios have the highest Sharpe ratios in-sample as they are constructed by the least constrained model. However, all L2 and RO portfolios show higher (or at least similar) Sharpe ratios out-of-sample than MV portfolios. These results clearly show that L2 and RO portfolios are more robust than MV portfolios.

4.2.2. Rate of return

MV portfolios, in general, have the largest rates of returns both in-sample and out-of-sample, but these are not of great importance as stated at the beginning of Section 4.2. It would be important to note that the ranges of differences between returns of MV and robust (L2 and RO) portfolios are 5% to 40% in-sample, but they become 1% to 7% out-of-sample. In other words, these results indicate that L2 and RO portfolios are more robust to the investment environment changes than MV portfolios. In-sample and out-of-sample annual returns of the test portfolios are presented in the Appendix.

4.2.3. Risk measures

Figures 3 and 4 summarize volatilities and VaR at the 95% confidence level of the test portfolios (annualized), respectively, which are the most common risk measures used in both academia and industry. With respect to all risk measures, regardless of in-sample or out-of-sample, MV portfolios are the most risky portfolios, except for some cases when the risk preference parameter λ is very small

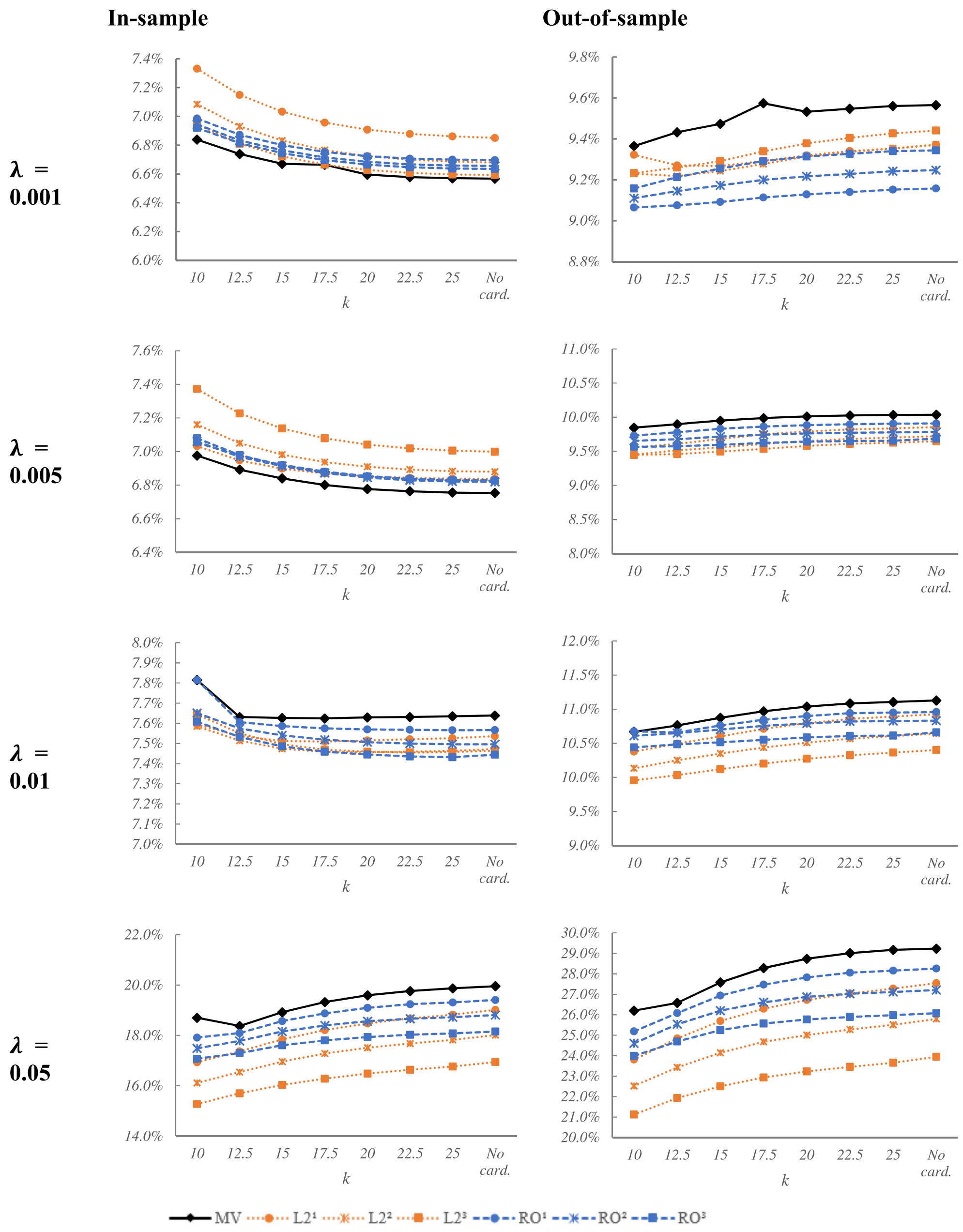

Figure 3. Average portfolio volatilities of MV, L2, and RO.

