



Enhancing Numerical Stability in Portfolio Optimization via Numerical Rank: The δ -Stable Part of the Efficient Frontier

Claudia Fassino¹ · Pierpaolo Uberti² 

Received: 9 February 2026 / Accepted: 21 September 2026
© The Author(s) 2026

Abstract

The fact that the solution of the Markowitz model is numerically unstable is recognized as one of the principal reasons for its poor out-of-sample results when applied in practice. In the present paper, we show that, alongside standard instability sources such as the estimation of the covariance matrix and returns, there is a structural source of instability. This source depends on the geometry of the admissible region defined by the linear constraints of the optimization problem namely, the budget constraint and the target expected return. We provide theoretical results based on numerical arguments and propose a set of suitable target returns that restricts the standard mean-variance efficient frontier to a subset of portfolios that are numerically stable with respect to a given parameter δ . The increased stability of these portfolios is achieved by imposing a condition that preserves the numerical rank of the constraint matrix. An application to various real-world financial databases highlights the effectiveness of our proposal in reducing the numerical instability of the optimal solution.

Keywords Numerical stability · Portfolio optimization · Numerical rank · Condition number

1 Introduction

The Markowitz model [1] is widely considered the foundation of modern portfolio theory. Since its inception, the model and its underlying assumptions have been

Claudia Fassino and Pierpaolo Uberti contributed equally to this work.

✉ Pierpaolo Uberti
pierpaolo.uberti@unimib.it

Claudia Fassino
fassino@dima.unige.it

¹ Department of Mathematics, University of Genoa, Via Dodecaneso 35, Genoa 16146, Italy

² Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Via Bicocca degli Arcimboldi 8, Milan 20126, Italy

extensively studied and debated in the literature. One of the primary challenges in applying the mean-variance framework in practice is the numerical instability of the optimal solution; specifically, small perturbations in the input parameters, the covariance matrix and the vector of expected returns, can produce large variations in the optimal allocation. Consequently, portfolios that are contiguous on the efficient frontier may exhibit radically different compositions, both in terms of weights and the specific assets selected, which complicates practical implementation. This instability is often cited as a primary reason for the model's poor out-of-sample performance when compared to naive allocation strategies, such as the $1/N$ rule [2]. Within the context of the Markowitz model, the uncertainty arising from parameter estimation has been rigorously examined (see, for example, [3, 4]).

Furthermore, in [5], the authors provide evidence that errors in estimating expected returns are roughly ten times more damaging than errors in variances and twenty times more damaging than errors in covariances. This explains why minimum variance portfolios which ignore expected returns often outperform full mean-variance models. In [4], the authors document that simply “plugging in” sample estimates leads to massive performance degradation, suggesting that parameter uncertainty is the primary obstacle to successful risk-based allocation.

Following these milestones, the academic debate has evolved in several directions which, upon deeper analysis, represent different perspectives on the same fundamental problem. A substantial body of research [6–9] continues to confirm the superiority of naive allocation over complex optimization in out-of-sample tests. Conversely, some authors defend the optimization framework [10–12], often emphasizing that the success of these models is contingent upon the accuracy of the parameter forecasts.

To mitigate numerical instability, numerous approaches have been developed, most of which are motivated by estimation uncertainty. These strategies can be broadly classified into two main categories. The first major stream focuses on the objective function, specifically the regularization of the covariance matrix. This includes Bayesian procedures [13], shrinkage estimation [14, 15] which has evolved from linear models to sophisticated nonlinear approaches [16, 17] and Lasso-type regularization or robust covariance techniques [18, 19]. A parallel branch focuses on robust optimization, which explicitly incorporates parameter uncertainty by assuming that true parameters lie within a predefined “uncertainty set” [19–21]. By seeking a portfolio that performs best under a “worst-case” scenario, these models attempt to shield the allocation from the extreme sensitivity typical of traditional optimizers (see [22, 23] for comprehensive reviews).

The 2008 global financial crisis further exposed the inability of standard models to handle extreme market conditions, leading to the rise of new paradigms: risk parity and maximum diversification. Risk parity shifts the focus from capital allocation to risk allocation, ensuring each asset contributes equally to total portfolio risk [24, 25]. Alternatively, diversification maximization aims to maximize the portfolio's “breadth” rather than minimizing a risk proxy. Strategies include optimizing the diversification ratio [26], maximizing the effective number of independent risk factors [27], or utilizing entropy-based measures [28].

The problem of numerical instability in asset allocation models remains a highly relevant topic in modern quantitative finance, with many recent studies proposing

alternative strategies to robustify portfolio weights and mitigate estimation error [29–32].

Despite these advances, the fact that numerical instability may also stem from the structure of the constraints has been only marginally discussed. For instance, [33] suggest that the estimation error of the mean is so significant that ignoring the mean constraint entirely can sometimes improve results. Similarly, [34] describes the tendency of the optimizer to over-allocate to assets with large estimation errors as “error-maximization.” While these works support the intuition that numerical issues could depend also on linear restrictions, there is a lack of approaches rooted in numerical mathematics to address this directly. Only recently, the structure of linear restrictions has been formally recognized as a distinct source of instability, with [35] proposing to reduce this sensitivity through optimal rescaling techniques.

This paper builds on that foundation by examining the numerical rank and conditioning of the linear constraints. Our approach highlights that numerical instability can arise even within a strictly deterministic framework, independent of the uncertainty associated with parameter estimation. Consequently, our results address a structural phenomenon not directly covered by standard approaches focusing on estimation error. We prove that the choice of the target expected return can, in itself, be a driver of numerical instability. The intuition behind our proposal is as follows: consider the two linear constraints, the budget constraint and the target expected return. In the framework of exact mathematics, two extreme cases exist. First, if the two vectors defining the constraints are linearly dependent, one constraint becomes redundant, and the efficient frontier collapses to the global minimum variance portfolio. Second, if they are linearly independent, the optimization is solvable for any target return and the entire frontier exists. We focus on the intermediate case between these extremes, which is highly relevant for practical applications. When the matrix of the constraints is ill-conditioned, choosing extreme values for the target return can “stress” the admissible region, causing specific segments of the efficient frontier to become numerically unstable. Therefore, we propose to restrict the admissible portfolio target returns to preserve the numerical rank of the matrices involved in the optimization. We highlight the consequences of this choice on the condition number of the constraint matrix and show, through both toy examples and real-world data, that our proposal effectively reduces numerical instability. The segment of the frontier that remains stable with respect to a parameter δ (which defines the numerical rank) is called the δ -stable frontier. Our empirical findings confirm that this approach significantly reduces instability in practice. We also demonstrate, through an out-of-sample experiment based on empirical data, that stable portfolios yield compelling risk-adjusted returns compared to less stable ones once transaction costs are accounted for.

The paper is organized as follows: Sect. 2 collects assumptions and preliminary results. Section 3 presents the main result regarding the δ -stable frontier. Section 4 contains the empirical application, including a discussion on the choice of δ , the data description, and the results. Section 5 concludes the paper. Appendices A and B contain the technical results and the graphical stability analysis for a selection of the databases, respectively.

2 Assumptions, Definitions, and Preliminary Results

We assume that $n > 2$ risky assets are available on a given market. The generic column vector $\mathbf{x} \in \mathbb{R}^n$ identifies a portfolio when it satisfies the budget constraint $\sum_{i=1}^n x_i = 1$; in that case, its i^{th} entry x_i ($i = 1, \dots, n$) represents the share invested in the i^{th} asset. The covariance matrix of the assets' returns, denoted by $V \in \mathbb{R}^{n \times n}$, is assumed to be square, symmetric, positive-definite, and consequently, non-singular. Finally, the column vector $\boldsymbol{\mu} \in \mathbb{R}^n$ contains the expected returns of the n assets, and $\mathbf{1} = (1, \dots, 1)^t \in \mathbb{R}^n$ represents the vector with unitary entries. We consider the following standard portfolio optimization problem:

Problem 1

$$\begin{aligned} & \text{Minimize}_{\mathbf{x}} \quad \sigma_p^2 = \mathbf{x}^t V \mathbf{x} \\ & \text{subject to} \quad \boldsymbol{\mu}^t \mathbf{x} = \mu_p \\ & \quad \quad \quad \mathbf{1}^t \mathbf{x} = 1 \\ & \quad \quad \quad x_i \geq 0 \quad \text{for } i = 1, \dots, n \end{aligned}$$

where σ_p^2 and μ_p are, respectively, the variance and the expected return of the portfolio p identified by the vector of weights $\mathbf{x}$. We denote $x_{lo}^*(\mu_p)$ as the solution to Problem 1, emphasizing that it is a function of μ_p since the solution is unique for each value of the parameter. This portfolio optimization problem is known to suffer from numerical instability (refer to the introduction for the literature on the point); that is, small perturbations in the input data can cause considerable variations in the optimal allocation. We show how the choice of μ_p impacts this stability by measuring the portfolio TurnOver (TO), a formal definition of which is provided in the application. In an economic context, TO is more than a mere metric for the sensitivity of the optimal solution to variations in μ_p ; it also captures the effective cost of implementing the investment strategy's allocation.

The efficient frontier F_{lo} is the set containing the solutions to Problem 1 as μ_p varies:

$$F_{lo} = \{x_{lo}^*(\mu_p) \in \mathbb{R}^n, \forall \mu_p \in [\min(\boldsymbol{\mu}), \max(\boldsymbol{\mu})]\}.$$

We note that to obtain F_{lo} , the choice of μ_p must be restricted to the interval between the minimum and maximum values of $\boldsymbol{\mu}$. We now introduce some technical concepts and results that will be used throughout the remainder of this paper. First, we recall the definition of the δ -numerical rank of a matrix (see [36]).

Definition 1 Given a real number $\delta \geq 0$ and a matrix $A \in \mathbb{R}^{m \times n}$, such that $\text{rank}(A) = p$, the δ -numerical rank of A , $P_\delta(A)$, is given by

$$P_\delta(A) = \text{Card} \{\lambda_i : \lambda_i > \delta, \text{ for } i = 1, \dots, p\}$$

where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p > 0$ are the singular values of A taken, as usual, in decreasing order and $\text{Card}(\cdot)$ is the cardinality function.

We note that for a generic matrix $A \in \mathbb{R}^{m \times n}$, $P_\delta(A) = \text{rank}(A)$ if $\delta = 0$ or when $0 \leq \delta \leq \lambda_p$. From the Eckart-Young theorem (see [37]), if $P_\delta(A) < \text{rank}(A)$, there

exists a matrix $\tilde{A}$ such that $\|A - \tilde{A}\|_2 \leq \delta$ and $\text{rank}(\tilde{A}) = P_\delta(A)$; that is, A is very close to a rank-deficient matrix.

We also recall the definition of the condition number of a matrix (see [38, Section 2.1]), which measures how sensitive the solutions of a linear system are with respect to perturbations in the input data and rounding errors in the solution process.

Definition 2 Let $A \in \mathbb{R}^{m \times n}$ such that $\text{rank}(A) = \min\{m, n\}$; the (2-norm) *condition number* of A , denoted by $K_2(A)$, is given by the following:

$$K_2(A) := \frac{\lambda_1(A)}{\lambda_{\min\{m,n\}}(A)}.$$

Subsequently, we investigate the behavior of the equality constraints of Problem 1 by analyzing their numerical rank and condition number.

The linear equality constraints of Problem 1 can be rewritten as $B\mathbf{x} = [\mu_p \ 1]^t$, where the matrix $B \in \mathbb{R}^{2 \times n}$ ($n > 2$) is defined as

$$B = \begin{pmatrix} \boldsymbol{\mu}^t \\ \mathbf{1}^t \end{pmatrix} = \begin{pmatrix} \beta \mathbf{1}^t + \boldsymbol{\rho}_\mu^t \\ \mathbf{1}^t \end{pmatrix},$$

and $\boldsymbol{\rho}_\mu = \boldsymbol{\mu} - \beta \mathbf{1} \in \mathbb{R}^n$ is the residual of the linear system $\mathbf{1}\beta = \boldsymbol{\mu}$. Consequently, $\beta = \frac{\mathbf{1}^t \boldsymbol{\mu}}{n}$ is the arithmetic mean of the entries of $\boldsymbol{\mu}$, and $\boldsymbol{\rho}_\mu^t \mathbf{1} = 0$.

In the following, we assess whether B is close to a rank-deficient matrix by examining its numerical rank and its condition number. The following proposition provides a computationally efficient way to detect if $P_\delta(B) = 1$, thereby avoiding the need to calculate the singular values of B .

