Smart network based portfolios

Gian Paolo Clemente[1] · Rosanna Grassi[2] · Asmerilda Hitaj[3]

Abstract

In this article we deal with the problem of portfolio allocation by enhancing network theory tools. We propose the use of the correlation network dependence structure in constructing some well-known risk-based models in which the estimation of the correlation matrix is a building block in the portfolio optimization. We formulate and solve all these portfolio allocation problems using both the standard approach and the network-based approach. Moreover, in constructing the network-based portfolios we propose the use of three different estimators for the covariance matrix: the sample, the shrinkage toward constant correlation and the depth-based estimators . All the strategies under analysis are implemented on three high-dimensional portfolios having different characteristics. We find that the network-based portfolio consistently performs better and has lower risk compared to the corresponding standard portfolio in an out-of-sample perspective.

Keywords Portfolio optimization · Mean-variance · Smart Beta strategies · Networks · Dependence · Interconnectedness

1 Introduction

Modern portfolio theory originates with the seminal work of Markowitz (1952). This work proposes the innovative idea of relating the return of an asset (the mean) and its risk (the variance) together with those of the other assets in the portfolio selection, through the mean-variance model. Nevertheless the prominent role in modern investment theory, this model, when applied in asset management setting, can lead to a poor out-of-sample portfolio per-

✉ Asmerilda Hitaj
  asmerilda.hitaj@uninsubria.it

  Gian Paolo Clemente
  gianpaolo.clemente@unicatt.it

  Rosanna Grassi
  rosanna.grassi@unimib.it

[1] Dipartimento di Discipline Matematiche, Finanza Matematica ed Econometria, Università Cattolica del Sacro Cuore, Milan, Italy

[2] Dipartimento di Statistica e Metodi Quantitativi, Università degli Studi di Milano-Bicocca, Milan, Italy

[3] Dipartimento di Economia, Università degli studidell'Insubria, Varese, Italy

formance, due to the estimation errors of the input parameters (see, for instance Merton 1980; Jobson and Korkie 1981). Furthermore, the risk, measured through the variance and the correlation, is based on expected values representing only a statistical statement about the future. Such measures often cannot capture the true statistical features of the risk and return which often follow highly skewed distributions.

To overcome these main drawbacks, several variations and extensions of the original methodology have been proposed in the literature. In Jagannathan and Ma (2003), a higher out-of-sample performance is derived by imposing specific constraints, and these results have been further confirmed in Behr et al. (2013) and Hitaj and Zambruno (2016). Alternative approaches deal with the problem of optimal portfolio choice by employing a Bayesian methodology to estimate unknown mean-variance parameters reducing the estimation errors. In this context, one of the most prominent is the Bayes-Stein approach based on the idea of shrinkage estimation (Jorion 1985, 1986; Bauder et al. 2018). The authors in Ledoit and Michael (2004) propose the shrinkage estimator toward the constant correlation, while in Martellini and Ziemann (2009) and Hitaj et al. (2012) this approach has been extended to higher moments such as skewness and kurtosis. Empirical analyses have shown that the use of shrinkage estimators for the mean-variance parameters often improves the out-of-sample performance (see Hitaj and Zambruno 2016, 2018). Robust optimization is widely used for dealing with uncertainty in parameters. For instance, Goldfarb and Iyengar (2003) proposed robust portfolio selection problems reformulated as second order cone programs under uncertainty structures for the market parameters. Similarly, robust mean-variance models have been put forward by Garlappi et al. (2007) and Fliege and Werner (2014). The authors in Pandolfo et al. (2020) proposed a robust estimation for mean and variance through the use of the weighted L^p depth function. In particular they considered $p = 2$ and performed empirical analysis using market portfolios. The authors concluded that the use of weighted L^p depth function for the estimation of mean and variance is a valuable alternative in a portfolio selection problem. Exhaustive reviews about robust portfolio selection are Fabozzi et al. (2010) and Scutella and Recchia (2013) where the authors focus on the application of robust optimization only in basic mean-variance, mean-CVaR, and mean-VaR problems. Theoretical contributions can be found for other models based on either different risk measures (see Zhu and Fukushima 2009; Zymler et al. 2013; Benati and Conde 2022), or alternative performance indicators (see Kapsos et al. 2014). Concerning the empirical aspect, recently (Georgantas et al. 2021) compare the performance of several models based on well-known risk measures.

It is well known that the effects of the estimation errors of the returns are higher than the effects of the estimation errors of the covariance matrix (see, among others, Chopra and Ziemba 2013). For this reason many portfolio strategies proposed in literature have put aside returns. These are called *risk based* strategies because they rely only on the estimation of the covariance matrix. Some well-known risk based strategies are Global Minimum Variance, Equally Weighted, (DeMiguel et al. 2007), Equal Risk Contribution (Qian 2006; Maillard et al. 2010) and Maximum Diversified Portfolio (Choueifaty and Coignard 2008). The *risk based* strategies are also called *Smart Beta*[1] strategies, as they are also proposed as alternatives to market capitalization-weighted indices, which are claimed to be not efficient (Choueifaty and Coignard 2008). The literature on constructing new portfolios able to beat a benchmark is vast and is not limited to the papers cited above. In particular, there is a wide literature on *enhanced indexing* where the objective is to outperform the index (see, e.g., Bruni et al. 2017;

---

[1] The term *Smart Beta* is popular to denote any strategy which attempts to take advantage of the benefits of traditional passive investments, adding a source of outperformance in order to beat traditional market capitalization-weighted indices. For more information on the *risk-based* strategies see e.g. Amenc and Goltz (2013) and the references therein.

de Paulo et al. 2016; Dentcheva and Ruszczynski 2003; Guastaroba et al. 2016; Roman et al. 2013). We highlight that in this paper we focus only the *Smart Beta* strategies.

In the last few years the problem of asset allocation has been discussed under a different perspective. Clustering methods for financial time series has been used to build a portfolio of assets selected by the resulting partition (see Iorio et al. 2018). Another approach consists in using network theory to represent the financial market. Indeed, in network-based portfolio models, the correlation matrix is included in the network structure, in order to reproduce the dependence among the assets (see, for instance, Mantegna 1999; Onnela et al. 2003b; Pozzi et al. 2013; Zhan et al. 2015), providing in this way useful insights in the portfolio selection process.

In particular, the minimum spanning tree has been used in Onnela et al. (2003b), the authors in Pozzi et al. (2013) use Planar Maximally Filtered Graphs, while in Zhan et al. (2015) hierarchical clustering trees and neighbor-nets have been applied in order to reduce the complexity of the network, characterizing the heterogeneous spreading of risk across a financial market. The work of Peralta and Zareei (2016) establishes a bridge between Markowitz's framework and the network theory, showing a negative relationship between optimal portfolio weights and the centrality of assets in the financial market network. As a result, the centrality measures of constructed networks can be used to facilitate the portfolio selection. A generalization to this approach has been provided in Vȳrost et al. (2019).

Recently, an alternative methodology to tackle the asset allocation problem using the network theory has been proposed in Clemente et al. (2019). Specifically, the authors catch how much a node is embedded in the system, by adapting to this context the clustering coefficient, a specific network index (see Barrat et al. 2004; Clemente and Grassi 2018; Fagiolo 2007; McAssey and Bijma 2015; Wasserman and Faust 1994; Watts and Strogatz 1998), meaningful in financial literature to assess systemic risk (Bongini et al. 2018; Minoiu and Reyes 2013; Tabak et al. 2014). The underlying structure of the financial market network is used as an effective tool in enhancing the portfolio selection process. In particular, the optimal allocation is obtained by maximizing a specific objective function that takes into account the interconnectedness of the system, unlike the classical global minimum variance model that is based only on the pairwise correlation between assets. Furthermore, in constructing the dependence structure of the portfolio network, various dependence measures are tested, namely, the Pearson correlation, Kendall correlation and lower tail dependence. All these measures are estimated using the sample approach. The results obtained in Clemente et al. (2019) show that, independently from the length of the rolling window and from the used dependence structure, the network-based portfolio leads to better out-of-sample performance compared with the classical approach.

The aim of this paper is to move one step further by enhancing the role of the network theory in solving portfolio allocation problem. The main contribution is the extension of the existing network-based approaches to different portfolio selection problems where the objective function depends on the variance-covariance matrix. In particular, we contribute to the existing literature along various dimensions.

On the one hand, we exploit the network theory constructing the Smart Beta strategies and the mean-variance portfolio, where alternative values of the trade-off parameter are considered. On the other hand, we extend the network-theory to the estimation of the variance-covariance matrix using alternative methodologies, as the shrinkage (toward the constant correlation) and the weighted L^p depth function.

The out-of-sample performance of the proposed methodology is empirically tested. Specifically, the Pearson correlation is used in order to capture the dependence structure of the portfolio network. Moreover, we apply the network theory to various well-known models in

which the estimation of the correlation matrix is a building block in the portfolio optimization. We consider in this paper Equally Risk Contribution, Maximum Diversified Portfolio, Global Minimum Variance, and the mean-variance model. In this last case we consider different levels of the trade-off parameter. Moreover, since recent academic papers and practitioner publications suggest that equal-weighted portfolios appear to outperform various other price-weighted or value-weighted strategies (see, e.g., DeMiguel et al. 2007), we also include the Equally Weighted (EW) portfolio in our analysis.