so that the MV model constructs conservative portfolios as well. Similar to the case of rates of returns, the differences in risk measures between MV and robust (L2 and RO) portfolios in out-of-sample results are larger than in in-sample results. This again shows that the L2 and RO models provide more robust portfolios than the MV model. Similar results are observed when measuring risk with maximum drawdowns (MDD), which are presented in the Appendix.

4.2.4. Further discussion on sparsity and robustness

It is clear from the in-sample results that there is a tradeoff between sparsity and robustness. Portfolios, regardless of MV, L2, or RO, show lower risk as their cardinality is increased. Moreover, higher cardinality leads to higher annual returns and, thus, higher Sharpe ratios. The out-of-sample tradeoffs are less evident since the difference caused by sparsity is often a few basis points per year. Nonetheless, out-of-sample patterns clearly demonstrate the robustness of L2 and RO. MV portfolios have higher volatility, VaR, and MDD when compared to the robust models, which confirms that the proposed models increase robustness while controlling portfolio cardinality.

As summarized above, the proposed models show clear strengths in achieving sparsity as well as robustness. One notable weakness of the formulations is that the cardinality constraint does not enforce the maximum number of assets but is a relaxed constraint that controls the sparsity level of the portfolio. However, a fixed limit on portfolio size is rarely necessary even when low cardinality is