Proposition 1 Given $0 < \delta < \sqrt{n}$, we have

$$P_\delta(B) = 1 \quad \text{if and only if} \quad \|\boldsymbol{\rho}_\mu\|_2 \leq \delta \sqrt{\frac{n}{n - \delta^2} \beta^2 + 1}.$$

Proof The singular values $\lambda_1 \geq \lambda_2 \geq 0$ of B correspond to the square roots of the eigenvalues of

$$BB^t = \begin{pmatrix} \|\boldsymbol{\rho}_\mu\|_2^2 + n\beta^2 & n\beta \\ n\beta & n \end{pmatrix},$$

and can be calculated by solving the characteristic equation:

$$\sigma^2 - \sigma \left(\|\boldsymbol{\rho}_\mu\|_2^2 + n(\beta^2 + 1) \right) + n\|\boldsymbol{\rho}_\mu\|_2^2 = 0,$$

which yields

$$\lambda_{1,2}^2 = \frac{(\|\boldsymbol{\rho}_\mu\|_2^2 + n(\beta^2 + 1)) \pm \sqrt{(\|\boldsymbol{\rho}_\mu\|_2^2 + n(\beta^2 + 1))^2 - 4n\|\boldsymbol{\rho}_\mu\|_2^2}}{2}. \quad (1)$$

The condition $\lambda_2^2 \leq \delta^2$ is equivalent to

$$\sqrt{(\|\rho_\mu\|_2^2 + n(\beta^2 + 1))^2 - 4n\|\rho_\mu\|_2^2} \geq \|\rho_\mu\|_2^2 + n(\beta^2 + 1) - 2\delta^2,$$

which simplifies to

$$\|\rho_\mu\|_2^2 \leq \delta^2 \frac{n(\beta^2 + 1) - \delta^2}{n - \delta^2} = \delta^2 \left(\frac{n}{n - \delta^2} \beta^2 + 1 \right).$$

Therefore, since $P_\delta(B) = 1$ if and only if $\lambda_2 \leq \delta$, we conclude that

$$P_\delta(B) = 1 \Leftrightarrow \|\rho_\mu\|_2 \leq \delta \sqrt{\frac{n}{n - \delta^2} \beta^2 + 1}.$$

□

From Eq. 1, it follows that when $\|\rho_\mu\|_2^2$ takes small values, λ_2 approaches zero, causing the matrix B to become ill-conditioned as $K_2(B) = \lambda_1/\lambda_2$ increases significantly. In the following, we analyze the scenario where $\|\rho_\mu\|_2^2$ is sufficiently small to require a dedicated numerical strategy to mitigate the instability stemming from the ill-conditioning of matrix B .

3 The Opportune Choice of μ_p

In this section, we address the issues that arise when the condition number of B is large due to the small magnitude of the vector μ . In such scenarios, μ_p must be selected within an appropriate interval to ensure that minor variations in its value do not induce severe perturbations in the optimal solution, as illustrated in the following example.

Example 1 Let the linear constraints be given by

$$\begin{cases} \varepsilon_1 x_1 + \cdots + \varepsilon_n x_n = \mu_p \\ x_1 + \cdots + x_n = 1 \end{cases} \quad \text{with } |\varepsilon_i| < \varepsilon \ll 1.$$

Roughly speaking, the coefficients of the first equation are negligible compared to those of the second equation, which causes B to be ill-conditioned. Consequently, a small perturbation in μ_p can lead to substantial differences in the solution.

For example, if $\mu_p = 0$, we have $\sum_{i=1}^n \varepsilon_i x_i = 0$. If we instead consider a perturbed value $\hat{\mu}_p \neq 0$, the corresponding perturbed solution $\hat{\mathbf{x}} = (\hat{x}_1, \dots, \hat{x}_n)^t$ satisfies $\sum_{i=1}^n \varepsilon_i \hat{x}_i = \hat{\mu}_p$. It follows that $\sum_{i=1}^n \varepsilon_i (\hat{x}_i - x_i) = \hat{\mu}_p$, and thus, $|\hat{\mu}_p| \leq \varepsilon \sum_{i=1}^n |\hat{x}_i - x_i|$. Under the assumption that $\|\mathbf{x}\|_1 = 1$, we obtain

$$\frac{\|\hat{\mathbf{x}} - \mathbf{x}\|_1}{\|\mathbf{x}\|_1} = \sum_{i=1}^n |\hat{x}_i - x_i| \geq \frac{|\hat{\mu}_p|}{\varepsilon}.$$

Given a small ε , minor perturbations in μ_p can cause a large variation in the entries of the solution. For this reason, in this example, the choice of $\widehat{\mu}_p$ cannot differ significantly from 0. Note that if $\widehat{\mu}_p$ is small enough, both matrices B and $\widehat{B}$,

$$\widehat{B} = \left(B \middle| \begin{smallmatrix} \widehat{\mu}_p \\ 1 \end{smallmatrix} \right) = \left(\begin{smallmatrix} \varepsilon_1 & \dots & \varepsilon_n \\ 1 & \dots & 1 \end{smallmatrix} \middle| \begin{smallmatrix} \widehat{\mu}_p \\ 1 \end{smallmatrix} \right)$$

are ill-conditioned.

The previous example suggests that μ_p should be chosen sufficiently close to the mean value β . To this end, we propose a strategy in the following to determine an interval, centered at β , from which a suitable μ_p can be selected. This strategy is inspired by the Rouché–Capelli theorem, which states that a linear system $A\mathbf{x} = \mathbf{b}$ is consistent (i.e., it admits at least one solution) if and only if $\text{rank}(A) = \text{rank}([A|\mathbf{b}])$. Accordingly, we suggest choosing the value of μ_p such that, for a given small threshold $\delta \geq 0$, if $P_\delta(B) = 1$ then $P_\delta(\widehat{B}) = 1$, where $\widehat{B} = [B | [\mu_p \ 1]^t]$ is the augmented matrix associated with the linear system $B\mathbf{x} = [\mu_p \ 1]^t$.¹ A key motivation behind this framework is to provide a criterion aimed at enhancing the numerical stability of the optimal allocation, whose practical effectiveness is supported in the application.

Let us consider the matrix B . Using the notation of the previous section, B can be rewritten as

$$B = \begin{pmatrix} \beta \mathbf{1}^t + \boldsymbol{\rho}_\mu^t \\ \mathbf{1}^t \end{pmatrix} \quad \text{with} \quad \beta = \frac{\boldsymbol{\mu}^t \mathbf{1}}{n}.$$

We express the matrix $\widehat{B}$ as

$$\widehat{B} = \begin{pmatrix} \widehat{\boldsymbol{\mu}}^t \\ \widehat{\mathbf{1}}^t \end{pmatrix} \quad \text{with} \quad \widehat{\boldsymbol{\mu}}^t = (\boldsymbol{\mu}^t, \mu_p) \in \mathbb{R}^{n+1}, \quad \widehat{\mathbf{1}}^t = (\mathbf{1}^t, 1) \in \mathbb{R}^{n+1}.$$

Theorem 2 states that if $P_\delta(B) = 1$, there exists an interval in $\mathbb{R}$ such that if μ_p belongs to this interval, then $P_\delta(\widehat{B}) = 1$ as well.

Theorem 2 Let $0 < \delta < \sqrt{n}$ and let

$$R_0 = \sqrt{1 + \frac{1}{n - \delta^2}} \sqrt{\delta^2 \left(\frac{n}{n - \delta^2} \beta^2 + 1 \right) - \|\boldsymbol{\rho}_\mu\|_2^2}.$$

If $\|\boldsymbol{\rho}_\mu\|_2^2 \leq \delta^2 \left(\frac{n}{n - \delta^2} \beta^2 + 1 \right)$ and $\mu_p \in \left[\frac{n}{n - \delta^2} \beta - R_0, \frac{n}{n - \delta^2} \beta + R_0 \right]$, then $P_\delta(B) = P_\delta(\widehat{B}) = 1$.

Proof First, since $0 < \delta < \sqrt{n}$, Proposition 1 implies that $P_\delta(B) = 1$.

¹ When $\delta = 0$, we require that $\text{rank}(B) = \text{rank}([B | [\mu_p \ 1]^t])$, which is the exact condition of the Rouché–Capelli theorem for the existence of a solution.

Applying the results of the previous section to the $(n + 1)$ -dimensional case, we can rewrite the matrix $\widehat{B}$ as

$$\widehat{B} = \begin{pmatrix} \widehat{\beta} \widehat{\mathbf{1}}^t + \boldsymbol{\rho}_{\widehat{\mu}}^t \\ \widehat{\mathbf{1}}^t \end{pmatrix} \quad \text{with} \quad \widehat{\beta} = \frac{\widehat{\mu}^t \widehat{\mathbf{1}}}{n + 1} = \frac{\mathbf{1}^t \boldsymbol{\mu} + \mu_p}{n + 1} = \frac{n\beta + \mu_p}{n + 1}.$$

As a consequence,

$$\begin{aligned} \boldsymbol{\rho}_{\widehat{\mu}} &= \widehat{\mu} - \widehat{\beta} \widehat{\mathbf{1}} = \begin{pmatrix} \boldsymbol{\mu} - \widehat{\beta} \mathbf{1} \\ \mu_p - \widehat{\beta} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\rho}_{\mu} + \beta \mathbf{1} - \widehat{\beta} \mathbf{1} \\ \mu_p - \frac{n\beta + \mu_p}{n + 1} \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{\rho}_{\mu} + (\beta - \widehat{\beta}) \mathbf{1} \\ \frac{n}{n + 1} (\mu_p - \beta) \end{pmatrix} = \begin{pmatrix} \boldsymbol{\rho}_{\mu} + \frac{1}{n + 1} (\beta - \mu_p) \mathbf{1} \\ \frac{n}{n + 1} (\mu_p - \beta) \end{pmatrix}. \end{aligned}$$

Since $\boldsymbol{\rho}_{\mu}^t \mathbf{1} = 0$, it holds that

$$\begin{aligned} \|\boldsymbol{\rho}_{\widehat{\mu}}\|_2^2 &= \boldsymbol{\rho}_{\widehat{\mu}}^t \boldsymbol{\rho}_{\widehat{\mu}} = \boldsymbol{\rho}_{\mu}^t \boldsymbol{\rho}_{\mu} + \frac{1}{(n + 1)^2} (\beta - \mu_p)^2 \mathbf{1}^t \mathbf{1} + \frac{n^2}{(n + 1)^2} (\mu_p - \beta)^2 \\ &= \|\boldsymbol{\rho}_{\mu}\|_2^2 + \frac{(\beta - \mu_p)^2}{(n + 1)^2} (n + n^2) = \|\boldsymbol{\rho}_{\mu}\|_2^2 + \frac{n}{n + 1} (\beta - \mu_p)^2. \end{aligned}$$

From Proposition 1, since $0 < \delta < \sqrt{n} < \sqrt{n + 1}$, $P_{\delta}(\widehat{B}) = 1$ if and only if

$$\|\boldsymbol{\rho}_{\widehat{\mu}}\|_2^2 \leq \delta^2 \left(\frac{n + 1}{n + 1 - \delta^2} \widehat{\beta}^2 + 1 \right).$$

Setting

$$R_0 = \sqrt{1 + \frac{1}{n - \delta^2}} \sqrt{\delta^2 \left(\frac{n}{n - \delta^2} \beta^2 + 1 \right) - \|\boldsymbol{\rho}_{\mu}\|_2^2},$$

which is well defined since $\|\boldsymbol{\rho}_{\mu}\|_2^2 \leq \delta^2 \left(\frac{n}{n - \delta^2} \beta^2 + 1 \right)$, the condition

$$\|\boldsymbol{\rho}_{\widehat{\mu}}\|_2^2 \leq \delta^2 \left(\frac{n + 1}{n + 1 - \delta^2} \widehat{\beta}^2 + 1 \right)$$

is satisfied (see Appendix 1, Prop. 4) for any $\mu_p \in \left[\frac{n}{n - \delta^2} \beta - R_0, \frac{n}{n - \delta^2} \beta + R_0 \right]$.

In conclusion, choosing $\mu_p \in \left[\frac{n}{n - \delta^2} \beta - R_0, \frac{n}{n - \delta^2} \beta + R_0 \right]$ ensures that $\|\boldsymbol{\rho}_{\widehat{\mu}}\|_2 \leq \delta \sqrt{\frac{n + 1}{n + 1 - \delta^2} \widehat{\beta}^2 + 1}$, and thus, by Proposition 1, $P_{\delta}(\widehat{B}) = 1$. $\square$

We emphasize that the previous result is useful only when, given a small parameter δ , $P_{\delta}(B) = 1$, i.e., when B is a δ -numerical rank-deficient matrix, a case that is very common in practical applications. Conversely, when $P_{\delta}(B) = 2$ and B has full

δ -numerical rank, no numerical issues arising from the choice of μ_p are expected, allowing the parameter to be set at any desired level.