Through empirical analyses, we test the impact of the estimation method on both the *standard* and *network-based* portfolios. For the sake of completeness, three different high-dimensional datasets with different characteristics are considered. The first dataset is composed by 266 among largest banks and insurance companies in the world, whose daily returns have been collected in the time-period ranging from January 2001 to December 2017. The second dataset is composed by the components of the *S&P 100* index (with Bloomberg ticker OEX) and the third dataset includes the components of the *Nikkei-225 Stock Average* (with Bloomberg ticker NKY) . The *OEX* and the *NKY* datasets contain daily returns in the time-period ranging from January 2001 to July 2021. All the obtained portfolios are compared in an out-of-sample perspective using some well-known performance measures. Main results show that, in the majority of the cases, the use of network-based approach leads to higher out-of-sample performances and lower volatility with respect to the corresponding *sample* strategy. The network-based portfolio is more robust with respect to the standard approach being only slightly affected by the estimation method of the covariance matrix. The out of-sample results suggest that the network-based strategy represents a viable alternative to classical portfolio strategies.

The remainder of the paper is organized as follows. Sect. 2 briefly recalls the investor's problems for each strategy under analysis. Section 2 explains the two estimation methods used for the covariance matrix. Section 3 explains in detail the approach of portfolio selection via network theory. Section 4 presents the empirical analysis and Sect. 5 draws main conclusions.

2 Portfolio selection strategies

In this section, we briefly set out the strategies used in the rest of the paper for the empirical analysis. We first introduce what we refer to as standard strategies. We start with the mean-variance problem and then we describe the most important Smart Beta approaches proposed as alternatives to the market capitalization-weighted indices, in the equity world.

Mean-variance (MV)

Let us first introduce the standard *mean-variance* model for a portfolio with N risky assets. Let R_i be the random variable (r.v.) of daily log returns. Let $\mathbf{r} = [r_i]_{i=1,...,N}$ be the returns' vector observed in a specific time period/window (w) and $\boldsymbol{\mu}$ ($\boldsymbol{\Sigma}$) be the mean vector (variance-covariance matrix) between assets estimated in the same period. Let $\mathbf{e}$ and $\mathbf{x} = [x_i]_{i=1,...,N}$ be, respectively, the vector of ones and the vector of portfolio weights, i.e. the proportional investments in the N risky assets. We denote with $\mu_p = \mathbf{x}^T \boldsymbol{\mu}$, σ_i and $\sigma_P = \sqrt{\mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x}}$ the portfolio mean, the standard deviation of the ith asset and the standard deviation of the portfolio, respectively. We recall that all optimization models considered in this paper include realistic investment constraints such as budged constraint (i.e. $\mathbf{e}^T \mathbf{x} = 1$) and non-short selling

constraints $x_i \geq 0 \ \forall i = 1, \ldots, N$ since many institutional investors are restricted to long positions only.

The mean-variance model proposed in Markowitz (1952) consists in optimizing a trade-off between risk and return. The standard mean-variance (MV) portfolio optimization problem is given by:

$$\begin{cases} \min_{\mathbf{x}} \quad \lambda \mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x} - (1 - \lambda) \mathbf{x}^T \boldsymbol{\mu} \\ \mathbf{e}^T \mathbf{x} = 1 \\ 0 \leq x_i \leq 1, \qquad\qquad\qquad i = 1, \ldots, N \end{cases}, \qquad (1)$$

where $\lambda \in [0, \ 1]$ expresses the trade-off between risk and return of the portfolio. It is possible to compute alternative points on the efficient frontier by solving problem (1) for different levels of λ. The level of λ plays a crucial role in the portfolio diversification and it can be set by the decision manager according to its preferences. A risk-prone investor may choose a low level of λ. The extreme case $\lambda = 0$ will lead to a highly concentrated portfolio as the investor is ignoring the risk and the portfolio will be concentrated to the asset with the higher mean. On the contrary, for $\lambda = 1$ the portfolio will be diversified as in this case the manager is ignoring the portfolio return and seeking the portfolio with the lower risk. Therefore, the lower is λ the less diversified may be the portfolio. For this reason in the empirical analysis we considered four alternative levels, $\lambda \in \{0.2 \ 0.4 \ 0.6 \ 0.8\}$

It is evident from (1) that the MV portfolio optimization relies on estimators of the means and covariances of the asset returns. This means that the MV portfolio strongly depends on the input data, see among others Jorion (1992) and Chopra and Ziemba (2013). The estimation methods used in this paper will be explained in more details in Sect. 2.1

Global minimum-variance portfolio (GMV)

The GMV strategy selects weights that minimize the variance of the portfolio ignoring completely the portfolio return. The GMV optimization problem is formulated as:

$$\begin{cases} \min_{\mathbf{x}} \quad \mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x} \\ \mathbf{e}^T \mathbf{x} = 1 \\ 0 \leq x_i \leq 1, \ i = 1, \ldots, N \end{cases}, \qquad (2)$$

Equally weighted portfolio (EW)

The EW strategy consists in holding a portfolio characterized by the same weight $\frac{1}{N}$ in each component. In the literature, it has been empirically showed that the EW portfolios perform better than many other quantitative models, with higher Sharpe Ratio and Certainty Equivalent return (see DeMiguel et al. 2007). Being the weights equally allocated among the assets, this strategy disregards the data and, of course, it does not require any optimization or estimation procedure.

Equal risk contribution portfolio (ERC)

The equal risk contribution strategy (ERC) is characterized by weights such that each asset provides the same contribution to the risk of the portfolio. The marginal contribution of the

asset i to the portfolio risk is:

$$\partial_{x_i} \sigma_P = \frac{\partial \sigma_P}{\partial x_i} = \frac{(\boldsymbol{\Sigma} \, \mathbf{x})_i}{\sqrt{\mathbf{x}^T \boldsymbol{\Sigma} \, \mathbf{x}}}.$$

Hence, $\sigma_i(x) = x_i \partial_{x_i} \sigma_P$ represents the risk contribution of the ith asset to the portfolio P. The authors in Maillard et al. (2010) proved that the portfolio risk can be expressed as:

$$\sigma_P = \sum_i^N \sigma_i(x)$$

that is the sum of risk contributions of the assets. The characterizing property of the ERC strategy is that weights are such that $\sigma_i(x) = \sigma_j(x) \, \forall \, i, j$. The result is a portfolio extremely diversified in terms of risk.

To obtain the optimal weights we have to solve an optimization problem consisting in minimizing the sum of all squared deviations under budged and non short-selling constraints. The mathematical formulation is the following:

$$\begin{cases} \min_{\mathbf{x}} \sum_{i=1}^N \sum_{j=1}^N \left(x_i (\boldsymbol{\Sigma} \, \mathbf{x})_i - x_j (\boldsymbol{\Sigma} \, \mathbf{x})_j \right)^2 \\ \mathbf{e}^T \mathbf{x} = 1 \\ x_i \geq 0 \;\; i = 1, ...N \end{cases}, \tag{3}$$

Maximum diversified portfolio (MDP)

The Maximum Diversification approach aims to construct a portfolio that maximizes the benefits from diversification. This goal can be achieved by solving a maximization problem where the objective function is given by the so-called Diversification Ratio $DR = \frac{\sum_{i=1}^N x_i \sigma_i}{\sigma_P}$, under the usual constraints. The mathematical formulation for the MDP strategy is:

$$\begin{cases} \max_{\mathbf{x}} \frac{\sum_{i=1}^N x_i \sigma_i}{\sqrt{\mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x}}} \\ \mathbf{e}^T \mathbf{x} = 1 \\ x_i \geq 0 \;\; i = 1, ...N \end{cases} \tag{4}$$

This approach creates portfolios characterized by minimally correlated assets, providing lower risk levels and higher returns than market cap-weighted portfolios strategies (see Choueifaty and Coignard 2008).

2.1 Estimation methods for covariance matrix

Unlike the investment problems (2), (3) and (4), in which only the estimate of the covariance matrix between assets in a given time interval is needed, we have to estimate both the covariance matrix and the mean vector in order to solve problem (1). A common way to estimate them is through the *sample approach*. This method allows to obtain each component $\hat{\mu}_i$ and $\hat{\sigma}_{i,j}$ by means of classical unbiased estimators. However, it is well known that the sample estimator of historical returns is likely to generate high sampling error. For this reason several methods have been introduced in order to improve the estimation of moments and comoments (see, among others, (Jorion 1985, 1986; Ledoit and Michael 2004; Martellini and Ziemann 2009) and Pandolfo et al. 2020). These methods are grounded on the idea of imposing some

structure on the moments (comoments) with the aim to reduce the number of parameters, leading in this way to a reduction of the sampling error at the cost of specification error.

It is worth stressing that in this work we mainly focus on the estimation of the covariance matrix. For this reason, we estimate the mean only through the sample approach, whereas we pay more attention to the estimate of the covariance matrix, considering also the shrinkage toward the constant correlation (CC) method (see Elton and Gruber 1973; Ledoit and Michael 2004) and *wighted L^P data depth function* (see Pandolfo et al. 2020). The idea of the CC approach is to estimate the covariance based on the fact that the correlation is assumed constant for each pair of assets, and it is given by the average of all the sample correlation coefficients (see Elton and Gruber 1973). The covariance between two assets is then computed as

$$\sigma_{i,j}^{CC} = \hat{\sigma}_i \hat{\sigma}_j \frac{1}{N(N-1)} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} (\hat{\rho}_{i,j}), \tag{5}$$

where $\hat{\rho}_{i,j}$ is the sample correlation between assets i and j.