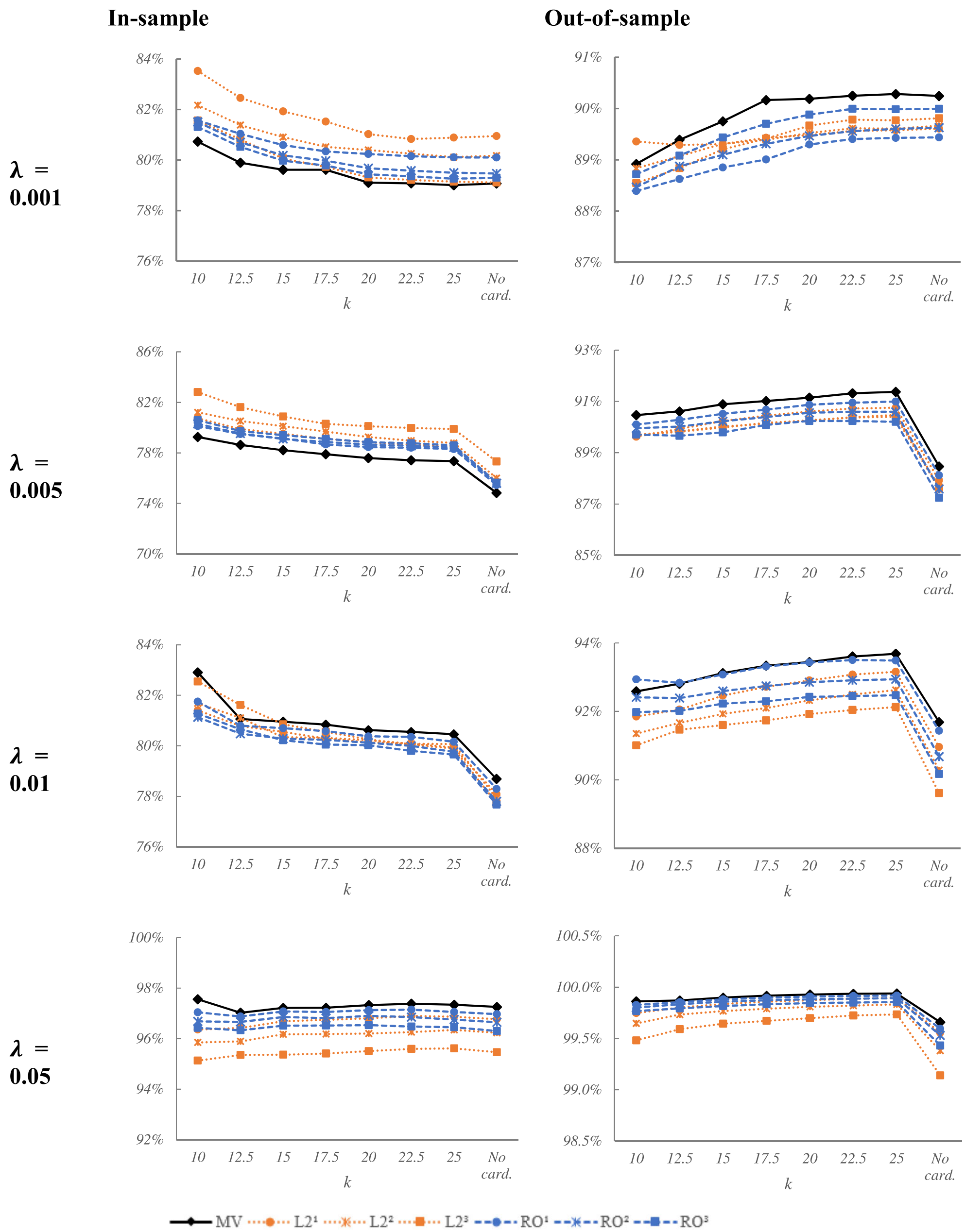

Figure 4. Average portfolio VaRs of MV, L2, and RO.

desired. For example, reducing cardinality with a relaxed constraint can help lower transaction costs, and construct index-tracking or mimicking portfolios.

Between the two approaches, L_2-norm regularization and worst-case optimization, no significant difference in in-sample and out-of-sample performance is observed. Especially, the out-of-sample performance tend to be steady for various parameter values. One observed difference is that the formulation with L_2-norm regularization seems to be slightly better at reducing portfolio cardinality for small values of k as shown in Figure 1.

We also note that the values of k, λ, δ, and ε would be set exogenously. This is similar to solving norm-constrained portfolio optimization in general, where the threshold parameter for portfolio norm can be calibrated based on investment objectives (DeMiguel et al., 2009). In our formulation, the values of k and λ will be dependent on the candidate assets and the investor preference. In particular, the target cardinality can be reached by adjusting the cardinality control parameter k as shown in Figure 1. Moreover, the values of δ and ε can be determined based on reference points as we have illustrated in Table 1; we set the value of δ as a percentage of the norm of a reference portfolio and the value of ε as a percentage of sample portfolio estimation error variance.

5. Conclusion

In this study, we developed two sparse and robust portfolio selection models that are both convex. A sparse mean-variance portfolio selection model was first derived using the semi-definite relaxation method by d'Aspremont et al. (2007) and Kim et al. (2016), and the model was further extended to arrive at two sparse and robust portfolio selection models using two methods, L_2-norm regularization and worst-case optimization. Empirical analyses confirmed that both the proposed models successfully construct sparse portfolios, and especially, that the two robust models provide portfolios that are more stable (or robust) than the sparse MV portfolios.

The main contribution of the proposed sparse and robust portfolio selection models is providing portfolios that are sparse and robust at the same time. The importance of portfolio robustness has almost become a tenet in investment, especially after the recent worldwide financial crisis. Sparse portfolios are gaining more attention nowadays due to the rise of robo-advisors (online platforms for automated private wealth management) that try to dramatically lower management costs, including transaction costs and taxes, to attract a large volume of non-high-net-worth individuals. Kim, Kwon, Lee, Kim, and Li (2019) noted that automated investment services should be able to provide at least quasi-real-time online interactions with customers in order for the customers to test whether or not their financial goals are feasible by trial and error; because there are no human experts to directly communicate with the customers. In this regard, our models are suitable because our models are convex programming problems that can be solved efficiently.