We now define the concept of the δ -stable part of the frontier.

Definition 3 Given $\delta > 0$, and the set F_{lo} defined above, the δ -stable part of the frontier, denoted by F_{lo}^δ , is defined as

$$F_{lo}^\delta = \left\{ x_{lo}^*(\mu_p) \in F_{lo} : \mu_p \in \left[\frac{n}{n-\delta^2}\beta - R_0, \frac{n}{n-\delta^2}\beta + R_0 \right] \cap [\min(\mu), \max(\mu)] \right\}.$$

Remark 1 From a strictly economic perspective, we cannot rule out the possibility that the δ -stable part of the frontier is located on a dominated portion of the efficient frontier. In other words, it may be possible to find portfolios with higher returns and lower risks elsewhere. Under short-sale constraints ($\mathbf{x} \geq \mathbf{0}$), the dominant portion of the efficient frontier corresponds to target return levels $\mu_p \geq \mu_{GMV}$, where μ_{GMV} is the expected return of the constrained global minimum variance (GMV) portfolio obtained from Problem 1, when the short-sale constraints and the expected return constraint are avoided or non-active.

In the absence of short-sale constraints (or when nonnegativity constraints are non-binding), the expected return of the GMV portfolio simplifies to the closed-form expression $\mu_{GMV} = \frac{\mu^t V^{-1} \mathbf{1}}{\mathbf{1}^t V^{-1} \mathbf{1}}$. Setting $\mu_p = \frac{\mu^t V^{-1} \mathbf{1}}{\mathbf{1}^t V^{-1} \mathbf{1}}$, and using our previous notation, we can express this benchmark return as follows:

$$\mu_p = \frac{(\beta \mathbf{1}^t + \rho_\mu^t) V^{-1} \mathbf{1}}{\mathbf{1}^t V^{-1} \mathbf{1}} = \beta + \frac{\rho_\mu^t V^{-1} \mathbf{1}}{\mathbf{1}^t V^{-1} \mathbf{1}}$$

From this, it straightforwardly follows that if $\rho_\mu^t V^{-1} \mathbf{1} \leq 0$, there exists a segment of the efficient frontier composed of dominant portfolios that is also δ -stable. Conversely, if $\rho_\mu^t V^{-1} \mathbf{1} > 0$, a δ -stable part of the frontier exists only when $\frac{\mu^t V^{-1} \mathbf{1}}{\mathbf{1}^t V^{-1} \mathbf{1}} \leq \beta + R_0$.

While partially counter-intuitive, this approach suggests choosing portfolios that may be dominated in the standard unconstrained mean-variance sense to prioritize numerical stability. This idea aligns with a broad stream of research highlighting that standard allocation techniques often yield poor out-of-sample results due to allocation instability, where the resulting transaction costs erode any potential performance gains from active management. For example, in [2], the authors show that the equally weighted stable allocation is highly competitive with standard minimization approaches, while in [15], the authors demonstrate the increased stability related to the estimation of the covariance matrix. Another allocation explicitly designed for stability is the risk parity portfolio, where risk is equally distributed across asset classes; see [24].

The following toy example makes our proposal more explicit.

Example 2 We set $n = 10$ and

$$B = \begin{pmatrix} \mu^t \\ \mathbf{1}^t \end{pmatrix}, \quad \text{with}$$

$$\mu^t = 10^{-4}(1.693, 2.435, 0.337, 1.395, 0.467, 2.137, 2.601, 1.669, -0.879, 1.798).$$

Table 1 Different values of μ_p and the associated $\|\rho_{\hat{\mu}}\|_2$ and $K_2(\hat{B})$

$\mu_p = 10^{-4}$.	0.962	1.0965	1.231	1.3654	1.500	1.6345	1.7690
$\ \rho_{\hat{\mu}}\ _2 = 10^{-4}$.	3.2716	3.2590	3.2514	3.2489	3.2514	3.2590	3.2716
$K_2(\hat{B}) = 10^3$.	10.138	10.177	10.200	10.208	10.200	10.177	10.138

Despite the illustrative nature of this toy example, the values in μ are calculated from real financial data; specifically, they represent the average historical daily returns of the ten sectors of the S&P500.² It follows that $\beta = 1.3654 \cdot 10^{-4}$, and $\|\rho_{\mu}\|_2 = 3.2490 \cdot 10^{-4}$. Choosing $\delta = 3.2717 \cdot 10^{-4}$ yields $P_{\delta}(B) = 1$, since $\|\rho_{\mu}\|_2 < \delta \sqrt{\beta^2 \frac{n}{n-\delta^2} + 1} \approx 3.2717 \cdot 10^{-4}$. Furthermore, from Theorem 2,

$$R_0 = \sqrt{1 + \frac{1}{n - \delta^2}} \sqrt{\delta^2 \left(\frac{n}{n - \delta^2} \beta^2 + 1 \right) - \|\rho_{\mu}\|_2^2} \approx 4.035 \cdot 10^{-5}.$$

Therefore, we restrict μ_p to the interval $[\beta \frac{n}{n-\delta^2} - R_0, \beta \frac{n}{n-\delta^2} + R_0] = 10^{-4} \cdot [0.962, 1.769]$ to ensure that the matrix

$$\hat{B} = \begin{pmatrix} \mu^t & \mu_p \\ \mathbf{1}^t & 1 \end{pmatrix}$$

maintains a δ -numerical rank equal to 1, since, as highlighted in Table 1, $\|\rho_{\hat{\mu}}\|_2 \leq \delta = 3.2719 \cdot 10^{-4}$, and so $\|\rho_{\hat{\mu}}\|_2 \leq \delta \sqrt{\beta^2 \frac{n+1}{n+1-\delta^2} + 1}$. Furthermore, Table 1 shows that $K_2(\hat{B})$ is greater than $K_2(B) = 9.733 \cdot 10^3$, confirming that both matrices are ill-conditioned.

The bold entries in Table 1 correspond to the case $\mu_p = \beta$. In Fig. 1, we report a graphical representation of the δ -stable part of the frontier (in bold in the figure) according to the data of the example.

The principal idea behind this paper is that to mitigate allocation instability, the value of the parameter μ_p is selected such that $P_{\delta}(B) = P_{\delta}(\hat{B})$, establishing a rank condition whose association with lower target return sensitivity is examined empirically. As shown in Table 1, this choice has interesting implications for the condition numbers of B and $\hat{B}$. Specifically, as proved in the following theorem, if $\mu_p = \beta$, then $K_2(\hat{B}) \geq K_2(B)$. Since the condition number increases when the rows of a matrix are nearly linearly dependent, the subsequent theorem demonstrates that choosing $\mu_p = \beta$ remains admissible even when the rows of B are nearly collinear and/or badly scaled.

Theorem 3 Let $\mu_p = \beta$. It holds that $K_2(B) \leq K_2(\hat{B})$ if and only if $0 < \frac{\|\rho_{\mu}\|_2^2}{n} \leq \sqrt{\frac{n+1}{n}}(\beta^2 + 1)$. Equality holds if and only if $\frac{\|\rho_{\mu}\|_2^2}{n} = \sqrt{\frac{n+1}{n}}(\beta^2 + 1)$.

² A detailed description of the databases used in the real-data application is provided in the next section.

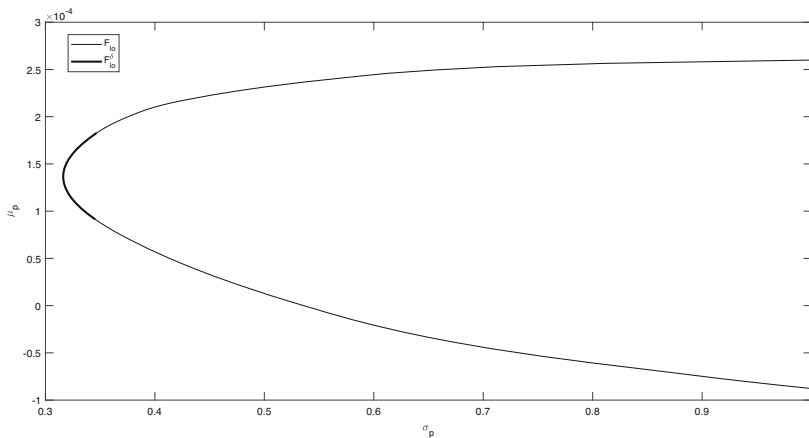


Fig. 1 The frontier calculated for Problem 1 and, in bold, the δ stable part

Furthermore, if $\lim_{n \rightarrow +\infty} \frac{\|\rho_\mu\|_2^2}{n(n+1)} = 0$, then $\lim_{n \rightarrow +\infty} K_2(\widehat{B}) = K_2(B)$.

Proof For $\mu_p = \beta$, we have $\widehat{\beta} = \beta$ and $\|\rho_\mu\|_2^2 = \|\rho_{\widehat{\mu}}\|_2^2$ (see the proof of Theorem 2). It follows that the matrices BB^t and $\widehat{B}\widehat{B}^t$ can be expressed as

$$BB^t = \begin{pmatrix} (\beta \mathbf{1} + \rho_\mu)^t (\beta \mathbf{1} + \rho_\mu) & \mathbf{1}^t \mu \\ \mathbf{1}^t \mu & n \end{pmatrix} = n \begin{pmatrix} \beta^2 + \frac{\|\rho_\mu\|_2^2}{n} & \beta \\ \beta & 1 \end{pmatrix},$$

$$\widehat{B}\widehat{B}^t = \begin{pmatrix} (\beta \widehat{\mathbf{1}} + [\rho_\mu; \beta])^t (\beta \widehat{\mathbf{1}} + [\rho_\mu; \beta]) & \widehat{\mathbf{1}}^t \widehat{\mu} \\ \widehat{\mathbf{1}}^t \widehat{\mu} & n+1 \end{pmatrix} = (n+1) \begin{pmatrix} \beta^2 + \frac{\|\rho_\mu\|_2^2}{n+1} & \beta \\ \beta & 1 \end{pmatrix}.$$

Let us define the quantities $q = \frac{\|\rho_\mu\|_2^2}{n(n+1)}$ and $a = \beta^2 + \frac{\|\rho_\mu\|_2^2}{n}$. Since

$$\beta^2 + \frac{\|\rho_\mu\|_2^2}{n+1} = \beta^2 + \frac{\|\rho_\mu\|_2^2}{n} - \frac{\|\rho_\mu\|_2^2}{n(n+1)} = a - q,$$

the matrices above can be rewritten as

$$BB^t = n \begin{pmatrix} a & \beta \\ \beta & 1 \end{pmatrix} \quad \text{and} \quad \widehat{B}\widehat{B}^t = (n+1) \begin{pmatrix} a-q & \beta \\ \beta & 1 \end{pmatrix}.$$

The eigenvalues of BB^t and $\widehat{B}\widehat{B}^t$, denoted by $\lambda_{1,2}$ and $\widehat{\lambda}_{1,2}$ respectively, are given by

$$\lambda_{1,2} = n \frac{(a+1) \pm \sqrt{(a-1)^2 + 4\beta^2}}{2}$$

$$\widehat{\lambda}_{1,2} = (n+1) \frac{(a+1-q) \pm \sqrt{(a-1)^2 + 4\beta^2 + q(q-2a+2)}}{2}.$$

If we consider the condition $K_2(B) \leq K_2(\widehat{B})$, which is equivalent to

$$\frac{\sqrt{\lambda_1}}{\sqrt{\lambda_2}} \leq \frac{\sqrt{\widehat{\lambda}_1}}{\sqrt{\widehat{\lambda}_2}},$$

we must solve the inequality $\lambda_1 \widehat{\lambda}_2 \leq \widehat{\lambda}_1 \lambda_2$. This yields

$$[(a+1) + \sqrt{(a-1)^2 + 4\beta^2}][(a+1-q) - \sqrt{(a-1)^2 + 4\beta^2 + q(q-2a+2)}] \leq$$

$$\leq [(a+1) - \sqrt{(a-1)^2 + 4\beta^2}][(a+1-q) + \sqrt{(a-1)^2 + 4\beta^2 + q(q-2a+2)}],$$

which, since $q \neq 0$, simplifies to $q(\beta^2 - a) + (\beta^2 - a)^2 - (\beta^2 + 1)^2 \leq 0$. Substituting the previously defined values of a and q , we obtain

$$\frac{\|\rho\|_2^4}{n(n+1)} \leq (\beta^2 + 1)^2 \quad \text{or, equivalently,} \quad \frac{\|\rho\|_2^2}{n} \leq \sqrt{\frac{n+1}{n}}(\beta^2 + 1).$$

This final condition is satisfied by hypothesis, thereby establishing that $K_2(B) \leq K_2(\widehat{B})$.