This approach resizes the problem, as only one correlation coefficient and N standard deviations have to be estimated. The $\mathbf{\Sigma}^{CC}$ covariance matrix, constructed by using previous formula, is characterized by a lower estimation risk due to the assumed structure, nevertheless it involves some misspecification in the artificial structure imposed by this estimator. In the attempt to find a trade-off between the sample risk and the model risk, the authors in Ledoit and Michael (2004) introduce the asymptotically optimal linear combination of the sample estimator and the structured estimator (in our case, the CC estimator) in the context of the covariance matrix, with the weight given by the optimal shrinkage intensity κ.[2] Therefore, the shrinkage toward CC covariance matrix is given by:

$$\sigma_{i,j}^{shrink} = \kappa \sigma_{i,j}^{CC} + (1 - \kappa)\hat{\sigma}_{i,j}. \tag{6}$$

Recently, in Pandolfo et al. (2020) the robust estimation of the mean and variance based on statistical data depth functions has been used in finance. Let $\mathcal{F}$ be the class of distributions on the Borel sets of $\mathbb{R}^N$ ($N > 1$) and $F_{\mathbf{R}}$ be the joint distribution of a random vector $\mathbf{R} \in \mathbb{R}^N$

A statistical data depth function is a bounded and non-negative function $D(\cdot; \cdot) : \mathbb{R}^N \times \mathcal{F} \to \mathbb{R}$ which satisfies the following desirable properties:

- *Affine Invariance* $D(\mathbf{A}\mathbf{r} + \mathbf{b}; F_{\mathbf{A}\mathbf{R}+\mathbf{b}}) = D(\mathbf{r}, F_{\mathbf{R}})$ holds for any vector $\mathbf{R} \in \mathbb{R}^N$, for any nonsingular $N \times N$ matrix $\mathbf{A}$ and for any vector $\mathbf{b}, \mathbf{r} \in \mathbb{R}^N$. This implies that the depth of $\mathbf{r}$ should be invariant to scale and location.
- *Maximality at Center* Let $\mathbf{r}^\star \in \mathbb{R}^N$ be a uniquely defined *center* (e.g., the point of symmetry with respect to some notion of symmetry) of $F_{\mathbf{R}}$ then $D(\mathbf{r}^\star; F_{\mathbf{R}}) = \sup_{\mathbf{r} \in \mathbb{R}^N} D(\mathbf{r}; F_{\mathbf{R}})$; this means that D attains the maximum at the center.
- *Monotonicity Relative to Deepest Point* $D(\mathbf{r}; F_{\mathbf{R}}) \leq D(\mathbf{r}^\star + \beta(\mathbf{r} - \mathbf{r}^\star); F_{\mathbf{R}})$ for any $\beta \in [0, 1]$ and $\mathbf{r} \in \mathbb{R}^N$. This implies that as the $\mathbf{r}$ moves away from the center ($\mathbf{r}^\star$) the depth at $\mathbf{r}$ should decrease monotonically.
- *Vanishing at Infinity* $D(\mathbf{r}; F_{\mathbf{R}}) \to 0$ as $||\mathbf{r}|| \to \infty$. The depth of $\mathbf{r}$ should approaches 0 as $||\mathbf{r}||$ approaches infinity.

As in Pandolfo et al. (2020), in the empirical part of this paper we use a statistical data depth function based on the distance approach, where the non negative distance function belongs to the L^p norm ($p > 0$).

---

[2] For further information on how is estimated the shrinkage intensity κ see Ledoit and Michael (2004).

A distance function of $\mathbf{r} \in \mathbb{R}^N$ from $\mathbf{R}$, based on the L^p norm, can be written as:

$$D_{L^p}(\mathbf{r}, F_{\mathbf{R}}) = \frac{1}{1 + \mathbb{E}\,||\mathbf{r} - \mathbf{R}||_p}. \tag{7}$$

We recall that the distance D_{L^p} does not possess the affine invariance property. As reported in Zuo (2004) and Zuo et al. (2004), different distances with respect to the data may not have the same importance. For this reason in order to obtain location and scatter estimators, designed to achieve greater robustness, the authors proposed a weighted L^p depth function, given by:

$$WD_{L^p}(\mathbf{r}, F_{\mathbf{R}}) = \frac{1}{1 + \mathbb{E}\left[W\left(||\mathbf{r} - \mathbf{R}||_p\right)\right]}, \tag{8}$$

where $W(\cdot)$ is a weight function, non-decreasing and continuous on $[0, \infty)$, that down-weights outlying observations.

Let $(\mathbf{r}_1, \ldots, \mathbf{r}_d)'$ be a N-variate simple random sample of size d from $\mathbf{R} \sim F_{\mathbf{R}}$. The authors in Pandolfo et al. (2020) used the weighted L^2 data depth function to obtain robust estimates of the mean ($\bar{}^{WD_{L^2}}$) and covariance matrix ($\Sigma^{WD_{L^2}}$) of the asset returns, given respectively by:

$$\bar{}^{WD_{L^2}} = \frac{\sum_{t=1}^{d} W_1\left(D(\mathbf{r}_t, F_{\mathbf{R}})\right) \mathbf{r}_t}{\sum_{t=1}^{d} W_1\left(D(\mathbf{r}_t, F_{\mathbf{R}})\right)} \tag{9}$$

and

$$\Sigma^{WD_{L^2}} = \frac{\sum_{t=1}^{d} W_2\left(D(\mathbf{r}_t, F_{\mathbf{R}})\right)\left(\mathbf{r}_t - \bar{}^{WD_{L^2}}\right)\left(\mathbf{r}_t - \bar{}^{WD_{L^2}}\right)'}{\sum_{t=1}^{d} W_2\left(D(\mathbf{r}_t, F_{\mathbf{R}})\right)} \tag{10}$$

where $W_1\left(D(\mathbf{r}_t, F_{\mathbf{R}})\right)$ and $W_2\left(D(\mathbf{r}_t, F_{\mathbf{R}})\right)$ are two weighted functions non decreasing and continuous on $[0, \infty]$. $W_j\left(D(\cdot, \cdot)\right)$, for $j = 1, 2$, is given in Pandolfo et al. (2020) by:

$$W_j(h) = \begin{cases} \frac{\exp\left(-k\left(1 - (\frac{h}{c})^{2j}\right)^{2j}\right) - \exp(-k)}{1 - \exp(-k)} & \text{if} \quad h < c, \\ 1 & \text{if} \quad h \geq c \end{cases} \tag{11}$$

where c is the median of the data depth function and the value k determines how heavily the weight function penalizes as h get away from c. Following (Pandolfo et al. 2020), in the empirical part we set $p = 2$ and $k = 3$.

The advantages of the robust approaches considered in this paper are that both are non-parametric approaches and are less sensitive to changes in the asset return distribution compared to the sample estimate.

We highlight that in the empirical part the depth data function is $D(\mathbf{r};\ \hat{F}_{\mathbf{R},d})$, where $\hat{F}_{\mathbf{R},d}$ is the empirical joint distribution of $\mathbf{R}$ estimated from the observed returns $(\mathbf{r}_1, \ldots, \mathbf{r}_d)'$.

In the following section we recall the network correlation-based portfolio model and explain how the three estimators of the variance-covariance matrix (sample, shrinkage toward CC and WD_{L^2}) can be used.

3 Optimal portfolio via network theory

The portfolio selection problem and its variants can be formulated in a networks context and several researchers dealt with the assets allocation problem using network theory tools, contributing to the related literature (Peralta and Zareei 2016; Clemente et al. 2019; Li et al. 2019; Pozzi et al. 2013). All these articles share the same framework, namely the financial market is represented as a network, in which nodes are assets and weights on the edges identify a dependence measure between returns.

We describe in this section the approach proposed by Clemente et al. (2019). The authors formulate an investment strategy that benefits from the knowledge of the dependency structure that characterizes the market. Unlike the risk-based strategies, based on an objective function that accounts for pairwise correlations among assets, the objective function considers here the interconnectedness of the whole system.

In order to make the paper self-consistent, we briefly remind some preliminary definitions and notations about networks. A network $G = (V, E)$ consists in a set V of nodes and a set E of edges between nodes, where the edge (i, j) connects a pair of nodes i and j. If $(j, i) \in E$ whenever $(i, j) \in E$, the network is undirected. A network is complete if every pair of vertices is connected by an edge. We denote with $\mathbf{A}$ the real N-square matrix whose elements are $a_{ij} = 1$ whenever $(i, j) \in E$ and 0 otherwise (the adjacency matrix). A network is weighted if a weight $w_{ij} \in \mathbb{R}$ is associated to each edge (i, j). In this case, both adjacency relationships between vertices of G and weights on the edges are described by a non negative, real N-square matrix $\mathbf{W}$ (the weighted adjacency matrix). We denote with k_i and s_i the degree and strength of the node i $(i = 1, ..., N)$, respectively.

Relationships between assets are quantified through three different levels of dependence. For the sake of brevity, we report here only the approach referring to the classical linear correlation network, which is the most used dependence measure in the literature. Since all assets are correlated in the market, the correlation structure is represented through a weighted, complete and undirected network G, where weights on the edges are given by the Pearson correlation coefficient between them, that is $w_{ij} = \rho\left(R_i, R_j\right) \forall i \neq j$. In order to assure nonnegative weights, a distance can be associated with the correlation coefficient (see Giudici and Spelta 2016; Mantegna 1999; Onnela et al. 2003a). In our case, this transformation does not affect the results in terms of optimal portfolio.

The extension of the pairwise correlations, included in the quadratic form of the problem (2), to a general intercorrelation among all stocks at the same time is obtained optimizing a function that includes the clustering coefficient. The classic clustering coefficient and its variants defined in the literature (see Watts and Strogatz 1998; Cerqueti et al. 2018; Fagiolo 2007; Clemente and Grassi 2018) are not computable for complete networks, then we have to adapt its formulation to this framework.