Notes

1. See Kim et al. (2019) for more detailed descriptions on automated portfolio managements.
2. Controlling portfolio cardinality can also be obtained by other approaches on measuring portfolio risk (e.g., Bruni, Cesarone, Scozzari, & Tardella, 2015).
3. Data available at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
4. The selection of the risk preference parameter λ is critical to the investment, and Fabozzi, Kolm, Pachamanova, and Focardi (2007) discuss in detail about the risk-aversion formulation of mean-variance model in Chapter 2. Note that the meaning of λ is exactly the opposite in Fabozzi et al. (2007), since λ is attached to the variance term in Fabozzi et al. (2007) whereas it is attached to the mean term throughout this study.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2018R1C1B6004271) and the New Faculty Fund (1.180089.01) of UNIST (Ulsan National Institute of Science and Technology).

References

Bertsimas, D., & Shioda, R. (2009). Algorithm for cardinality-constrained quadratic optimization. *Computational Optimization and Applications*, *43*(1), 1–22. doi:10.1007/s10589-007-9126-9

Best, M. J., & Grauer, R. R. (1991). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. *Review of Financial Studies*, *4*(2), 315–342. doi:10.1093/rfs/4.2.315

Bienstock, D. (1996). Computational study of a family of mixed-integer quadratic programming problems. *Mathematical Programming*, *74*(2), 121–140. doi:10.1007/BF02592208

Birge, J. R., & Louveaux, F. (2011). *Introduction to stochastic programming*. New York: Springer Science & Business Media.

Black, F., & Litterman, R. (1992). Global portfolio optimization. *Financial Analysts Journal*, *48*(5), 28–43. doi:10.2469/faj.v48.n5.28

Black, F., & Litterman, R. B. (1991). Asset allocation: Combining investor views with market equilibrium. *The Journal of Fixed Income*, *1*(2), 7–18. doi:10.3905/jfi.1991.408013

Borchers, B., & Mitchell, J. E. (1994). An improved branch and bound algorithm for mixed integer nonlinear programs. *Computers & Operations Research*, *21*, 359–367. doi:10.1016/0305-0548(94)90024-8

Borchers, B., & Mitchell, J. E. (1997). A computational comparison of branch and bound and outer approximation algorithms for 0-1 mixed integer nonlinear programs. *Computers & Operations Research*, *24*, 699–701. doi:10.1016/S0305-0548(97)00002-6

Brodie, J., Daubechies, I., De Mol, C., Giannone, D., & Loris, I. (2009). Sparse and stable Markowitz portfolios. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(30), 12267–12272. doi:10.1073/pnas.0904287106

Bruni, R., Cesarone, F., Scozzari, A., & Tardella, F. (2015). A linear risk-return model for enhanced indexation in portfolio optimization. *OR Spectrum*, *37*(3), 735–759. doi:10.1007/s00291-014-0383-6

Bruni, R., Cesarone, F., Scozzari, A., & Tardella, F. (2017). On exact and approximate stochastic dominance strategies for portfolio selection. *European Journal of Operational Research*, *259*(1), 322–329. doi:10.1016/j.ejor.2016.10.006

Chang, T. J., Meade, N., Beasley, J. E., & Sharaiha, Y. M. (2000). Heuristics for cardinality constrained portfolio optimisation. *Computers & Operations Research*, *27*(13), 1271–1302. doi:10.1016/S0305-0548(99)00074-X

Chen, C., Li, X., Tolman, C., Wang, S., & Ye, Y. (2013). Sparse Portfolio Selection via Quasi-Norm Regularization. *arXiv Preprint arXiv*, *1312*, 6350.

Chopra, V. K. (1993). Improving optimization. *The Journal of Investing*, *2*(3), 51–59. doi:10.3905/joi.2.3.51

Chopra, V. K., & Ziemba, W. T. (1993). The effect of errors in means, variances, and covariances on optimal portfolio choice. *The Journal of Portfolio Management*, *19*(2), 6–11. doi:10.3905/jpm.1993.409440

Dantzig, G. B., & Infanger, G. (1993). Multi-stage stochastic linear programs for portfolio optimization. *Annals of Operations Research*, *45*(1), 59–76. doi:10.1007/BF02282041

d'Aspremont, A., El Ghaoui, L., Jordan, M. I., & Lanckriet, G. R. (2007). A direct formulation for sparse PCA using semi-definite programming. *SIAM Review*, *49*(3), 434–448. doi:10.1137/050645506

DeMiguel, V., Garlappi, L., Nogales, F. J., & Uppal, R. (2009). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science*, *55*(5), 798–812. doi:10.1287/mnsc.1080.0986

Fabozzi, F. J., Kolm, P. N., Pachamanova, D., & Focardi, S. M. (2007). *Robust Portfolio Optimization and Management*. Hoboken, New Jersey: Wiley.