Finally, if $\lim_{n \rightarrow +\infty} \frac{\|\rho\|_2^2}{n(n+1)} = 0$, meaning that $\lim_{n \rightarrow +\infty} q = 0$, it follows that $\lim_{n \rightarrow +\infty} \widehat{\lambda}_1 = \lambda_1$ and $\lim_{n \rightarrow +\infty} \widehat{\lambda}_2 = \lambda_2$. We therefore conclude that $\lim_{n \rightarrow +\infty} K_2(\widehat{B}) = K_2(B)$. $\square$

Choosing $\delta^2 < n/2$, if $\|\rho_\mu\|_2 \leq \delta \sqrt{\beta^2 \frac{n}{n-\delta^2} + 1}$, the hypotheses of Theorem 2 are satisfied, and thus $P_\delta(B) = P_\delta(\widehat{B}) = 1$ provided that μ_p is sufficiently close to β . Furthermore, since $\frac{\|\rho_\mu\|_2^2}{n} \leq \delta^2(\beta^2/(n-\delta^2) + 1/n) \leq \frac{n}{2}(2\beta^2/n + 1/n) \leq \beta^2 + 1$, the hypotheses of Theorem 3 are also satisfied, implying that $K_2(B) - K_2(\widehat{B}) \leq 0$. Since the matrix condition number is a continuous function of its entries, there exists a neighborhood around β such that $K_2(B) \leq K_2(\widehat{B})$ for every μ_p within this neighborhood. Therefore, when μ_p is sufficiently close to β , not only does $P_\delta(B) = P_\delta(\widehat{B}) = 1$ hold, but $\widehat{B}$ is also more ill-conditioned than B , a result clearly demonstrated in Table 1.

Remark 2 Using the Maclaurin expansion, we have

$$\frac{1}{n - \delta^2} = \frac{1}{n} + \sum_{j=1}^{\infty} \frac{\delta^{2j}}{n^{j+1}} = \frac{1}{n} + O\left(\frac{\delta^2}{n^2}\right).$$

Hence, provided that

$$\frac{\delta^2}{n} \ll 1,$$

the higher-order terms in the expansion can be neglected, and we may use the approximation

$$\frac{1}{n - \delta^2} \simeq \frac{1}{n}.$$

In the remainder of the paper, this approximation is adopted throughout both the theoretical analysis and the illustrative examples. Under this assumption, the condition

$$P_\delta(B) = 1$$

is equivalent, up to the aforementioned approximation, to

$$\|\rho_\mu\|_2 \leq \delta\sqrt{\beta^2 + 1}.$$

Moreover, the corresponding quantity R_0 is approximated by

$$R = \sqrt{1 + \frac{1}{n}} \sqrt{\delta^2(\beta^2 + 1) - \|\rho_\mu\|_2^2}.$$

Consequently, within the regime $\delta^2/n \ll 1$, we characterize the condition

$$P_\delta(B) = P_\delta(\widehat{B}) = 1$$

by the simultaneous requirements

$$\|\rho_\mu\|_2^2 \leq \delta^2(\beta^2 + 1) \quad \text{and} \quad \mu_p \in [\beta - R, \beta + R].$$

4 Application

The empirical analysis is divided into three subsections: the first subsection discusses the selection of δ with respect to the number of assets n ; the second describes the real financial data used for the application and defines the criteria for evaluating the results; and the third presents the numerical results on real data along with the corresponding discussion.

4.1 The Choice of δ

In the proposed approach, we deal with ill-conditioned matrices B . Conversely, if $\|\rho_\mu\|_2$ is not small enough, no significant numerical issues are expected, and consequently, no restrictions on the value of μ_p are necessary. In the cases under consideration, $\|\rho_\mu\|_2$ is a small value and δ is a user-defined parameter. If $\delta\sqrt{\beta^2 + 1} \geq \|\rho_\mu\|_2$ and $\delta^2/n \ll 1$, we can apply the results established in the previous sections to compute the optimal interval for μ_p . Thus, the choice of δ clearly plays a central role in this methodology. Through the following example, we demonstrate that the assumption $\delta^2/n \ll 1$, where n denotes the number of assets in the portfolio, ensures the avoidance of scenarios where the minimum singular value of B remains relatively large despite B being ill-conditioned.

Example 3 Without loss of generality, we set an even integer n and define a matrix $B \in \mathbb{R}^{2 \times n}$ as

$$B = \begin{pmatrix} \epsilon & 2\epsilon & \epsilon & 2\epsilon & \dots & \epsilon & 2\epsilon \\ 1 & 1 & 1 & 1 & \dots & 1 & 1 \end{pmatrix} \quad \text{with } 0 < \epsilon \ll 1.$$

By construction, it holds that $\beta = 1.5\epsilon$ and $\rho_\mu = 0.5(-\epsilon, \epsilon, -\epsilon, \epsilon, \dots, -\epsilon, \epsilon)^t$, from which $\|\rho_\mu\|_2 = 0.5\epsilon\sqrt{n}$. Furthermore, using Eq. 1 and neglecting the term $9\epsilon^4$, we have

$$\lambda_{1,2}^2 = \frac{n}{4} \left(5\epsilon^2 + 2 \pm \sqrt{(2 + 4\epsilon^2)^2 + 9\epsilon^4} \right) \approx \frac{n}{4} \left(5\epsilon^2 + 2 \pm (2 + 4\epsilon^2) \right)$$

and so $\lambda_1 \approx \frac{\sqrt{n}}{2} \sqrt{9\epsilon^2 + 4}$ and $\lambda_2 \approx \frac{\sqrt{n}}{2} \epsilon$.

It follows that $K_2(B) \approx \frac{\sqrt{9\epsilon^2 + 4}}{\epsilon}$, meaning that B is ill-conditioned independently of n .

Conversely, for large values of n , λ_2 can be relatively large even if B remains ill-conditioned.

We choose δ such that $\|\rho_\mu\|_2^2 \leq \delta^2(1 + \beta^2)$, which yields $\delta^2 \geq \frac{n\epsilon^2}{4 + 9\epsilon^2}$. Choosing $\delta = \epsilon \frac{\sqrt{1.01n}}{\sqrt{4 + 9\epsilon^2}}$, we have $\delta^2/n = 1.01 \frac{\epsilon^2}{4 + 9\epsilon^2} \ll 1$, and we obtain

$$R = \sqrt{\frac{n+1}{n}} \sqrt{\delta^2(1 + \beta^2) - \|\rho_\mu\|_2^2} = \sqrt{\frac{n+1}{n}} \sqrt{0.01 \frac{n\epsilon^2}{4}} = \epsilon \frac{\sqrt{n+1}}{20}.$$

If $n = 500$ and $\epsilon = 0.001$, then $K_2(B) \approx 2000$ and $\lambda_2(B) \approx 0.01118$. Choosing $\delta = \epsilon \frac{\sqrt{1.01n}}{\sqrt{4 + 9\epsilon^2}} \approx 0.011236$, we have $\lambda_2(B) < \delta$, which implies that $P_\delta(B) = 1$. Furthermore, we obtain $R \approx 0.001119$. Consequently, if $\mu_p \in [\beta - R, \beta + R] = [0.00038, 0.0026]$, then $P_\delta(\hat{B}) = 1$. For example, if $\mu_p = 0.002$, then $\lambda_2(\hat{B}) = 0.0111914 < \delta \approx 0.011236$.

Example 3 highlights the following fact: when n is large, it is possible to obtain a non-negligible value for λ_2 even when the matrix B is severely ill-conditioned. This outcome strongly underscores the necessity of linking the choice of δ to the value of n when applying the proposed strategy in practical applications.

In the following, we propose a suitable choice of δ for the scenario under consideration, namely when $P_\delta(B) = 1$ and $\delta \ll \sqrt{n}$. The value of this parameter is derived from both theoretical considerations and mathematical arguments. Let B and $\hat{B} = [B \mid [\mu_p, 1]^t]$. By the interlacing property of singular values [39], the following relationship holds:

$$\lambda_1(\hat{B}) \geq \lambda_1(B) \geq \lambda_2(\hat{B}) \geq \lambda_2(B).$$

Assume that the hypotheses of Theorem 2 are satisfied, meaning that $\lambda_2(B)$ and $\lambda_2(\hat{B})$ are both strictly less than δ ; consequently, we have

$$\lambda_2(B) \leq \lambda_2(\hat{B}) \leq \delta.$$

Since the smallest singular value of a matrix coincides with its distance in the 2-norm to the nearest rank-deficient matrix, choosing δ sufficiently close to $\lambda_2(B)$ ensures that the respective distances of B and $\hat{B}$ from rank deficiency remain tightly close.

Since we are concerned with the case in which $\lambda_2(B) \leq \delta \ll \sqrt{n}$, and recalling from Proposition 1 that

$$\lambda_2(B) \leq \delta \Leftrightarrow \frac{\|\rho_\mu\|_2}{\sqrt{\beta^2 + 1}} \leq \delta,$$

a reasonable choice of δ is obtained by requiring it to be slightly larger than, yet sufficiently close to, $\lambda_2(B)$. Accordingly, one may set

$$\delta = 1.005 \frac{\|\rho_\mu\|_2}{\sqrt{\beta^2 + 1}}. \quad (2)$$

Observe that the choice $\delta = \frac{\|\rho_\mu\|_2}{\sqrt{\beta^2 + 1}}$ is not acceptable, as it forces $R = 0$, thereby making $\mu_p = \beta$ the only feasible value. On the other hand, selecting a value of δ that is too large distorts the concept of δ -numerical rank, which is precisely intended to indicate the presence of a rank-deficient matrix in the immediate neighborhood of B .

4.2 Data Description

The empirical application in the next section is performed on five different datasets. All datasets contain daily asset returns from 1/3/2000 to 9/17/2020. For simplicity, assets without a complete time series over the chosen period have been removed. Consequently, the number n of effective constituents in the financial indexes under analysis is smaller than the nominal one. The datasets are the following:

- S&P500 sectors ($n = 10$) are composed of the ten sector portfolios of the S&P500 index, obtained using the Global Industry Classification Standard (GICS): Energy, Materials, Industrials, Consumer Discretionary, Consumer Staples, Health Care, Financials, Information Technology, Communication Services, and Utilities.
- DAX ($n = 23$) is the stock index of the major German blue-chip companies trading on the Frankfurt Stock Exchange.
- ESX ($n = 38$), EURO STOXX 50, is a stock index for the Eurozone composed of European blue-chip companies considered leaders in their respective sectors.
- FTSE ($n = 69$) is the Financial Times Stock Exchange 100 Index, composed of the companies listed on the London Stock Exchange with the highest market capitalization.
- Nikkei ($n = 181$) is the Nikkei 225, a stock market index for the Tokyo Stock Exchange composed of large, publicly owned companies in Japan.

The datasets allow us to test our proposal for different values of the portfolio size n .

In the next section, the results are discussed on the basis of the Turnover (TO). Given a small positive number ϵ , $TO_\epsilon(\mu_p)$ is defined as

$$TO_\epsilon(\mu_p) = \sum_{i=1}^n |x_i^*(\mu_p) - x_i^*(\mu_p + \epsilon)|. \quad (3)$$

Both the mathematical and economic interpretations of the TO are immediate. From a strict mathematical point of view, the TO measures the numerical sensitivity of the optimal solution to small variations ϵ in the parameter μ_p . From an economic point of view, the TO provides a measure of the transaction costs required to adjust the investment portfolio with respect to a small variation ϵ in μ_p . Moreover, to fully capture the economic consequences of numerical instability, we incorporate transaction costs of 10 basis points (as supported by [40]) proportional to the TO . This ensures a realistic performance comparison between active portfolio strategies in the out-of-sample rolling-window exercise.

4.3 Real Data Application

The empirical application is divided into two main experiments, each detailed in a dedicated subsection. The first experiment provides a descriptive analysis of the numerical stability of the optimal portfolios when the expected return is chosen within the suggested interval (see Section 4.4). The second experiment quantifies the economic impact of this numerical stability in terms of out-of-sample risk-adjusted returns (see Section 4.5). Both applications are implemented using the datasets described in the previous section.