Following a similar procedure to that proposed by McAssey and Bijma (2015), a threshold $s \in [-1, 1]$ is introduced on the matrix $\mathbf{W}$ in order to define the new matrix $\mathbf{A}_s$, whose elements a_{ij}^s are

$$a_{ij}^s = \begin{cases} 1 & \text{if } w_{ij} \geq s \\ 0 & \text{otherwise} \end{cases} . \tag{12}$$

$\mathbf{A}_s$ is the adjacency matrix describing the existing links in the network with weights w_{ij} at, or above the threshold s. Through this matrix we are selecting the strongest edges, namely those greater than a given threshold, bringing out the mean cluster prevalence of the network. On this new network we compute the clustering coefficient proposed in Watts and Strogatz

([1998](#)) and then we repeat the process, varying the threshold s. The clustering coefficient C_i for a node i corresponding to the graph is the average of $C_i(\mathbf{A}_s)$ overall $s \in [-1, 1]$:

$$C_i = \int_{-1}^{1} C_i(\mathbf{A}_s) ds \tag{13}$$

Since $0 \leq C_i \leq 1$, C_i is well-defined. Now, we define the N-square matrix $\mathbf{C}$, of entries

$$c_{ij} = \begin{cases} C_i C_j & \text{if } i \neq j \\ 1 & \text{otherwise} \end{cases}. \tag{14}$$

This matrix accounts for the level of interconnection for all pairs with the whole system, therefore, it can be used to construct the matrix

$$\mathbf{H} = \mathbf{\Delta}^T \mathbf{C} \mathbf{\Delta} \tag{15}$$

where $\mathbf{\Delta} = diag(s_i)$ is a diagonal matrix with diagonal entries

$$s_i = \frac{\hat{\sigma}_i}{\sqrt{\sum_{i=1}^{N} \hat{\sigma}_i^2}}.$$

Notice that the element s_i considers the contribute of the standard deviation of the return i to the total standard deviation, computed in case of independence. In Clemente et al. ([2019](#)), the authors solve the optimization problem defined in ([2](#)) replacing the covariance matrix $\mathbf{\Sigma}$ with $\mathbf{H}$.

The main difference between the classical and the network portfolio selection problem is due to the use of the interconnectedness matrix in order to consider how much each couple of assets is related to the system. In particular, being $\mathbf{C}$ dependent on a network-based measure of systemic risk (i.e. the clustering coefficient), we are implicitly including a measure of the state of stress of the financial system in the time period.

4 Dataset description and empirical analysis

4.1 Dataset description

The goal of this section is to examine the out-of-sample properties of the Smart Beta and mean-variance network-based portfolios in which the covariance matrix is estimated using the network theory though the methodology described in Sect. [3](#). In particular, we make a comparison between network-based portfolios and standard portfolio strategies, where the three previously described estimators of the covariance matrix are considered: the sample, the shrinkage and the weighted L^2 depth function. We summarize in Table [1](#) the alternative asset allocation models applied in this analysis.

As a robustness check we consider three large-dimensional portfolios with different characteristics. The investment universe of the first portfolio is composed by 266 among largest banks and insurance companies in the world.[3] In particular we have 120 insurers and 144 banks. The dataset of this portfolio contains daily returns in the time-period ranging from January 2001 to December 2017. The investment universe of the second portfolio consists of the components of the S&P 100 index. A third portfolio consists of the components of

---

[3] The greatest firms by market capitalization in the banking and insurance sector are considered.

Table 1 List of asset allocation models considered in the empirical study

Number	Model	Label
1	Standard sample based mean-variance (MV)	S-sample MV
2	Standard shrinkage toward constant correlation based mean-variance (MV)	S-shrinkage MV
3	Standard WD_{L2} based mean-variance (MV)	S-WD_{L2} MV
4	Network sample based mean-variance (MV)	NB-sample MV
5	Network shrinkage toward constant correlation based mean-variance (MV)	NB-shrinkage MV
6	Network WD_{L2} based mean-variance (MV)	NB-WD_{L2} MV
7	Standard sample based Maximum Diversified Portfolio (MDP)	S-sample MDP
8	Standard shrinkage toward constant correlation based MDP	S-shrinkage MDP
9	Standard WD_{L2} based Maximum Diversified Portfolio (MDP)	S-WD_{L2} MDP
10	Network sample based MDP	NB-sample MDP
11	Network shrinkage toward constant correlation based MDP	NB-shrinkage MDP
12	Network WD_{L2} based Maximum Diversified Portfolio (MDP)	NB-WD_{L2} MDP
13	Standard sample based Equally Risk Contribution (ERC)	S-sample ERC
14	Standard shrinkage toward constant correlation based ERC	S-shrinkage ERC
15	Standard WD_{L2} based Equally Risk Contribution (ERC)	S-WD_{L2} ERC
16	Network sample based ERC	NB-sample ERC
17	Network shrinkage toward constant correlation based ERC	NB-shrinkage ERC
18	Network WD_{L2} based Equally Risk Contribution (ERC)	NB-WD_{L2} ERC
19	Standard sample based Global Minimum Variance (GMV)	S-sample GMV
20	Standard shrinkage toward constant correlation based GMV	S-shrinkage GMV
21	Standard WD_{L2} based Global Minimum Variance (GMV)	S-WD_{L2} GMV
22	Network sample based GMV	NB-sample GMV
23	Network shrinkage toward constant correlation based GMV	NB-shrinkage GMV
24	Network WD_{L2} based Global Minimum Variance (GMV)	NB-WD_{L2} GMV
25	Equally Weighted	EW

The last column of the table indicates the label used to refer to each strategy in the empirical section, where the performance of the various approaches is compared

Nikkei-225 Stock Average, that considers the 225 stocks selected from domestic common stocks in the first section of the Tokyo Stock Exchange. Last two portfolios contain daily returns in the time-period ranging from January 2001 to August 2021.[4]

All the portfolios discussed in this paper are analysed and compared in an out-of-sample perspective. In particular the first four moments, the Sharpe Ratio (SR), the Omega Ratio (OR) the Information Ratio (IR) and the out-of-sample performance are used to compare the portfolios. All these aspects are investigated through a rolling window methodology, which is characterized by an in-sample period of length d and an out-of-sample period of length k.[5] This means that the first in-sample window of width d contains the observations of all the components in the portfolio from $t = 1$ to $t = d$. The dataset of the first in-sample window is used to estimate the optimal weights, using the different portfolio selection models considered in this paper and listed in Table 1. These optimal weights are then invested in the

---

[4] All the data have been downloaded from Bloomberg (2012).

[5] In the literature different lengths for the in-sample and out-of-sample period has been used, see among others Cesarone et al. (2013) and Consigli et al. (2018).

out-of-sample period, from $t = d + 1$ to $t = d + k$, where the out-of-sample performance is computed. The process is repeated rolling the window k steps forward. Hence, weights are updated by solving the optimal allocation problem in the new subsample and the performance is estimated once again using data from $t = d + k + 1$ to $d + 2k$. Repeating these steps until the end of the dataset is reached, we buy-and-hold the portfolios and we record out-of-sample portfolio returns in each rebalancing period.

To ensure the robustness of our results, we analyse three alternative estimation windows; namely, 6 months in-sample and 1 month out-of-sample, 1 year in-sample and 1 month out-of-sample and two years in-sample and 1 month out-of-sample.

4.2 Performance measures

In order to assess the magnitude of potential gains/losses that can be attained by an investor adopting a network-based portfolio selection, we implement an out-of-sample analysis. For this reason several performance measures are calculated. First, for each optimization strategy, we compute the first four moments of the out-of-sample portfolio returns. Further, for each strategy j, we determine the out-of-sample Sharpe Ratio of the optimal portfolio[6]:

$$SR_j^\star = \frac{\mu_{p_j}^\star - \mu_f}{\sigma_{p_j}^\star},$$

where $\mu_{p_j}^\star$ and $\sigma_{p_j}^\star$ are, respectively, the average return and the standard deviation of the optimal portfolio according to the strategy j and μ_f indicates the average risk-free rate.[7] This ratio measures the average return of a portfolio in excess of the risk-free rate, also called the risk premium, as a fraction of the portfolio total risk, measured by its standard deviation. As alternatives performance measures we also calculate the Information ratio (IR) and the Omega ratio (OR). The Information ratio of the optimal portfolio is defined as:

$$IR_j^\star = \frac{\mu_{p_j}^\star - \mu_{p_{ref}}^\star}{\sigma\left(r_{p,j}^\star - r_{p_{ref}}^\star\right)}.$$

where $\mu_{p_{ref}}^\star$ is the average return of the reference portfolio and $r_{p,j}^\star, r_{p_{ref}}^\star$ represent the out-of-sample time series of optimal portfolio returns corresponding to a strategy j and the reference strategy, respectively.

Once identified the reference portfolio, managers seek to maximize IR_j, i.e. to reconcile a high residual return and a low tracking error. This ratio allows to check if the risk taken by the manager in deviating from the reference portfolio is sufficiently rewarded.

The *Omega Ratio* has been introduced by Keating and Shadwick in Keating and Shadwick (2002) and it is defined as:

$$OR_j^\star = \frac{\int_\epsilon^{+\infty}(1 - F_j(x))\,dx}{\int_{-\infty}^\epsilon F_j(x)dx} = \frac{\mathbb{E}\left(r_{p_j}^\star - \epsilon\right)^+}{\mathbb{E}\left(\epsilon - r_{p_j}^\star\right)^+},$$

---

[6] In general, the superscript $\star$ indicates that the statistic is calculated using the out-of-sample time series of returns of the optimal portfolio.