Frost, P. A., & Savarino, J. E. (1988). For better performance: Constrain portfolio weights. *The Journal of Portfolio Management*, *15*(1), 29–34. doi:10.3905/jpm.1988.409181

Gao, J., & Li, D. (2011). Cardinality constrained linear-quadratic optimal control. *IEEE Transactions on Automatic Control*, *56*(8), 1936–1941. doi:10.1109/TAC.2011.2140770

Gao, J., & Li, D. (2013a). A polynomial case of the cardinality-constrained quadratic optimization problem. *Journal of Global Optimization*, *56*(4), 1441–1455. doi:10.1007/s10898-012-9853-z

Gao, J., & Li, D. (2013b). Optimal cardinality constrained portfolio selection. *Operations Research*, *61*(3), 745–761. doi:10.1287/opre.2013.1170

Garey, M. R., & Johnson, D. S. (1979). *Computers and intractability: A guide to the theory of NP-completeness*. New York, NY: W. H. Freeman & Co.

Garlappi, L., Uppal, R., & Wang, T. (2007). Portfolio selection with parameter and model uncertainty: A multi-prior approach. *Review of Financial Studies*, *20*(1), 41–81. doi:10.1093/rfs/hhl003

Goldfarb, D., & Iyengar, G. (2003). Robust portfolio selection problems. *Mathematics of Operations Research*, *28*(1), 1–38. doi:10.1287/moor.28.1.1.14260

Gotoh, J. Y., & Takeda, A. (2011). On the role of norm constraints in portfolio selection. *Computational Management Science*, *8*(4), 323. doi:10.1007/s10287-011-0130-2

Halldórsson, B. V., & Tütüncü, R. H. (2003). An interior-point method for a class of saddle-point problems. *Journal of Optimization Theory and Applications*, *116*(3), 559–590. doi:10.1023/A:1023065319772

Jorion, P. (1986). Bayes-Stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis*, *21*(3), 279–292. doi:10.2307/2331042

Kim, W. C., Kim, J. H., Ahn, S. H., & Fabozzi, F. J. (2013). What do robust equity portfolio models really do? *Annals of Operations Research*, *205*(1), 141–168. doi:10.1007/s10479-012-1247-6

Kim, W. C., Kim, J. H., & Fabozzi, F. J. (2014). Deciphering robust portfolios. *Journal of Banking & Finance*, *45*, 1–8. doi:10.1016/j.jbankfin.2014.04.021

Kim, W. C., Kwon, D.-G., Lee, Y., Kim, J. H., & Lin, C. (2019). *Personalized goal-based investment via multi-stage stochastic goal programming. KAIST Working Paper*.

Kim, M. J., Lee, Y. J., Kim, J. H., & Kim, W. C. (2016). Sparse tangent portfolio selection via semi-definite relaxation. *Operations Research Letters*, *44*(4), 540–543. doi:10.1016/j.orl.2016.05.012

Konno, H., & Yamamoto, R. (2005). Integer programming approaches in mean-risk models. *Computational Management Science*, *2*(4), 339–351. doi:10.1007/s10287-005-0038-9

Konno, H., & Yamazaki, H. (1991). Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market. *Management Science*, *37*(5), 519–531. doi:10.1287/mnsc.37.5.519

Ledoit, O., & Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance*, *10*(5), 603–621. doi:10.1016/S0927-5398(03)00007-0

Li, D., Sun, X., & Wang, J. (2006). Optimal lot solution to cardinality constrained mean–variance formulation for portfolio selection. *Mathematical Finance*, *16*(1), 83–101. doi:10.1111/j.1467-9965.2006.00262.x

Lobo, M. S., & Boyd, S. (2000). The worst-case risk of a portfolio. *Unpublished manuscript*. Retrieved from http://faculty. fuqua. duke. edu/% 7Emlobo/bio/researchfiles/rsk-bnd. pdf.

Maringer, D., & Kellerer, H. (2003). Optimization of cardinality constrained portfolios with a hybrid local search algorithm. *OR Spectrum*, *25*(4), 481–495. doi:10.1007/s00291-003-0139-1

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*(1), 77–91. doi:10.1111/j.1540-6261.1952.tb01525.x

Olivares-Nadal, A. V., & DeMiguel, V. (2018). A robust perspective on transaction costs in portfolio optimization. *Operations Research*, *66*(3), 733–739. https://doi.org/10.1287/opre.2017.1699.