4.4 Experiment One

In the first experiment, we demonstrate the stability of the optimal allocation when μ_p is selected within the proposed theoretical interval. This is achieved by constructing the efficient frontier for each dataset and evaluating the TO metric associated with contiguous points along the frontier. To further emphasize the role of the linear equality constraints as a primary source of numerical instability, the allocation model is implemented for each dataset using both the historical covariance matrix and, alternatively, the identity matrix. In the latter case, any potential numerical instability depends exclusively on the restrictions imposed by the optimization problem, given that the identity matrix represents the ideal benchmark in terms of conditioning. As noted in the introduction, a broad strand of literature has focused on alternative covariance estimators to reduce the numerical instability ascribed to sample estimation errors. While our focus centers on the instability caused by the linear equality constraints and the choice of μ_p , we emphasize that predicting how these two distinct sources of instability, the covariance matrix and the linear constraints, interact to determine the overall conditioning of the problem remains a challenging task. Interestingly, these

two sources may interact to mutually mitigate the overall instability or, in the worst-case scenarios, they may compound, resulting in amplified fragility. For this reason, we explicitly report the results obtained by replacing the covariance matrix with the identity matrix.

The analysis and discussion presented in this section are focused on two specific datasets: S&P500 Sectors and DAX. For both datasets, the stability analysis is presented both graphically and through a table displaying the exact optimal portfolio weights across different levels of μ_p , leveraging the fact that these asset universes include a limited number of constituents. For the remaining datasets, namely ESX, FTSE, and Nikkei, we limit our analysis to the graphical representations provided in Appendix 2 (Figs. 6, 7, 9, 10, 11), noting that the results are qualitatively identical to those discussed here. For each dataset, we solve the optimization problem formulated using the sample covariance matrix, as well as its counterpart obtained by replacing the covariance matrix with the identity matrix. As previously emphasized, this second framework allows us to isolate and focus on the role of the constraints in generating numerical instability. The graphical representations comprise three subfigures: the top panel illustrates the evolution of the optimal portfolio composition along the efficient frontier as a function of μ_p . The vertical green lines delineate the δ -stable region identified by our theoretical proposal. The middle panel shows the behavior of the

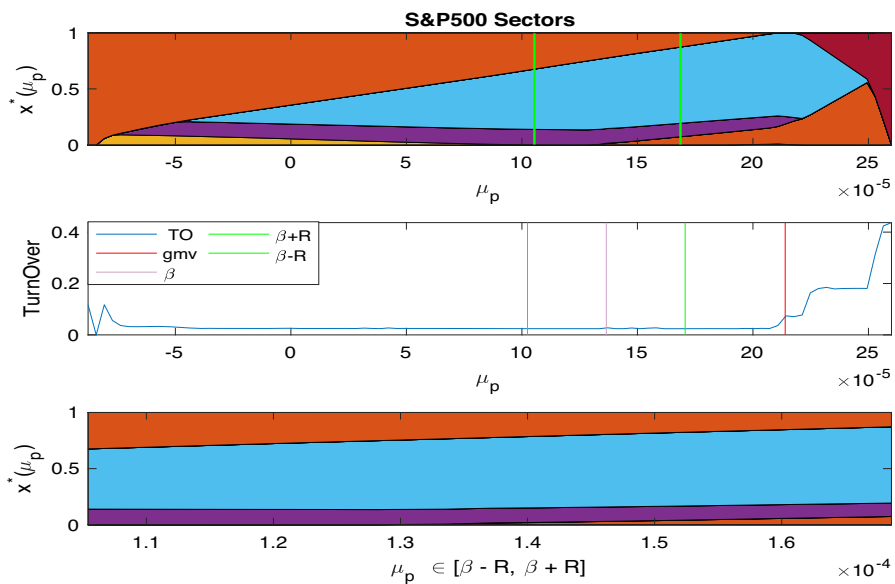


Fig. 2 Optimal portfolio structure and turnover across target returns. The dataset comprises S&P 500 sectors, and the optimization model relies on the sample covariance matrix. **Top panel:** Asset weight composition of the optimal portfolio across varying target returns (μ_p). **Middle panel:** Portfolio turnover (TO) plotted as a function of the target return μ_p . **Bottom panel:** Detailed asset weight composition restricted to the δ -stable region of the efficient frontier, corresponding to target returns $\mu_p \in [0.1054, 0.1686] \times 10^{-3}$. Across all panels, the vertical lines indicate key portfolio benchmarks: the global minimum variance (GMV) portfolio (red dashed line), the boundaries of the δ -stable interval (green lines), and the specific return level $\mu_p = \beta$ (pink line)

TO metric as μ_p varies. Again, the vertical green lines identify the interval where the portfolio composition along the frontier is most stable, while the vertical red and pink lines mark the global minimum variance portfolio and the portfolio with $\mu_p = \beta$, respectively. The location of the global minimum variance portfolio is of fundamental importance for the economic interpretation, given that the dominant portion of the frontier consists of all portfolios lying to the right of this point, which are characterized by higher expected returns. Finally, the bottom panel provides a specific view of the top plot, focusing specifically on the stability interval identified by the theoretical approach.

Figure 2 illustrates the stability analysis of the efficient frontier for the S&P500 Sectors dataset using the historical covariance matrix. From the top panel, it is evident that the optimal portfolio composition changes very rapidly for small variations in μ_p when the latter is close to its extreme feasible values. Under these circumstances, both the specific assets selected in the portfolio and their corresponding weights are highly sensitive to minor changes in μ_p . It is also worth noting that the optimal portfolio holds only a few asset classes at extreme values of μ_p , whereas a significantly larger number of assets are included for more central values of μ_p . The vertical green lines delineate the region of maximum stability identified by our theoretical approach. The middle panel shows the behavior of the TO as a function of μ_p . The most striking feature is that turnover is extremely high for extreme values of μ_p , while it remains remarkably low for choices of μ_p within the δ -stable interval. In this specific case, it is interesting to observe that the global minimum variance portfolio (indicated by the vertical red line in the middle panel) exactly overlaps with the portfolio obtained when $\mu_p = \beta$. This occurrence is coincidental, as will become evident from the subsequent examples. As a consequence of this coincidence, a portion of the efficient frontier exists that is simultaneously dominant in a mean-variance sense and numerically stable. The bottom panel, by focusing on the segment of the frontier identified by the theoretical interval, provides an immediate visual representation of the allocative stability in this region.

Table 2 Effective portfolio composition within the δ -stable region. The results are based on S&P 500 sectors return data, and the optimization problem is defined starting from the historical sample covariance matrix. This displays the effective asset weight allocation for optimal portfolios when the target return (μ_p) is restricted to the identified stable region. This highlights how asset weights stabilize and remain well-behaved inside the central interval compared to the unconstrained efficient frontier

$\mu_p = 10^{-3}$	0.1054	0.1124	0.1195	0.1265	0.1335	0.1405	0.1476	0.1546	0.1616	0.1686
x_1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0008	0.0000	0.0014	0.0000	0.0001
x_2	0.0000	0.0000	0.0001	0.0006	0.0055	0.0185	0.0320	0.0445	0.0598	0.0736
x_3	0.0014	0.0007	0.0003	0.0001	0.0001	0.0014	0.0000	0.0007	0.0000	0.0000
x_4	0.1379	0.1370	0.1361	0.1349	0.1333	0.1298	0.1279	0.1247	0.1223	0.1195
x_5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
x_6	0.5363	0.5605	0.5847	0.6083	0.6275	0.6372	0.6490	0.6588	0.6693	0.6795
x_7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0002	0.0000	0.0000
x_8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0000	0.0003	0.0000	0.0000
x_9	0.3244	0.3017	0.2789	0.2560	0.2337	0.2116	0.1911	0.1691	0.1486	0.1274
x_{10}	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0002	0.0000	0.0000

The nearly horizontal lines in this plot illustrate that the portfolio composition remains stable even when slightly modifying μ_p . Conversely, returning to the top panel, it can be seen that outside the stable region, the lines exhibit very steep slopes, corresponding to highly pronounced changes in the optimal portfolio composition relative to small variations in the expected return.

Table 2 reports the optimal portfolio composition as μ_p varies within the identified theoretical interval. This provides numerical, rather than just graphical, evidence of how gradually the optimal portfolio composition changes in response to small variations in μ_p .

Figure 3 illustrates the stability of the optimal portfolio composition for the S&P500 Sectors dataset when the covariance matrix is replaced with the identity matrix in the optimization problem. The most immediate observation compared to the previous case is the significant impact that the dependence structure provided by the covariance matrix exerts on the allocation. Specifically, when asset dependence and variances are omitted, a generally larger number of asset classes comprise the optimal portfolio. That being said, the observations regarding this figure are entirely analogous to those made previously. When μ_p takes on extreme values, the portfolio concentrates on a few assets, and small variations in μ_p heavily impact the composition. Conversely, when

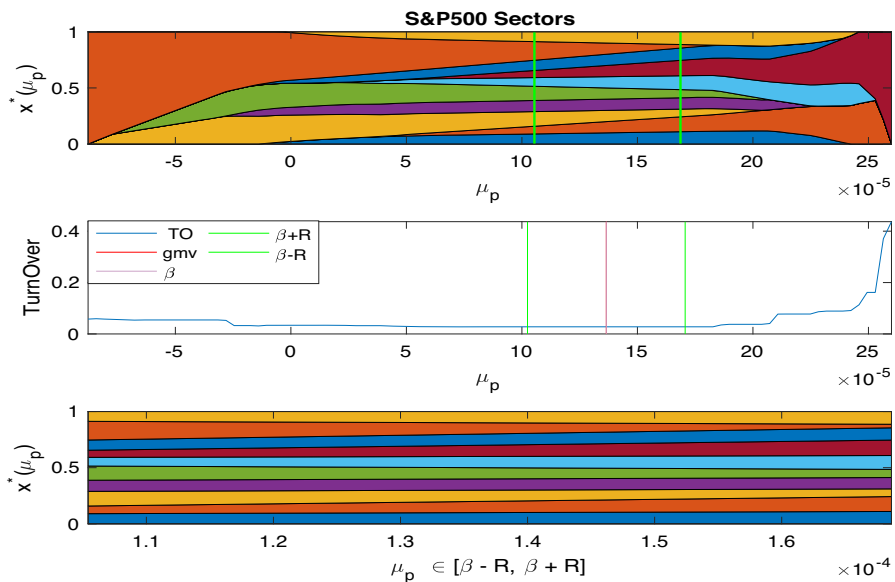


Fig. 3 Optimal portfolio structure and turnover across target returns. The dataset comprises S&P 500 sectors, and the optimization model relies on the identity matrix in place of the sample covariance matrix. **Top panel:** Asset weight composition of the optimal portfolio across varying target returns (μ_p). **Middle panel:** Portfolio turnover (TO) plotted as a function of the target return μ_p . **Bottom panel:** Detailed asset weight composition restricted to the δ -stable region of the efficient frontier, corresponding to target returns $\mu_p \in [0.1054, 0.1686] \times 10^{-3}$. Across all panels, the vertical lines indicate key portfolio benchmarks: the global minimum variance (GMV) portfolio (red dashed line), the boundaries of the δ -stable interval (green lines), and the specific return level $\mu_p = \beta$ (pink line)

Table 3 Effective portfolio composition within the δ -stable region. The dataset comprises S&P 500 sectors, and the optimization model relies on the identity matrix in place of the sample covariance matrix. This displays the effective asset weight allocation for optimal portfolios when the target return (μ_p) is restricted to the identified stable region. This highlights how asset weights stabilize and remain well-behaved inside the central interval compared to the unconstrained efficient frontier

$\mu_p = 10^{-3}$	0.1054	0.1124	0.1195	0.1265	0.1335	0.1405	0.1476	0.1546	0.1616	0.1686
x_1	0.0903	0.0925	0.0947	0.0969	0.0991	0.1012	0.1034	0.1056	0.1078	0.1100
x_2	0.0685	0.0756	0.0827	0.0898	0.0969	0.1041	0.1112	0.1183	0.1254	0.1325
x_3	0.1303	0.1235	0.1166	0.1098	0.1029	0.0961	0.0893	0.0824	0.0756	0.0687
x_4	0.0991	0.0993	0.0995	0.0997	0.0999	0.1001	0.1003	0.1005	0.1007	0.1009
x_5	0.1265	0.1205	0.1145	0.1086	0.1026	0.0966	0.0906	0.0846	0.0787	0.0727
x_6	0.0772	0.0824	0.0875	0.0926	0.0978	0.1029	0.1081	0.1132	0.1183	0.1235
x_7	0.0636	0.0718	0.0800	0.0882	0.0965	0.1047	0.1129	0.1211	0.1293	0.1376
x_8	0.0910	0.0931	0.0951	0.0971	0.0991	0.1011	0.1032	0.1052	0.1072	0.1092
x_9	0.1662	0.1512	0.1363	0.1214	0.1064	0.0915	0.0766	0.0616	0.0467	0.0318
x_{10}	0.0872	0.0901	0.0930	0.0959	0.0988	0.1016	0.1045	0.1074	0.1103	0.1132

μ_p is selected within the identified region, the composition appears more diversified and stable. This evidence is confirmed by Table 3, which, similarly to the previous case, numerically demonstrates how the optimal portfolio composition varies with μ_p .