[7] As a proxy for the risk-free rate the literature suggests the use of *1 month* or *3 months* maturity U.S. Treasury Bills (see e.g Deguest et al. 2013) or, alternatively, an exogenously given value (for instance, $\mu_f = 5\%$ is considered in Brennan 1998). In the empirical part of this paper, for illustration purposes, we set $\mu_f = 0\%$.

where $F_j(x)$ is the cumulative distribution function of the portfolio returns for a strategy j and ϵ is a specified threshold.[8] Returns below (respectively above) the threshold are considered as losses (respectively gains). In general, a value of the OR_j greater than 1 indicates that strategy j provides more expected gains than expected losses. The portfolio with the highest ratio will be preferred by the investor. The OR_j implicitly embodies all the moments of the return distribution without any a-priori assumption.

4.3 Empirical results

As previously explained, for robustness purposes we consider three large-dimensional datasets. The first is the Banks and Insurers dataset composed by 266 among largest banks and insurance companies in the world. The second is composed by the assets of the S&P 100 index and the third regards the constituents of the Nikkei-225 Stock Average (NKY). This empirical analysis is based on a buy and hold strategy. For the sake of completeness we consider three alternative strategies, with an in-sample period of two years, one year and six months, respectively, and an out-of-sample period of one month. For the sake of brevity, we report in the following the results obtained for the NKY dataset with a rolling window of one year in-sample and one month out-of-sample. However, all the detailed statistics of the three analysed portfolios, for all strategies and estimation methods used for the covariance matrix are reported in the Supplementary Material.

The proposed network approaches allow to visualize the portfolio composition and the dependence structure between assets. To have a preliminary idea of the results, we depict in Fig. 1a the structure of the network of the NKY dataset based on the correlation matrix, obtained via sample estimation, at different time periods. Each node represents an asset and the weighted edge (i, j) measures the correlation between assets i and j.

As shown in Fig. 1a (top, right) that covers the window January 2008-December 2008, it is noticeable a higher dependence and a higher volatility of the returns in periods of financial crisis. A prominent volatility is partially observed also in 2020 (see Fig. 1a (bottom, right)) due to the effects of Covid-19 announcements on the financial market.

As described in Sect. 3, we solve a network-based portfolio model where the clustering coefficient is used in the optimal problem to catch the structure of interconnections. Indeed, the intensity of the relations between assets is related to the pairwise correlations, that affect the value of the clustering coefficient and, therefore, the optimal solution.

We report in Fig. 1b the optimal solutions of the sample network-based GMV problem, i.e. NB-sample GMV, for the same windows w considered in Fig. 1a. In this network representation, the size of bullets is instead proportional to the allocated weight. We observe that the initial endowment is invested in a limited number of assets. As expected, the approach tends to allocate weights on assets characterized by a low volatility and with a preference on assets that are negatively correlated to the rest of the network.

As mentioned before, we will focus in this section on the NKY dataset, but similar results have been obtained for the other datasets. To provide an example, we depict in Fig. 2 the sample correlation network obtained in the window that covers the period December 2015-November 2017 for the Bank and Insurers dataset. In this case, the initial endowment is invested in only 26 firms, specifically 10 banks and 16 insurance companies. However, approximatively 94% of the total amount is invested in insurers that are characterized on average in this time period by both a lower volatility and a lower clustering coefficient.

---

[8] We point out that OR ratio is very sensitive to values of ϵ which can be different from investor to investor. In the empirical analysis ϵ is set equal to 0.
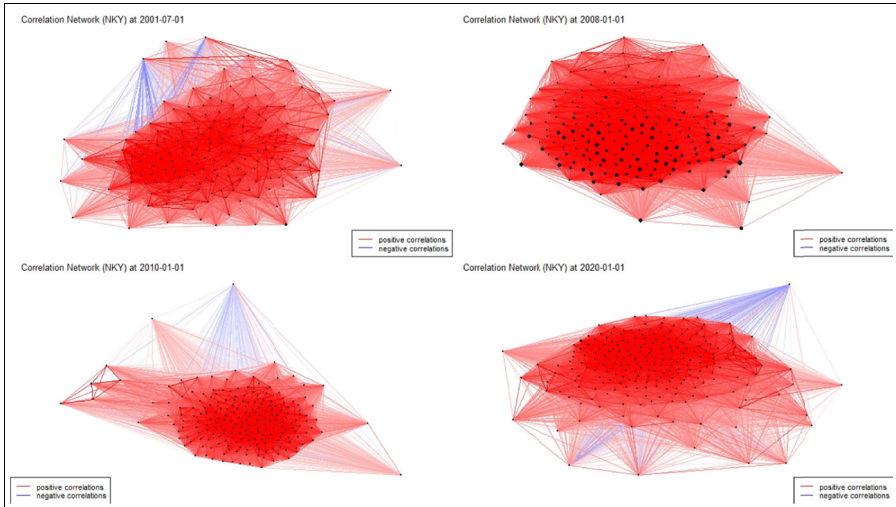
(a) Pearson Correlation Network computed by using returns of NKY dataset (based on sample estimation) referred to different time periods. The rolling window is one year in-sample. The date in the title is the initial period of the rolling window. Bullets size is proportional to the standard deviation of each firm. Edges opacity is proportional to edges weights (i.e. intensity of correlations).
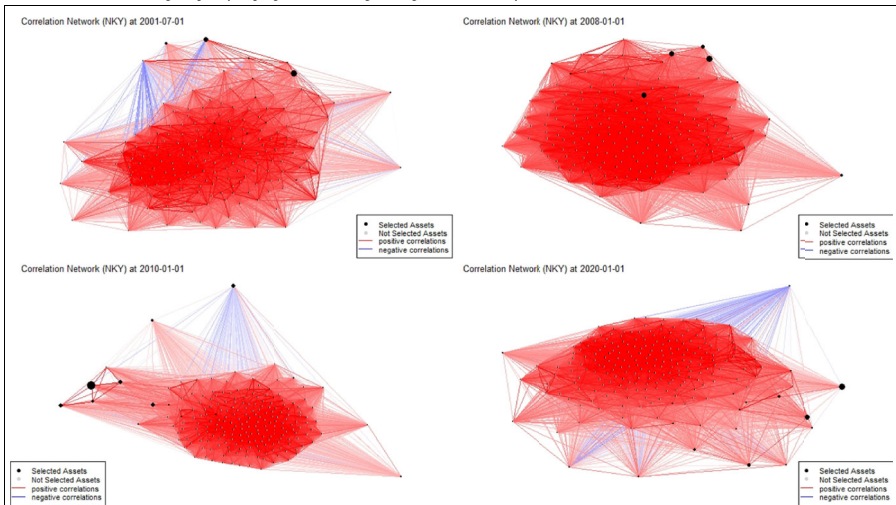
(b) The optimal network-based sample GMV portfolio referred to the same periods as in 1a, where the covariance matrix is estimated using the sample approach (NB-sample GMV). In this figure, the bullets size is proportional to the allocated weight. Edges opacity is proportional to edges weights.

Fig. 1 **a** Pearson Correlation Network computed by using returns of NKY dataset (based on sample estimation) referred to different time periods. The rolling window is one year in-sample. The date in the title is the initial period of the rolling window. Bullets size is proportional to the standard deviation of each firm. Edges opacity is proportional to edges weights (i.e. intensity of correlations). **b** The optimal network-based sample GMV portfolio referred to the same periods as in **a**, where the covariance matrix is estimated using the sample approach (NB-sample GMV). In this figure, the bullets size is proportional to the allocated weight. Edges opacity is proportional to edges weights

It is worth noting the case of two insurers, Nationwide Mutual Insurance Company and One America, which are characterized by the lowest standard deviations and a high proportion of negative pairwise correlations (for instance, approximatively 90% of correlations between
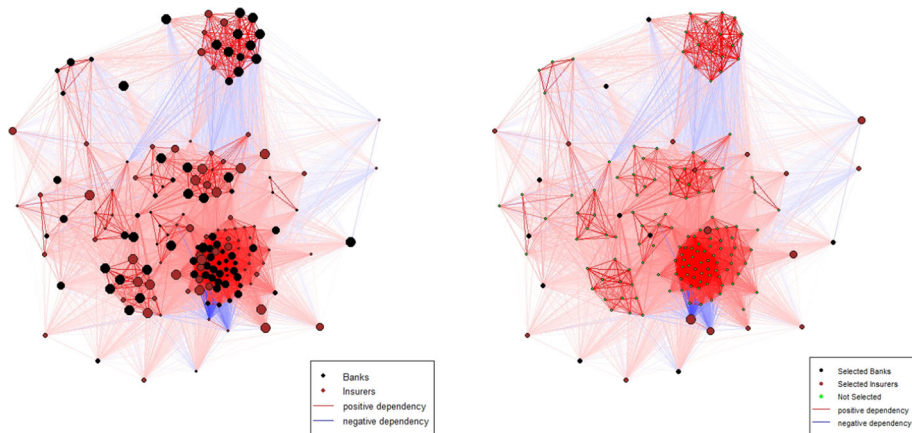
Fig. 2 On the left, Pearson Correlation Network computed by using returns of Banks and Insurers dataset (based on sample estimation) referred to the last window, from the beginning of December 2015 to the end of November 2017. Bullets size is proportional to the standard deviation of each firm. Edges opacity is proportional to edges weights (i.e. intensity of correlations). On the right, the optimal network-based sample GMV portfolio (NB-sample GMV) referred to the same period. In this figure, the bullets size is proportional to the allocated weight. Edges opacity is proportional to edges weights

Nationwide Mutual and other firms is lower than zero). As expected, the optimal portfolio allocates a high proportion of the initial endowment in these two firms (54% and 17% respectively).