Roman, D., Mitra, G., & Zverovich, V. (2013). Enhanced indexation based on second-order stochastic dominance. *European Journal of Operational Research*, *228*(1), 273–281. doi:10.1016/j.ejor.2013.01.035

Tütüncü, R. H., & Koenig, M. (2004). Robust asset allocation. *Annals of Operations Research*, *132*(1–4), 157–187. doi:10.1023/B:ANOR.0000045281.41041.ed

Wilding, T. (2003). Using genetic algorithms to construct portfolios. In *Advances in portfolio construction and implementation* (pp. 135–160). Oxford, UK: Elsevier.

Xu, H., Caramanis, C., & Mannor, S. (2009). Robust regression and lasso. In *Advances in Neural Information Processing Systems*, 1801–1808.

Appendix

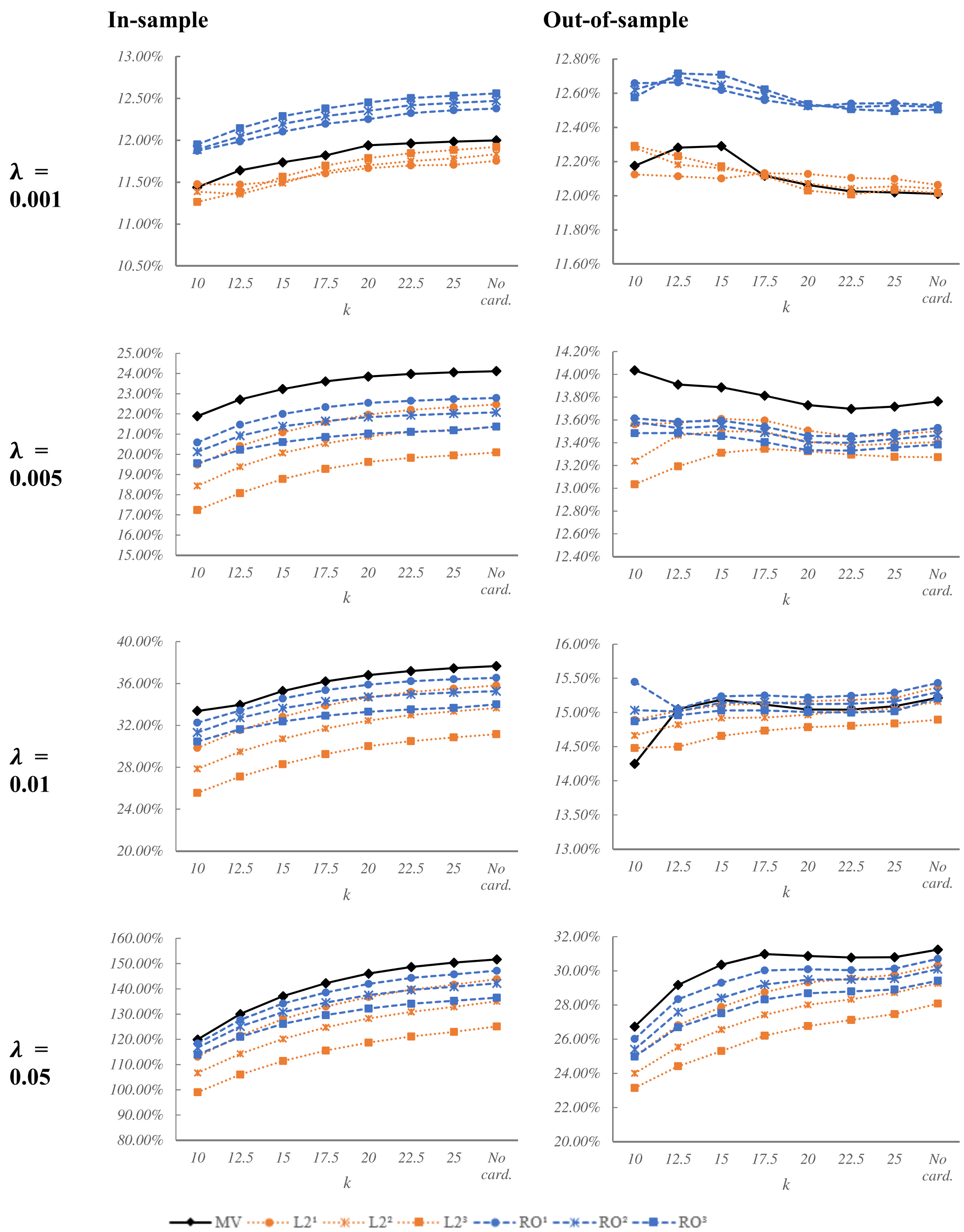

Figure A1. Average portfolio annual returns of MV, L2, and RO.