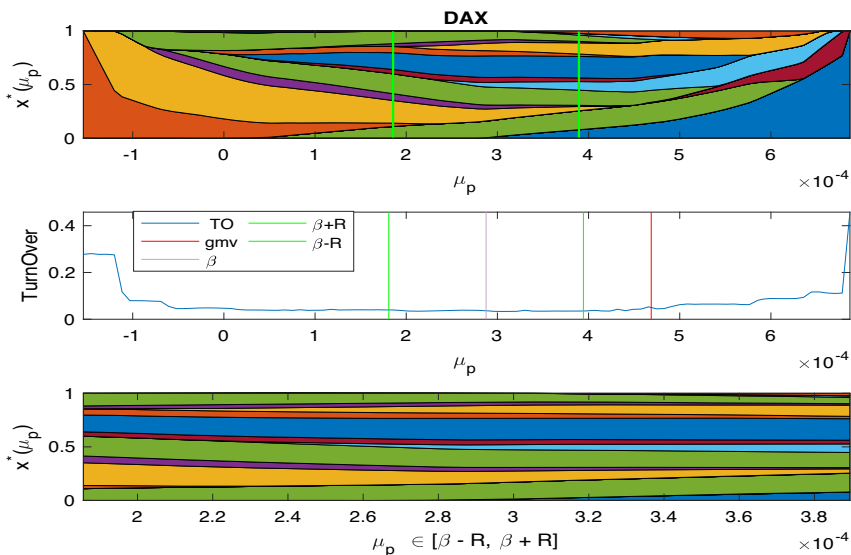


Fig. 4 Optimal portfolio structure and turnover across target returns. The dataset comprises DAX index constituents, and the optimization model relies on the sample covariance matrix. **Top panel:** Asset weight composition of the optimal portfolio across varying target returns (μ_p). **Middle panel:** Portfolio turnover (TO) plotted as a function of the target return μ_p . **Bottom panel:** Detailed asset weight composition restricted to the δ -stable region of the efficient frontier, corresponding to target returns $\mu_p \in [0.1856, 0.3894] \times 10^{-3}$. Across all panels, the vertical lines indicate key portfolio benchmarks: the global minimum variance (GMV) portfolio (red dashed line), the boundaries of the δ -stable interval (green lines), and the specific return level $\mu_p = \beta$ (pink line)

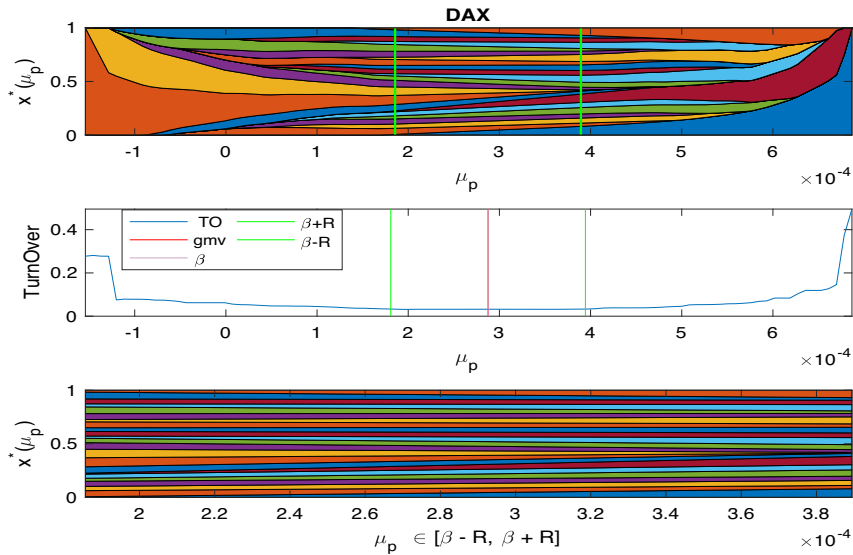


Fig. 5 Optimal portfolio structure and turnover across target returns. The dataset comprises the DAX index components, and the optimization model relies on the identity matrix in place of the sample covariance matrix. **Top panel:** Asset weight composition of the optimal portfolio across varying target returns (μ_p). **Middle panel:** Portfolio turnover (TO) plotted as a function of the target return μ_p . **Bottom panel:** Detailed asset weight composition restricted to the δ -stable region of the efficient frontier, corresponding to target returns $\mu_p \in [0.1856, 0.3894] \times 10^{-3}$. Across all panels, the vertical lines indicate key portfolio benchmarks: the global minimum variance (GMV) portfolio (red dashed line), the boundaries of the δ -stable interval (green lines), and the specific return level $\mu_p = \beta$ (pink line)

In a completely analogous manner to the results shown for the S&P500 Sectors dataset, Figs. 4 and 5, along with Tables 4 and 5, present the empirical findings for the DAX dataset using the historical covariance matrix and the identity matrix, respectively. In terms of numerical stability, it is difficult to identify any significantly different behavior in this example compared to the previous case. When utilizing the covariance matrix, the optimal portfolio concentrates on a subset of the available asset classes; for extreme values of μ_p , the portfolio composition is highly sensitive to small variations, whereas the allocation is clearly more stable within the region identified by the theoretical approach. Similarly, when the identity matrix replaces the covariance matrix, the optimal portfolio generally includes a larger number of assets, and once again, the theoretically identified stable region effectively corresponds to a more numerically stable allocation.

Remark 3 It is worth noting that, as anticipated, the stable interval for μ_p depends exclusively on the linear restrictions of the optimization problem. As an immediate consequence, this stable interval remains unchanged when the historical covariance matrix is replaced with the identity matrix.

The middle panel of Fig. 4 warrants a dedicated economic discussion. In this case, as theoretically anticipated in Remark 1, the numerically stable allocation region lies entirely to the left of the global minimum variance portfolio. This implies that if

Table 4 Effective portfolio composition within the δ -stable region. The dataset comprises DAX index constituents, and the optimization model relies on the sample covariance matrix. This displays the effective asset weight allocation for optimal portfolios when the target return (μ_p) is restricted to the identified stable region. This highlights how asset weights stabilize and remain well-behaved inside the central interval compared to the unconstrained efficient frontier

$\mu_p = 10^{-3}$	0.1856	0.2111	0.2366	0.2620	0.2875	0.3130	0.3385	0.3640	0.3894
x_1	0.0000	0.0000	0.0000	0.0000	0.0035	0.0246	0.0424	0.0607	0.0783
x_2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
x_3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0003	0.0003
x_4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0002	0.0002
x_5	0.1098	0.1224	0.1331	0.1441	0.1543	0.1603	0.1650	0.1699	0.1745
x_6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0006	0.0008	0.0006
x_7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0005	0.0008
x_8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0000
x_9	0.0294	0.0120	0.0010	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000
x_{10}	0.2121	0.1887	0.1650	0.1386	0.1131	0.0945	0.0762	0.0581	0.0398
x_{11}	0.0633	0.0588	0.0532	0.0453	0.0372	0.0298	0.0231	0.0168	0.0113
x_{12}	0.1850	0.1867	0.1832	0.1755	0.1679	0.1613	0.1548	0.1483	0.1417
x_{13}	0.0000	0.0002	0.0085	0.0259	0.0423	0.0531	0.0619	0.0709	0.0796
x_{14}	0.0408	0.0449	0.0471	0.0473	0.0469	0.0447	0.0414	0.0382	0.0348
x_{15}	0.1565	0.1688	0.1795	0.1903	0.1995	0.2007	0.2007	0.2007	0.2006
x_{16}	0.0580	0.0587	0.0577	0.0537	0.0492	0.0433	0.0362	0.0294	0.0222
x_{17}	0.0006	0.0167	0.0353	0.0533	0.0700	0.0801	0.0882	0.0964	0.1043
x_{18}	0.0267	0.0326	0.0350	0.0334	0.0312	0.0277	0.0226	0.0178	0.0134
x_{19}	0.1159	0.1085	0.1010	0.0926	0.0845	0.0783	0.0723	0.0660	0.0593
x_{20}	0.0000	0.0000	0.0000	0.0000	0.0001	0.0005	0.0059	0.0102	0.0154
x_{21}	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001
x_{22}	0.0019	0.0009	0.0003	0.0000	0.0000	0.0000	0.0002	0.0001	0.0001
x_{23}	0.0000	0.0000	0.0000	0.0000	0.0001	0.0010	0.0078	0.0144	0.0225

numerical stability is prioritized, the investor will select a portfolio that is dominated in a mean-variance sense by alternative allocations. This aspect is of such critical importance that the entire out-of-sample experiment in Section 4.5 is dedicated to analyzing this specific scenario by incorporating transaction costs.

4.5 Experiment Two

This subsection is dedicated to investigating the economic consequences of numerical instability by incorporating transaction costs into the analysis. This second experiment is conducted within a standard rolling-window out-of-sample framework. For a given estimation window of length w , the covariance matrix and the vector of expected returns are computed using historical data to determine the optimal asset allocation. The window length w is set to be strictly larger than the number of constituents in each

Table 5 Effective portfolio composition within the δ -stable region. The dataset comprises the DAX index components, and the optimization model relies on the identity matrix in place of the sample covariance matrix. This displays the effective asset weight allocation for optimal portfolios when the target return (μ_p) is restricted to the identified stable region. This highlights how asset weights stabilize and remain well-behaved inside the central interval compared to the unconstrained efficient frontier

$\mu_p = 10^{-3}$	0.1856	0.2111	0.2366	0.2620	0.2875	0.3130	0.3385	0.3640	0.3894
x_1	0.0061	0.0154	0.0248	0.0341	0.0434	0.0528	0.0621	0.0715	0.0808
x_2	0.0564	0.0532	0.0499	0.0467	0.0435	0.0403	0.0370	0.0338	0.0306
x_3	0.0399	0.0408	0.0417	0.0426	0.0435	0.0444	0.0453	0.0462	0.0471
x_4	0.0480	0.0468	0.0457	0.0446	0.0435	0.0424	0.0412	0.0401	0.0390
x_5	0.0279	0.0318	0.0357	0.0396	0.0435	0.0474	0.0513	0.0552	0.0591
x_6	0.0387	0.0399	0.0411	0.0423	0.0435	0.0447	0.0459	0.0470	0.0482
x_7	0.0093	0.0178	0.0264	0.0349	0.0435	0.0520	0.0605	0.0691	0.0776
x_8	0.0589	0.0551	0.0512	0.0473	0.0435	0.0396	0.0358	0.0319	0.0281
x_9	0.0849	0.0745	0.0642	0.0538	0.0435	0.0332	0.0228	0.0125	0.0022
x_{10}	0.0791	0.0702	0.0613	0.0524	0.0435	0.0346	0.0257	0.0168	0.0079
x_{11}	0.0617	0.0571	0.0526	0.0480	0.0435	0.0390	0.0344	0.0299	0.0253
x_{12}	0.0408	0.0415	0.0421	0.0428	0.0435	0.0441	0.0448	0.0455	0.0461
x_{13}	0.0209	0.0265	0.0322	0.0378	0.0435	0.0491	0.0547	0.0604	0.0660
x_{14}	0.0427	0.0429	0.0431	0.0433	0.0435	0.0437	0.0439	0.0440	0.0442
x_{15}	0.0347	0.0369	0.0391	0.0413	0.0435	0.0457	0.0479	0.0500	0.0522
x_{16}	0.0551	0.0522	0.0493	0.0464	0.0435	0.0406	0.0377	0.0348	0.0319
x_{17}	0.0217	0.0271	0.0326	0.0380	0.0435	0.0489	0.0543	0.0598	0.0652
x_{18}	0.0551	0.0522	0.0493	0.0464	0.0435	0.0406	0.0377	0.0348	0.0319
x_{19}	0.0627	0.0579	0.0531	0.0483	0.0435	0.0387	0.0339	0.0291	0.0243
x_{20}	0.0288	0.0325	0.0361	0.0398	0.0435	0.0471	0.0508	0.0545	0.0582
x_{21}	0.0459	0.0453	0.0447	0.0441	0.0435	0.0429	0.0423	0.0417	0.0411
x_{22}	0.0630	0.0581	0.0532	0.0484	0.0435	0.0386	0.0338	0.0289	0.0240
x_{23}	0.0178	0.0242	0.0306	0.0370	0.0435	0.0499	0.0563	0.0627	0.0691