In the following we consider the out-of-sample performances of all the models under investigation for the NKY dataset. We analyse the Smart Beta and the MV optimal portfolios obtained using standard and network-based approaches for sample, shrinkage and the weighted L^2 depth estimators.

In Fig. 3 the out-of-sample performances of the Smart Beta models under analysis are reported.

We observe that in all cases the EW strategy has the worst performance. Concerning the other strategies, focusing on MDP portfolios (see Fig. 3a) we have a remarkable prevalence of network-based approaches. The improved estimators lead to best performing models over time, with a prevalence of the network-based model (NB-WD_{L^2}). However, it should be pointed out that in all the dataset analysed strategies based on shrinkage and WD_{L^2} estimators tend to better perform during and immediately after the sovereign debt crisis.

Among the ERC portfolios (see Fig. 3b), the six alternative strategies show a similar pattern, with a slight preponderance of the NB-shrinkage and NB-WD_{L^2} approaches. Even for the GMV portfolio (see Fig. 3c), the six alternative strategies show a similar pattern, with a slight preponderance of the NB-shrinkage approach.

Figure 3d collects the best out-of-sample performances for each *risk based* analysed portfolio (i.e. MDP, ERC and GMV). What emerges is that the network-based approaches outperform classical strategies and the highest out-of-sample performance at the end of the period is assured by the NB-WD_{L^2} MDP portfolio.

Concluding, from the analysis of the performances of Smart Based portfolios with the alternative approaches, the network-based models almost always lead to higher out-of-sample performance compared to the corresponding classical ones. This result is confirmed also by
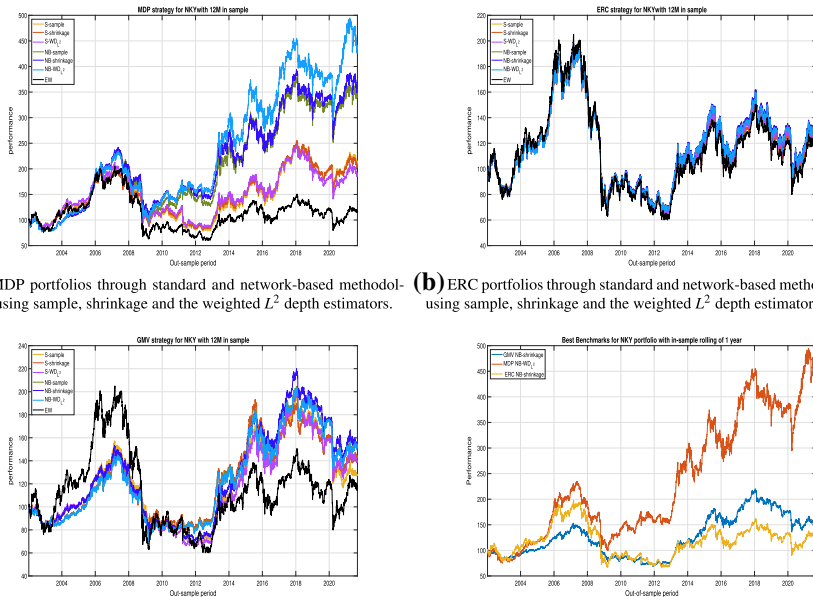
(a) MDP portfolios through standard and network-based methodology using sample, shrinkage and the weighted L^2 depth estimators.

(b) ERC portfolios through standard and network-based methodology using sample, shrinkage and the weighted L^2 depth estimators.

(c) GMV portfolios through standard and network-based methodology using sample, shrinkage and the weighted L^2 depth estimators.

(d) Best out-of-sample performances for each Smart Beta portfolio.

Fig. 3 Out-of-sample performances for NKY dataset with a rolling window of 1 year in-sample and 1 month out-sample. In **a**–**c**, we display the out-of-sample performances of EW, S-sample, S-Shrinkage, S-WD_{L^2}, NB-sample, NB-Shrinkage and NB-WD_{L^2} of MDP, ERC and GMV models, respectively. In **d** The best out-of-sample performances for each Smart Beta portfolio (MDP, ERC and GMV) are reported

the analysis carried on the other datasets, for all the estimators of the covariance matrix and for all the rolling windows strategies considered.

However, the simple inspection of the Fig. 3 is not enough in identifying the best portfolio selection strategy. To this end, in order to complete the analysis we report in Table 2 the four moments of the out-of-sample returns' distributions and values of alternative performance measures (namely, SR, $IR(EW)$[9] and OR). As well-known, these performance measures consider different characteristics of the portfolios and they could lead to different rankings between the models. However, by the inspection of Table 2, we can provide additional insights.

First, we observe that for all the strategies, the network-based approaches lead to a lower out-of-sample risk, measured by the standard deviation, regardless of the estimation method of the covariance matrix. Moreover, for each strategy the network-based approach almost always leads to a less relevant negative tail, due to a skewness closer to zero and a lower kurtosis with respect to the corresponding standard approach. These findings are further confirmed by SR and OR values. In particular, it results that the best portfolio is obtained by one of the *Network Based* approaches. In the majority of the results obtained there is a prevalence of the $NB - WD_{L^2}$ and $NB - shrinkage$. Hence, the results reported in Table 2 make us more confident in believing that using the network theory in building the *Smart Beta* strategies can be a good alternative to the standard approach not only for the easier visualization of the results (as reported in Figs. 1 and 3) but also for the better performances that they may reach in an out-of-sample perspective. The conclusions drawn for NKY portfolio with a monthly

---

[9] The $IR(EW)$ indicates the Information ratio where the reference portfolio is the Equally Weighted one.

Table 2 Out-of-sample statistics for the *risk-based* approaches in case of NKY portfolio with a buy and hold strategy of 12 months in-sample and 1 month out-of-sample

	S-sample	S-shrinkage	S-WD_{L2}	NB-sample	NB-shrinkage	NB-WD_{L2}	EW
ERC							
$\mu^\star$	0.00012	0.00013	0.00013	0.00014	**0.00015**	0.00014	0.00014
$\sigma^\star$	0.013	0.013	0.013	0.012	0.012	**0.012**	0.014
$skew^\star$	−0.513	−0.497	−0.520	−0.488	−0.488	−0.481	**−0.436**
$kurt^\star$	12.289	12.068	12.433	11.967	11.967	11.975	**10.949**
$SR^\star$	0.009	0.010	0.010	0.010	**0.012**	0.010	0.010
$IR^\star(EW)$	−0.003	−0.007	−0.004	0.001	**0.001**	0.000	
$OR^\star$	1.028	1.030	1.029	1.031	**1.032**	1.031	1.028
MDP							
$\mu^\star$	0.00022	0.00022	0.00022	0.00032	0.00033	**0.00037**	0.00014
$\sigma^\star$	0.012	0.011	0.012	0.011	0.011	**0.011**	0.014
$skew^\star$	−0.787	−0.787	−0.875	−0.655	−0.552	−0.591	**−0.436**
$kurt^\star$	12.612	12.918	13.772	11.588	11.355	11.334	**10.949**
$SR^\star$	0.018	0.019	0.018	0.026	0.027	**0.030**	0.010
$IR^\star(EW)$	0.011	0.014	0.016	0.020	0.023	**0.025**	
$OR^\star$	1.055	1.057	1.056	1.080	1.081	**1.090**	1.028
GMV							
$\mu^\star$	0.00012	0.00011	0.00011	0.00013	0.00013	0.00012	**0.00014**
$\sigma^\star$	0.009	0.009	0.010	0.009	**0.008**	**0.008**	0.014
$skew^\star$	−0.564	−0.477	−0.576	−0.404	−0.404	**−0.325**	−0.436
$kurt^\star$	18.439	18.579	18.769	15.801	15.766	16.428	**10.949**
$SR^\star$	0.013	0.012	0.012	**0.016**	**0.016**	0.015	0.010
$IR^\star(EW)$	−0.001	−0.005	−0.003	0.000	**0.000**	−0.002	
$OR^\star$	1.041	1.030	1.036	**1.042**	**1.043**	1.041	1.028

For each strategy (ERC, MDP and GMV), sample, shrinkage and WD_{L2} estimators are reported for classical and network-based models. The last column also considers results for EW. The best results are reported in bold for each measure. All the statistics are reported on daily bases

stepped one-year rolling window are still valid, in general, also for the other rolling windows and portfolios under analysis.[10]

Let us now analyse the results obtained in case of the Mean-Variance portfolio where different levels of trade-off between risk and return are considered. In particular we report the results for λ equal to 0.2, 0.4, 0.6, 0.8, respectively. A low level of λ indicates that the investor gives higher importance to the portfolio return. In particular, $\lambda = 0$ indicates that the decision maker is completely ignoring the risk of the portfolio. In this case, the optimal portfolio is usually concentrated only in the asset with the higher return.[11] On the contrary, high values of λ indicate a higher relevance to the risk with respect to the return. The extreme case of $\lambda = 1$ corresponds to the GMV portfolio, meaning that the decision maker completely ignores the portfolio return.

---

[10] Detailed results are reported in the Supplementary Material.

[11] The optimal portfolio when $\lambda = 0$ may not be unique, since more than one asset with the highest return can exist.