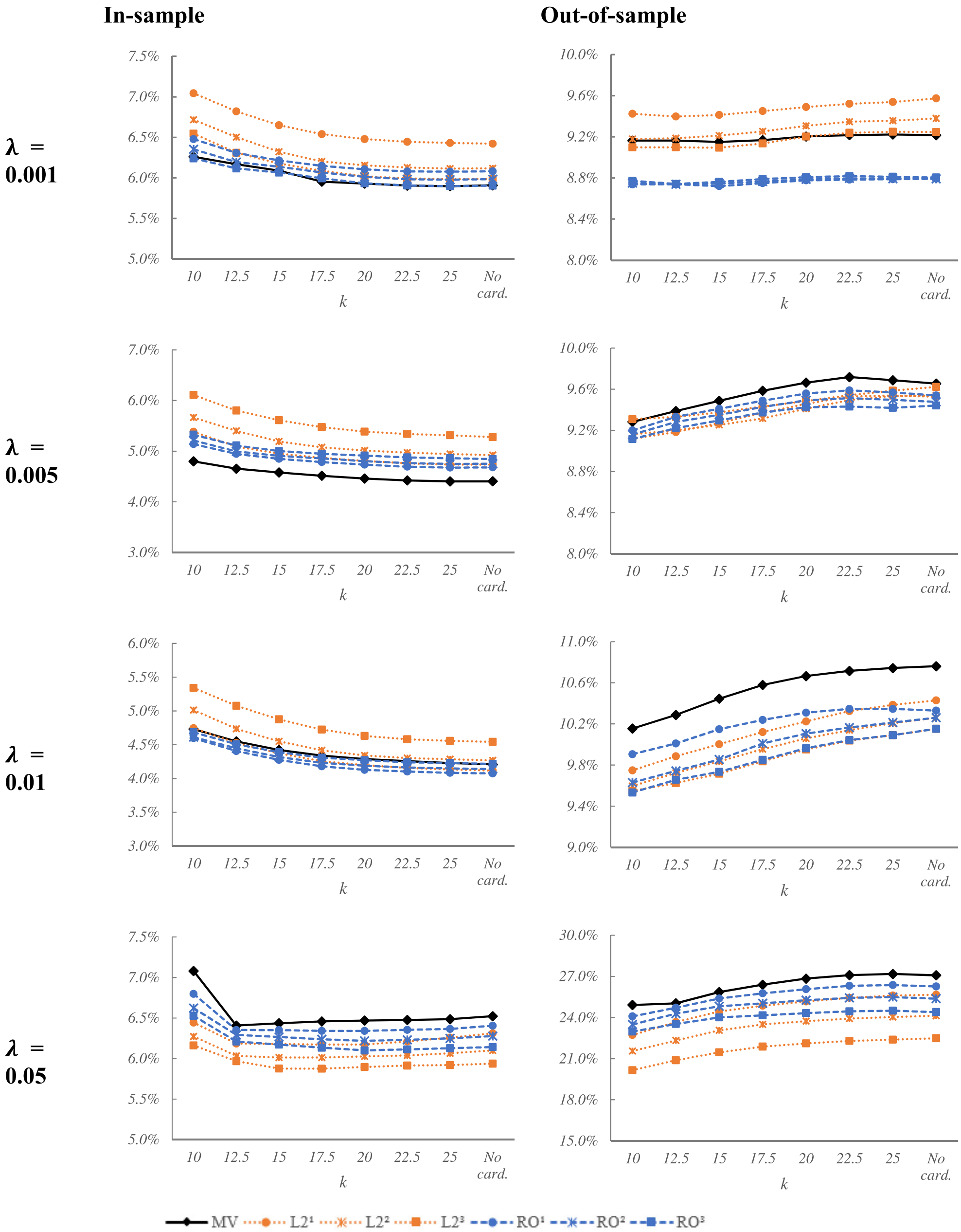

Figure A2. Average portfolio MDDs of MV, L2, and RO.