dataset to ensure a full-rank covariance matrix. In this setup, we omit testing the performance under the identity matrix substitution, as our primary objective is to evaluate the economic significance of the asset allocation strategy rather than its mathematical properties. The out-of-sample portfolio return is subsequently computed based on the realizations of the $(w + 1)$ -th period. The procedure is then iteratively repeated by shifting the estimation window one period forward. Transaction costs are incorporated as detailed in the previous section, specifically by multiplying the turnover (TO) metric by 10 basis points. In this rolling-window exercise, the TO corresponds to the $\mathcal{L}_1$ -norm distance between the optimal allocations of two consecutive periods. Following standard empirical practices, transaction costs, reflecting the rebalancing expenses required to realign the actual portfolio with the newly computed optimal weights, are deducted from the raw portfolio returns. This approach naturally accounts for market frictions, thereby penalizing highly active investment strategies. To demonstrate the

practical impact of a stable allocation in a realistic investment scenario, we compare the performance of two specific portfolios along the efficient frontier. The first portfolio is associated with $\mu_p = \beta$; as established analytically in this paper, this allocation exhibits the highest stability with respect to the choice of μ_p . The benchmark portfolio selected for comparison is chosen from the dominant portion of the frontier, setting μ_p to identify the portfolio located halfway between the global minimum variance portfolio and the portfolio with the theoretical maximum expected return. This choice ensures that the benchmark belongs to the dominant region of the efficient frontier, while simultaneously avoiding extreme allocations characterized by severe instability, as demonstrated in the first experiment. The two portfolios, denoted as $\mathbf{x}_\beta^*$ and $\mathbf{x}_b^*$, are compared in terms of their out-of-sample Sharpe ratios (SR). Recall that the Sharpe ratio is defined as the average return in excess of the risk-free rate divided by the standard deviation of returns, serving as the most widely used risk-adjusted performance metric in practice; see [41].

Table 6 displays the out-of-sample Sharpe ratios for the two portfolios across estimation window lengths ($w = 120, 180, 240, 480$), corresponding approximately to 6 months, 9 months, 1 year, and 2 years of daily data. Across all datasets and window sizes, the stable portfolio $\mathbf{x}_\beta^*$ consistently achieves higher Sharpe ratios than the benchmark portfolio $\mathbf{x}_b^*$. Looking across estimation windows, performance trends vary by dataset. For the S&P Sectors and DAX, Sharpe ratios generally improve as w increases, reaching their highest values at $w = 480$. Conversely, for ESX and FTSE, performance peaks at an intermediate window length ($w = 240$) before declining slightly at $w = 480$. These empirical patterns reflect fundamental trade-offs in portfolio construction. While a longer estimation window provides a larger sample size that stabilizes covariance estimates and smooths out portfolio turnover, thereby mitigating transaction costs, it simultaneously relies on older historical data that may suffer from reduced forecasting power in changing market environments. Conversely, shorter windows adapt more quickly to recent market dynamics but are inherently subject to higher estimation noise and increased turnover. Overall, these results suggest that the observed performance differentials stem from a combination of transaction costs, the

Table 6 Out-of-sample Sharpe ratios across asset universes and estimation horizons. This table reports the out-of-sample Sharpe ratios achieved by the optimal portfolios across five major international equity markets: S&P 500 Sectors, DAX, ESX, FTSE, and Nikkei. Performance is evaluated using rolling estimation windows (w) of varying lengths ($w \in \{120, 180, 240, 480\}$ days). All metrics are calculated out-of-sample to assess the stability and empirical performance of the proposed strategy under different market conditions and estimation horizons

	w = 120		w = 180		w = 240		w = 480	
	SR($\mathbf{x}_\beta^*$)	SR($\mathbf{x}_b^*$)	SR($\mathbf{x}_\beta^*$)	SR($\mathbf{x}_b^*$)	SR($\mathbf{x}_\beta^*$)	SR($\mathbf{x}_b^*$)	SR($\mathbf{x}_\beta^*$)	SR($\mathbf{x}_b^*$)
S&P Sectors	0.0069	−0.0007	0.009	-10^{-7}	0.0125	0.0058	0.016	0.011
DAX	0.0099	−0.0061	0.013	−0.0104	0.0206	0.005	0.0312	0.0242
ESX	0.014	−0.0218	0.0175	−0.0076	0.0192	−0.0056	0.0172	0.0084
FTSE	0.0318	−0.007	0.033	0.0027	0.0413	0.0091	0.0389	0.0125
Nikkei	—	—	—	—	0.0059	−0.005	0.0162	0.0043

turnover-reducing effects of longer rolling windows, and the inherent balance between parameter estimation stability and data timeliness.

5 Conclusions

In this paper, we investigate how the target expected return influences the numerical instability of optimal portfolio allocation. From a theoretical perspective, we establish a link between the numerical rank of the matrix defining the optimization problem's linear constraints and the choice of the expected return. Furthermore, we demonstrate that numerical stability is preserved when the target expected return is selected such that it avoids degrading the condition number of the linear constraints matrix. By introducing a parameter δ , calibrated as a function of the number of assets, we restrict μ_p to a suitable interval, thereby isolating what we define as the δ -stable region of the efficient frontier. The empirical analysis on financial datasets validates this theoretical framework, explicitly demonstrating how the choice of μ_p drives portfolio stability. We evaluate this allocation stability in a descriptive framework by computing the turnover between contiguous portfolios along the efficient frontier. Additionally, a rolling-window backtest reveals that these stable portfolios outperform traditional dominant portfolios once transaction costs are considered, reinforcing a broad body of empirical literature that emphasizes the practical importance of allocation stability. Finally, while traditional literature primarily identifies covariance matrix estimation and model uncertainty as the main drivers of numerical instability, our approach shifts the focus toward constraints-driven instability. Future research will focus on the examination of how these distinct sources of instability interact to determine the overall robustness of the optimal allocation problem, as well as establishing a formal, scale-invariant link between the numerical rank and KKT conditioning of the equality constraints and parametric optimizer sensitivity.

Appendix 1: Technical results

The appendix contains instrumental technical results that are removed from the main body of the paper to preserve the logical structure of the proposal.

Proposition 4 *Let $0 < \delta < \sqrt{n}$. The condition $\|\rho_{\hat{\mu}}\|_2^2 \leq \delta^2 \left(\frac{n+1}{n+1-\delta^2} \hat{\beta}^2 + 1 \right)$ is satisfied by $\mu_p \in \left[\frac{n}{n-\delta^2} \beta - R_0, \frac{n}{n-\delta^2} \beta + R_0 \right]$, where*

$$R_0 = \sqrt{1 + \frac{1}{n - \delta^2}} \sqrt{\delta^2 \left(\frac{n}{n - \delta^2} \beta^2 + 1 \right) - \|\rho_{\mu}\|_2^2}.$$

Proof The condition $\|\rho_{\mu}\|_2^2 \leq \delta^2 \left(\frac{n+1}{n+1-\delta^2} \widehat{\beta}^2 + 1 \right)$ can be rewritten as

$$\|\rho_{\mu}\|_2^2 + \frac{n}{n+1}(\beta - \mu_p)^2 \leq \delta^2 \left(1 + \frac{(n\beta + \mu_p)^2}{(n+1)(n+1-\delta^2)} \right)$$

which corresponds to

$$\mu_p^2(n - \delta^2) - 2n\beta\mu_p + (n+1-\delta^2)(\|\rho_{\mu}\|_2^2 - \delta^2) + n\beta^2(1 - \delta^2) \leq 0. \quad (4)$$

First, we solve the second-order equation associated with Inequality Eq. 4, whose discriminant Δ_{μ} is

$$\Delta_{\mu} = (n+1-\delta^2) \left(n\beta^2\delta^2 - (n-\delta^2) \left(\|\rho_{\mu}\|_2^2 - \delta^2 \right) \right)$$

and so

$$\mu_p^{(1,2)} = \frac{n\beta}{n-\delta^2} \pm \sqrt{\frac{n+1-\delta^2}{n-\delta^2} \sqrt{\frac{n\delta^2\beta^2}{n-\delta^2} - \|\rho_{\mu}\|_2^2 + \delta^2}} = \frac{n\beta}{n-\delta^2} \pm R_0.$$

Since Inequality Eq. 4 is satisfied for $\mu_p^{(1)} \leq \mu_p \leq \mu_p^{(2)}$, the proof is complete. $\square$

Appendix 2: Stability analysis for the remaining databases

In this appendix, we provide a graphical analysis of the stability of the δ -stable part of the efficient frontier for the ESX, FTSE, and Nikkei datasets. For all the datasets, we show what happens both when the optimization problem is solved using the historical covariance matrix and the identity matrix. Due to the large number of assets in these datasets, the stability analysis is presented only in a graphical way, avoiding to show the composition of the optimal portfolios. As already underlined in the main text, the results are qualitatively equivalent to the ones described in Subsection 4.4. When the covariance matrix is replaced with the identity matrix for the two largest datasets, FTSE and Nikkei, the number of assets included in the optimal portfolio becomes so large that the corresponding area plots lose visual clarity. For this reason, in these two scenarios, we omit the plot displaying the turnover (TO) as a function of μ_p , thereby allocating more space to the graphs illustrating the optimal portfolio composition. Under the identity matrix substitution, the asset allocation becomes independent of both the correlation and risk structures, solely accounting for expected returns. Consequently, each available asset enters the optimal allocation, resulting in a portfolio that remains remarkably close to the equally weighted benchmark.

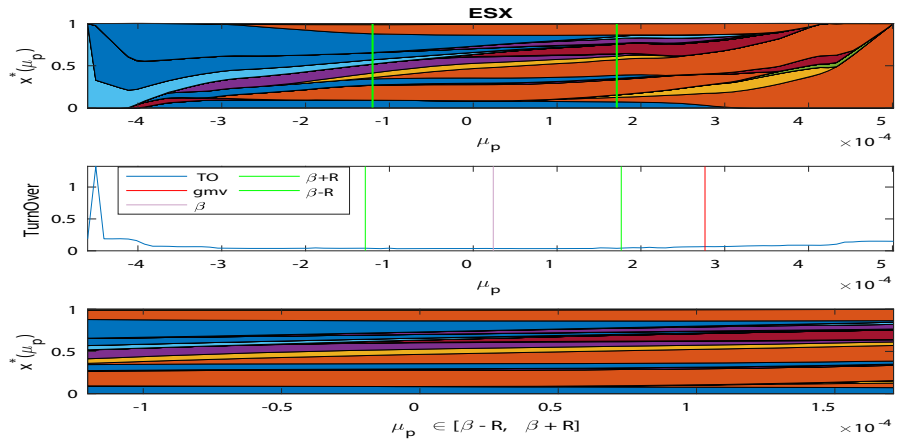


Fig. 6 Optimal portfolio structure and turnover across target returns. The dataset comprises ESX index constituents, and the optimization model relies on the sample covariance matrix. **Top panel:** Asset weight composition of the optimal portfolio across varying target returns (μ_p). **Middle panel:** Portfolio turnover (TO) plotted as a function of the target return μ_p . **Bottom panel:** Detailed asset weight composition restricted to the δ -stable region of the efficient frontier, corresponding to target returns $\mu_p \in [-0.1201, 0.1712] \times 10^{-3}$. Across all panels, the vertical lines indicate key portfolio benchmarks: the global minimum variance (GMV) portfolio (red dashed line), the boundaries of the δ -stable interval (green lines), and the specific return level $\mu_p = \beta$ (pink line)