(a) $\lambda = 0.2$ **(b)** $\lambda = 0.4$

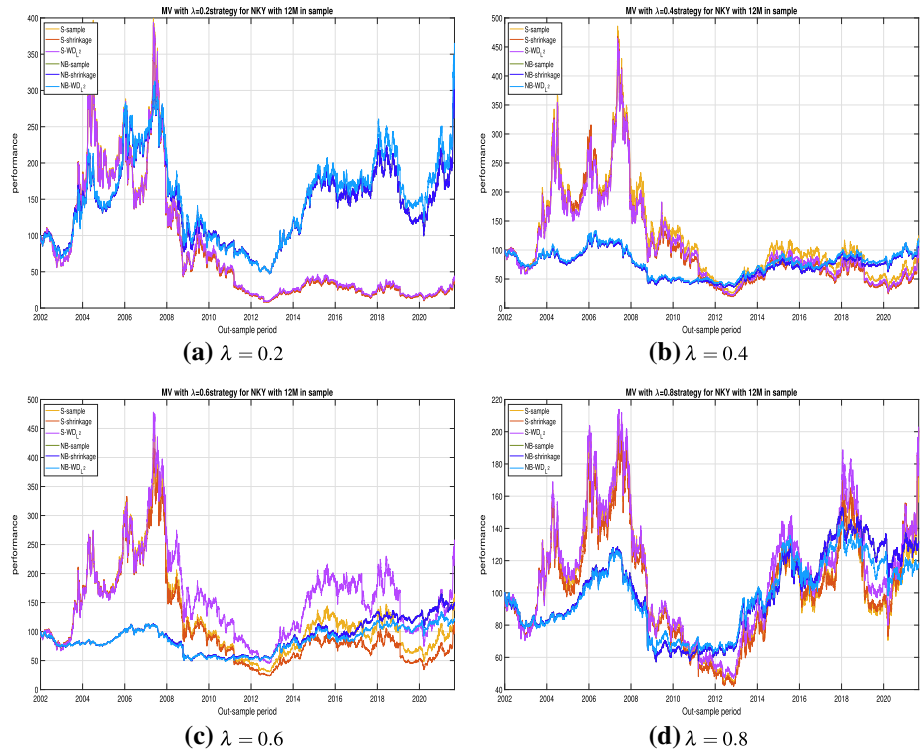**(c)** $\lambda = 0.6$ **(d)** $\lambda = 0.8$

Fig. 4 Out-of-sample performances for NKY portfolio with a rolling window of 12 months in-sample and 1 month out-sample. In **a–d** we report the out-of-sample performances for $S - sample$, $S - Shrinkage$, $S - WD_{L^2}$, $NB - sample$, $NB - Shrinkage$ and $NB - WD_{L^2}$ strategies according to alternative values of the trade-off parameter (namely, $\lambda = 0.2$, $\lambda = 0.4$, $\lambda = 0.6$ and $\lambda = 0.8$ respectively)

To this end, Fig. 4 reports the out-of-sample performances of the NKY dataset, with a buy and hold strategy of 12 months in-sample and one month out-of-sample, in case of the MV model. As previously described in Sect. 2.1, for this strategy μ and Σ have to be estimated. We estimate μ using the sample approach while Σ is estimated using the *sample*, the *shrinkage toward constant correlation* and the *weighted-depth L^2* methods. Notice that the matrix Σ is also used in the network-based approaches to construct the network and to obtain the interconnectedness matrix C. Figure 4a displays the out-of-sample performances obtained setting $\lambda = 0.2$, which means that the investor tends to prefer high potential returns with respect to low levels of uncertainty. Although it is not possible to define a univocal ranking between methods in terms of performance, we observe higher returns at the end of the period with the network-based approaches. Indeed, it is noticeable a fast decrease in the out-of-sample performance of the *Standard approaches* after the crisis of 2008 while the performance of the *Network Based approaches* is much more stable leading at the end of the period at a prevalence of NB models.

Post-crisis effect is also evident for the other values of λ (see Fig. 4b–d).

To have a complete view of the effect of the network based strategies on the MV model, we report in Table 3 the first four moments and alternative performance measures.

The results in Table 3 clearly shows that, for all considered values of λ, the MV network-based portfolio has a lower volatility than the corresponding standard approach. Moreover,

Table 3 Out-of-sample statistics for the MV model, in case of the NKY portfolio with one year rolling estimation window for the mean vector and the covariance matrix and one month for out-of-sample returns

	S-sample	S-shrinkage	S-WD_{L2}	NB-sample	NB-shrinkage	NB-WD_{L2}	S-sample	S-shrinkage	S-WD_{L2}	NB- sample	NB-shrinkage	NB-WD_{L2}
$\lambda = 0.2$							$\lambda = 0.4$					
$\mu^{\star}$	0.00032	0.00031	**0.00033**	0.00025	0.00025	0.00028	**0.00043**	0.00038	0.00041	0.00020	0.00024	0.00027
$\sigma^{\star}$	0.031	0.031	0.031	0.021	0.021	**0.021**	0.028	0.028	0.028	0.015	**0.015**	**0.015**
$skew^{\star}$	−0.744	−0.744	−0.737	−0.269	**−0.264**	−0.298	−0.460	−0.467	−0.470	−0.445	**−0.445**	−0.483
$kurt^{\star}$	13.138	13.090	13.061	7.555	**7.510**	7.632	8.179	8.149	**8.097**	9.002	9.002	9.303
$SR^{\star}$	0.010	0.010	0.011	0.012	0.012	**0.013**	0.016	0.014	0.014	0.014	**0.016**	**0.018**
$IR^{\star}(EW)$	0.007	0.006	0.007	0.018	0.018	**0.019**	0.010	0.011	0.012	0.011	**0.013**	**0.013**
$OR^{\star}$	1.031	1.029	1.032	1.060	1.060	**1.064**	1.046	1.040	1.045	1.039	1.044	**1.046**
$\lambda = 0.6$							$\lambda = 0.8$					
$\mu^{\star}$	0.00038	0.00032	**0.00047**	0.00018	0.00021	0.00024	0.00028	0.00030	**0.00031**	0.00019	0.00019	0.00022
$\sigma^{\star}$	0.024	0.024	0.024	**0.012**	**0.012**	**0.012**	0.018	0.018	0.019	**0.010**	**0.010**	**0.010**
$skew^{\star}$	−0.504	−0.483	**−0.422**	−0.636	−0.636	−0.584	−0.514	−0.485	−0.484	−0.532	−0.532	**−0.431**
$kurt^{\star}$	8.329	8.311	**7.466**	12.288	12.288	12.227	8.691	8.671	**8.282**	13.807	13.807	14.158
$SR^{\star}$	0.016	0.013	**0.020**	0.016	0.018	**0.021**	0.016	0.017	0.017	0.018	0.018	**0.021**
$IR^{\star}(EW)$	0.013	0.009	0.017	**0.018**	0.017	0.016	0.011	0.012	**0.013**	0.012	0.012	**0.013**
$OR^{\star}$	1.047	1.039	1.058	**1.059**	1.044	1.058	1.046	1.049	1.049	1.045	1.045	**1.054**

The best results are reported in bold for each measure. All the statistics are reported on daily bases

in the majority of the cases the network-based portfolios lead also to higher out-of-sample performances in terms of Sharpe ratio and Omega ratio. These results are in line with those obtained for the *risk-based* approaches, presented in Fig. 3 and Table 2, confirming that applying network tools to portfolio selection models may enhance the portfolio selection process.

The use of network tools to manage the optimal portfolio selection proved to be effective, especially in the case of *risk-based* strategies. Looking at the results of the MV portfolios, in the Supplementary Material we can observe that the network-based portfolios lead to better out-of-sample results for some levels of the trade-off parameters. However, there is not an absolute dominance of these approaches, since in some cases the standard methods behave better for specific values of the trade-off parameter. This behaviour depends on the trade-off parameter value and to the estimator used for the portfolio mean.

5 Conclusions

In this work we applied network tools to the most used portfolio models characterized by an objective function depending on the covariance matrix of assets. Following Clemente et al. (2019), we took advantage of the correlations network to capture the interconnectedness between assets, that explicitly enters through the clustering coefficient in the objective function. We extended the approach of Clemente et al. (2019), tested to the GMV model, proposing the application of network theory also for the most used Smart Beta models, as well as the MV model. We estimated the network-correlation structure through the sample (as in Clemente et al. 2019), the shrinkage toward the constant correlation and the weighted depth - L^2 approaches.

To test the robustness of our methodology, we performed numerical analyses, based on three large-dimensional portfolios. We implemented both the standard and the network-based models, using sample, shrinkage WD_{L^2} estimators for the covariance matrix, and we compared the out-of-sample performances based on a rolling sample optimization. The results obtained show in most cases the effectiveness of network-based portfolios compared to the standard approaches and to the equally weighted portfolios. Network-based strategies show higher out-of-sample performances and lower out-of-sample volatility, reducing the risk. Results appeared significant especially for Smart Beta strategies, which are based only on the risk measure, that refers to the covariance matrix or the interconnectedness matrix. Results in case of mean-variance portfolio do not provide a univocal ranking of the models. However, the network-based approaches lead to better results in the most cases, although there is a good percentage of cases in which the standard approaches behave better. This behaviour depends obviously on the trade-off parameter value and to the estimator used for the portfolio mean. We believe that better out-of-sample results can be obtained in case the network theory is used for the estimation of not only of the risk measure but also of the performance measure, which is left for future research.

In conclusion, we hope that this empirical analysis will help to shed some light on how network theory can be implemented in portfolio selection problems and encourage portfolio managers in considering and testing the network-based portfolio selection models as an alternative to the standard approaches.

References

Amenc, N., & Goltz, F. (2013). Smart Beta 2.0. *The Journal of Index Investing, 4*(3), 15–23.

Barrat, A., Barthélemy, M., Pastor-Satorras, R., & Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences, 101*(11), 3747–3752.

Bauder, D., Bodnar, T., Parolya, N., & Schmid, W. (2018). Bayesian mean-variance analysis: Optimal portfolio selection under parameter uncertainty. arXiv:1803.03573

Behr, P., Guettler, A., & Miebs, F. (2013). On portfolio optimization: Imposing the right constraints. *Journal of Banking & Finance, 37*(4), 1232–1242.

Benati, S., & Conde, E. (2022). A relative robust approach on expected returns with bounded CVar for portfolio selection. *European Journal of Operational Research, 296*(1), 332–352.

Bloomberg. (2012). *Bloomberg professional*. [online]. Subscription Service. Accessed November 30, 2012.

Bongini, P., Clemente, G., & Grassi, R. (2018). Interconnectedness, G-SIBs and network dynamics of global banking. *Finance Research Letters, 27*, 185–192.

Brennan, M. (1998). The role of learning in dynamic portfolio decisions. *European Finance Review, 1*, 295–306.

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.

Cerqueti, R., Ferraro, G., & Iovanella, A. (2018). A new measure for community structure through indirect social connections. *Expert Systems with Applications, 114*, 196–209.

Cesarone, F., Gheno, A., & Tardella, F. (2013). Learning & holding periods for portfolio selection models: A sensitivity analysis. *Applied Mathematical Sciences, 7*(100), 4981–4999.

Chopra, V., & Ziemba, W. (2013). The effect of errors in means, variances, and covariances on optimal portfolio choice. In: L. C. MacLean & W. T. Ziemba (Eds.), *Handbook of the fundamentals of financial decision making: Part I* (pp. 365–373). World Scientific.

Choueifaty, Y., & Coignard, Y. (2008). Towards maximum diversification. *Journal of Portfolio Management, 35*(1), 40–51.

Clemente, G., & Grassi, R. (2018). Directed clustering in weighted networks: A new perspective. *Chaos, Solitons & Fractals, 107*, 26–38.

Clemente, G., Grassi, R., & Hitaj, A. (2019). Asset allocation: New evidence through network approaches. *Annals of Operations Research, 299*, 61–80.

Consigli, G., Hitaj, A., & Mastrogiacomo, E. (2018). Portfolio choice under cumulative prospect theory: Sensitivity analysis and an empirical study. *Computational Management Science, 16*(4), 1–26.

de Paulo, W. L., de Oliveira, E. M., & do Valle Costa, O. L. (2016). Enhanced index tracking optimal portfolio selection. *Finance Research Letters, 16*, 93–102.

Deguest, R., Martellini, L., & Meucci, A. (2013). *Risk parity and beyond-from asset allocation to risk allocation decisions*. Working paper.

DeMiguel, V., Garlappi, L., & Uppal, R. (2007). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The Review of Financial studies, 22*(5), 1915–1953.

Dentcheva, D., & Ruszczynski, A. (2003). Optimization with stochastic dominance constraints. *SIAM Journal on Optimization, 14*(2), 548–566.

Elton, E., & Gruber, M. (1973). Estimating the dependence structure of share prices—Implications for portfolio selection. *The Journal of Finance, 28*(5), 1203–1232.

Fabozzi, F. J., Huang, D., & Zhou, G. (2010). Robust portfolios: Contributions from operations research and finance. *Annals of Operations Research, 176*(1), 191–220.

Fagiolo, G. (2007). Clustering in complex directed networks. *Physical Review E, 76*(2), 026107.

Fliege, J., & Werner, R. (2014). Robust multiobjective optimization & applications in portfolio optimization. *European Journal of Operational Research, 234*(2), 422–433.

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

Georgantas, A., Doumpos, M., & Zopounidis, C (2021). Robust optimization approaches for portfolio selection: a comparative analysis. Annals of Operations Research. https://doi.org/10.1007/s10479-021-04177-y

Giudici, P., & Spelta, A. (2016). Graphical network models for international financial flows. *Journal of Business & Economic Statistics, 34*(1), 128–138.

Goldfarb, D., & Iyengar, G. (2003). Robust portfolio selection problems. *Mathematics of Operations Research, 28*(1), 1–38.

Guastaroba, G., Mansini, R., Ogryczak, W., & Speranza, M. G. (2016). Linear programming models based on omega ratio for the enhanced index tracking problem. *European Journal of Operational Research, 251*(3), 938–956.

Hitaj, A., Martellini, L., & Zambruno, G. (2012). Optimal hedge fund allocation with improved estimates for coskewness and cokurtosis parameters. *The Journal of Alternative Investments, 14*(3), 6–16.

Hitaj, A., & Zambruno, G. (2016). Are smart beta strategies suitable for hedge fund portfolios? *Review of Financial Economics, 29*, 37–51.

Hitaj, A., & Zambruno, G. (2018). Portfolio optimization using modified herfindahl constraint. In G. Consigli, S. Stefani, & G. Zambruno (Eds.), *Handbook of Recent Advances in Commodity and Financial Modeling* (pp. 211–239). Springer.

Iorio, C., Frasso, G., D'Ambrosio, A., & Siciliano, R. (2018). A p-spline based clustering approach for portfolio selection. *Expert Systems with Applications, 95*, 88–103.

Jagannathan, R., & Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance, 58*(4), 1651–1683.

Jobson, J. D., & Korkie, R. M. (1981). Putting Markowitz theory to work. *The Journal of Portfolio Management, 7*(4), 70–74.

Jorion, P. (1985). International portfolio diversification with estimation risk. *Journal of Business*, 259–278.

Jorion, P. (1986). Bayes-stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis, 21*(3), 279–292.

Jorion, P. (1992). Portfolio optimization in practice. *Financial Analysts Journal, 48*(1), 68–74.

Kapsos, M., Christofides, N., & Rustem, B. (2014). Worst-case robust omega ratio. *European Journal of Operational Research, 234*(2), 499–507.

Keating, C., & Shadwick, W. F. (2002). A universal performance measure. *Journal of Performance Measurement, 6*(3), 59–84.

Ledoit, O., and Michael, W. (2004)*H*oney, I shrunk the sample covariance matrix. The Journal of Portfolio Management *30 (4) : 110-119.*

Li, Y., Jiang, X. F., Tian, Y., Li, S. P., & Zheng, B. (2019). Portfolio optimization based on network topology. *Physica A: Statistical Mechanics and its Applications, 515*, 671–681.

Maillard, S., Roncalli, T., & Teiletche, J. (2010). On the properties of equally-weighted risk contributions portfolios. *Journal of Portfolio Management, 36*(4), 60–70.

Mantegna, R. N. (1999). Hierarchical structure in financial markets. *The European Physical Journal B-Condensed Matter and Complex Systems, 11*(1), 193–197.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance, 7*(1), 77–91.

Martellini, L., & Ziemann, V. (2009). Improved estimates of higher-order comoments and implications for portfolio selection. *The Review of Financial Studies, 23*(4), 1467–1502.

McAssey, M. P., & Bijma, F. (2015). A clustering coefficient for complete weighted networks. *Network Science, 3*(2), 183–195.

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics, 8*(4), 323–361.

Minoiu, C., & Reyes, J. (2013). A network analysis of global banking: 1978–2010. *Journal of Financial Stability, 9*(2), 168–184. https://doi.org/10.1209/0295-5075/115/18002

Onnela, J., Chakraborti, A., Kaski, K., Kertesz, J., & Kanto, A. (2003a). Asset trees and asset graphs in financial markets. *Physica Scripta, 2003*(T106), 48.

Onnela, J. P., Chakraborti, A., Kaski, K., Kertész, J., & Kanto, A. (2003b). Dynamics of market correlations: Taxonomy and portfolio analysis. *Physical Review E, 68*, 056110.

Pandolfo, G., Iorio, C., Siciliano, R., & D'Ambrosio, A. (2020). Robust mean-variance portfolio through the weighted lp depth function. *Annals of Operations Research, 292*(1), 519–531.

Peralta, G., & Zareei, A. (2016). A network approach to portfolio selection. *Journal of Empirical Finance, 38*, 157–180.

Pozzi, F., Di Matteo, T., & Aste, T. (2013). Spread of risk across financial markets: Better to invest in the peripheries. *Scientific Reports, 3*, 1665.

Qian, E. (2006). On the financial interpretation of risk contribution: Risk budgets do add up. *Journal of Investment Management, 4*(4), 41–51.

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

Scutella, M. G., & Recchia, R. (2013). Robust portfolio asset allocation and risk measures. *Annals of Operations Research, 204*(1), 145–169.

Tabak, B., Takamib, M., Rochac, J., Cajueirod, D., & Souzae, S. (2014). Directed clustering coefficient as a measure of systemic risk in complex banking networks. *Physica A: Statistical Mechanics and its Applications, 394*, 211–216.

Výrost, T., Lyócsa, Š., & Baumöhl, E. (2019). Network-based asset allocation strategies. *The North American Journal of Economics and Finance, 47*, 516–536.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge University Press.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393*(6684), 440.

Zhan, H. C. J., Rea, W., & Rea, A. (2015). A comparision of three network portfolio selection methods—Evidence from the dow jones. arXiv:1512.01905

Zhu, S., & Fukushima, M. (2009). Worst-case conditional value-at-risk with application to robust portfolio management. *Operations Research, 57*(5), 1155–1168.

Zuo, Y. (2004). Robustness of weighted l p-depth and l p-median. *Allgemeines Statistisches Archiv, 88*(2), 215–234.

Zuo, Y., Cui, H., & He, X. (2004). On the Stahel–Donoho estimator and depth-weighted means of multivariate data. *The Annals of Statistics, 32*(1), 167–188.

Zymler, S., Kuhn, D., & Rustem, B. (2013). Worst-case value at risk of nonlinear portfolios. *Management Science, 59*(1), 172–188.