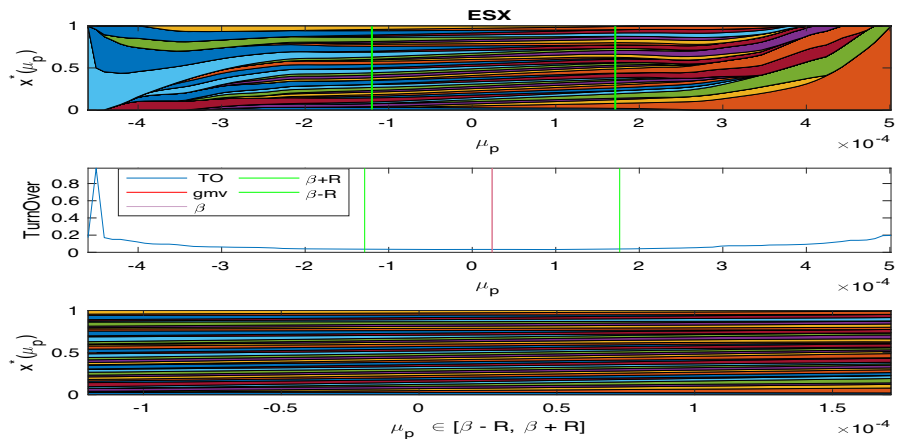


Fig. 7 Optimal portfolio structure and turnover across target returns. The dataset comprises the ESX index components, and the optimization model relies on the identity matrix in place of the sample covariance matrix. **Top panel:** Asset weight composition of the optimal portfolio across varying target returns (μ_p). **Middle panel:** Portfolio turnover (TO) plotted as a function of the target return μ_p . **Bottom panel:** Detailed asset weight composition restricted to the δ -stable region of the efficient frontier, corresponding to target returns $\mu_p \in [-0.1201, 0.1712] \times 10^{-3}$. Across all panels, the vertical lines indicate key portfolio benchmarks: the global minimum variance (GMV) portfolio (red dashed line), the boundaries of the δ -stable interval (green lines), and the specific return level $\mu_p = \beta$ (pink line)

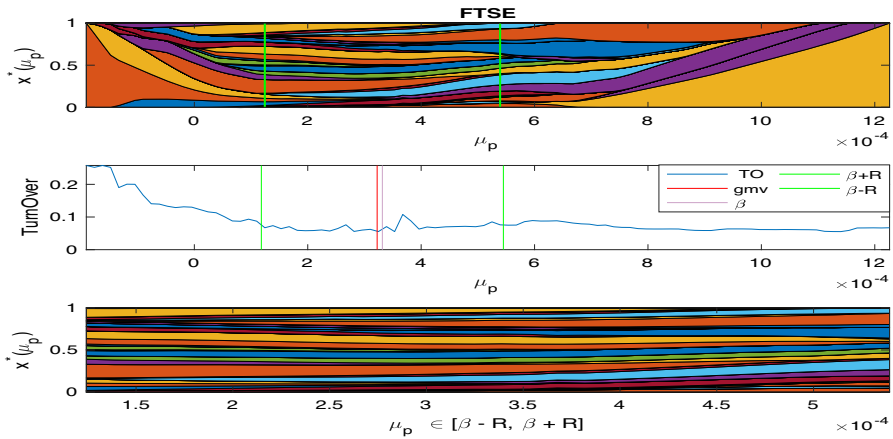


Fig. 8 Optimal portfolio structure and turnover across target returns. The dataset comprises FTSE index constituents, and the optimization model relies on the sample covariance matrix. **Top panel:** Asset weight composition of the optimal portfolio across varying target returns (μ_p). **Middle panel:** Portfolio turnover (TO) plotted as a function of the target return μ_p . **Bottom panel:** Detailed asset weight composition restricted to the δ -stable region of the efficient frontier, corresponding to target returns $\mu_p \in [0.1243, 0.5393] \times 10^{-3}$. Across all panels, the vertical lines indicate key portfolio benchmarks: the global minimum variance (GMV) portfolio (red dashed line), the boundaries of the δ -stable interval (green lines), and the specific return level $\mu_p = \beta$ (pink line).labelFTSEcov

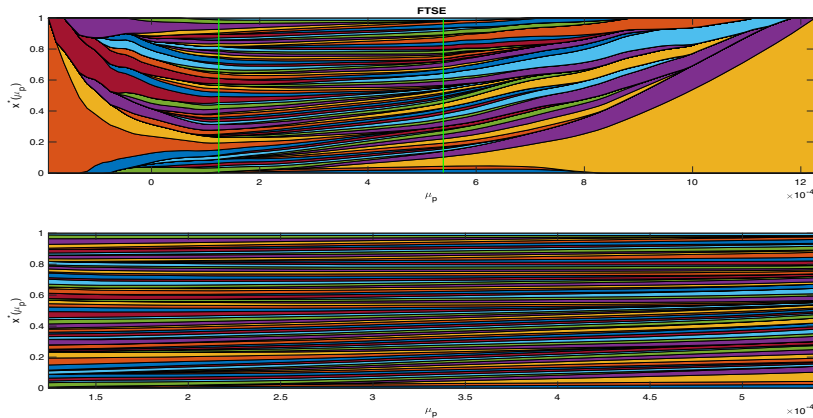


Fig. 9 Optimal portfolio structure and turnover across target returns. The dataset comprises the FTSE index components, and the optimization model relies on the identity matrix in place of the sample covariance matrix. **Top panel:** Asset weight composition of the optimal portfolio across varying target returns (μ_p). **Bottom panel:** Detailed asset weight composition restricted to the δ -stable region of the efficient frontier, corresponding to target returns $\mu_p \in [0.1243, 0.5393] \times 10^{-3}$. Across all panels, the vertical lines indicate key portfolio benchmarks: the global minimum variance (GMV) portfolio (red dashed line), the boundaries of the δ -stable interval (green lines), and the specific return level $\mu_p = \beta$ (pink line)

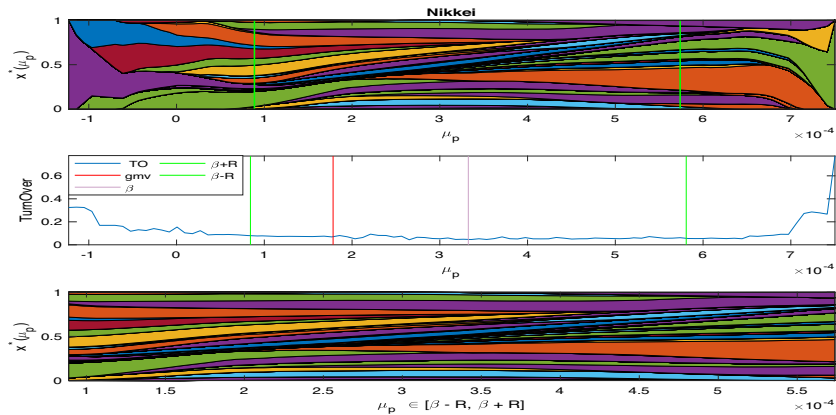


Fig. 10 Optimal portfolio structure and turnover across target returns. The dataset comprises Nikkei index constituents, and the optimization model relies on the sample covariance matrix. **Top panel:** Asset weight composition of the optimal portfolio across varying target returns (μ_p). **Middle panel:** Portfolio turnover (TO) plotted as a function of the target return μ_p . **Bottom panel:** Detailed asset weight composition restricted to the δ -stable region of the efficient frontier, corresponding to target returns $\mu_p \in [0.0888, 0.5739] \times 10^{-3}$. Across all panels, the vertical lines indicate key portfolio benchmarks: the global minimum variance (GMV) portfolio (red dashed line), the boundaries of the δ -stable interval (green lines), and the specific return level $\mu_p = \beta$ (pink line)

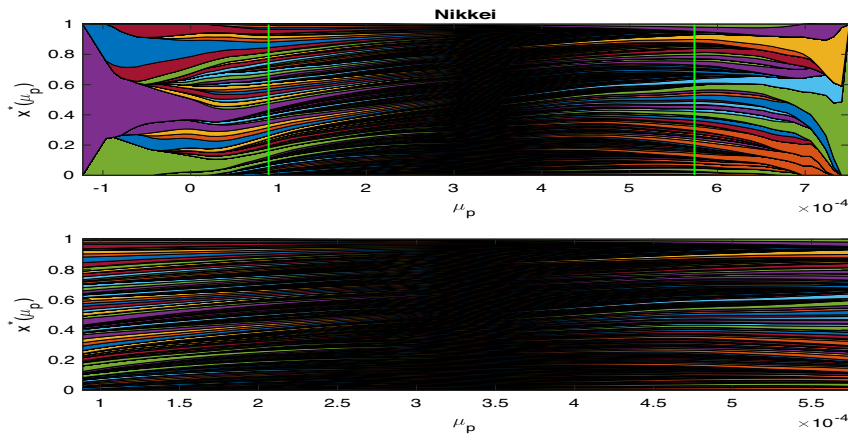


Fig. 11 Optimal portfolio structure and turnover across target returns. The dataset comprises the Nikkei index components, and the optimization model relies on the identity matrix in place of the sample covariance matrix. **Top panel:** Asset weight composition of the optimal portfolio across varying target returns (μ_p). **Bottom panel:** Detailed asset weight composition restricted to the δ -stable region of the efficient frontier, corresponding to target returns $\mu_p \in [0.0888, 0.5739] \times 10^{-3}$. Across all panels, the vertical lines indicate key portfolio benchmarks: the global minimum variance (GMV) portfolio (red dashed line), the boundaries of the δ -stable interval (green lines), and the specific return level $\mu_p = \beta$ (pink line)

Author Contribution The authors contributed equally to this paper.

Funding Open access funding provided by Università degli Studi di Milano - Bicocca within the CRUI-CARE Agreement. The authors did not receive any funding for the present research.

Data Availability The data are available upon request.

Declarations

Competing Interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Markowitz H (1952) Portfolio selection. *J Financ* 7(1):77–91
2. DeMiguel V, Garlappi L, Uppal R (2009) Optimal versus naive diversification: how inefficient is the 1/n portfolio strategy? *Rev Finan Stud* 22(5):1915–1953
3. Pflug GC, Pichler A, Wozabal D (2012) The 1/n investment strategy is optimal under high model ambiguity. *J Bank Financ* 36(2):410–417
4. Kan R, Zhou G (2007) Optimal portfolio selection with parameter uncertainty. *J Finan Quant Anal* 42(3):621–656
5. Chopra VK, Ziemba WT et al (1993) The effect of errors in means, variances, and covariances on optimal portfolio choice. *J Portf Manag* 19(2):6–11
6. Gelmini M, Uberti P (2024) The equally weighted portfolio still remains a challenging benchmark. *Int Econ* 179:100525
7. Malladi R, Fabozzi FJ (2017) Equal-weighted strategy: why it outperforms value-weighted strategies? Theory and evidence. *J Asset Manag* 18(3):188–208
8. Zakamulin V (2017) Superiority of optimized portfolios to naive diversification: fact or fiction? *Financ Res Lett* 22:122–128
9. Yuan M, Zhou G (2024) Why naive diversification is not so naive, and how to beat it? *J Finan Quant Anal* 59(8):3601–3632
10. DeMiguel V, Garlappi L, Nogales FJ, Uppal R (2009) A generalized approach to portfolio optimization: improving performance by constraining portfolio norms. *Manage Sci* 55(5):798–812
11. Allen D, Lizieri C, Satchell S (2019) In defense of portfolio optimization: what if we can forecast? *Financ Anal J* 75(3):20–38
12. Kritzman M, Page S, Turkington D (2010) In defense of optimization: the fallacy of 1/n. *Financ Anal J* 66(2):31–39
13. Frost PA, Savarino JE (1986) An empirical bayes approach to efficient portfolio selection. *J Finan Quant Anal* 21(3):293–305
14. Ledoit O, Wolf M (2004) A well-conditioned estimator for large-dimensional covariance matrices. *J Multivar Anal* 88(2):365–411
15. Ledoit O, Wolf M (2003) Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *J Empir Financ* 10(5):603–621
16. Ledoit O, Wolf M (2017) Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. *Rev Finan Stud* 30(12):4349–4388

17. Ledoit O, Wolf M (2022) The power of (non-) linear shrinking: a review and guide to covariance matrix estimation. *J Financ Econ* 20(1):187–218
18. Brodie J, Daubechies I, De Mol C, Giannone D, Loris I (2009) Sparse and stable Markowitz portfolios. *Proc Natl Acad Sci* 106(30):12267–12272
19. Kim JH, Kim WC, Fabozzi FJ (2014) Recent developments in robust portfolios with a worst-case approach. *J Optim Theory Appl* 161:103–121
20. Xidonas P, Steuer R, Hassapis C (2020) Robust portfolio optimization: a categorized bibliographic review. *Ann Oper Res* 292(1):533–552
21. Ghahtarani A, Saif A, Ghasemi A (2022) Robust portfolio selection problems: a comprehensive review. *Oper Res Int J* 22(4):3203–3264
22. Pflug GC, Pohl M (2018) A review on ambiguity in stochastic portfolio optimization. *Set-Valued and Variational Analysis* 26:733–757
23. Fabozzi FJ, Huang D, Zhou G (2010) Robust portfolios: contributions from operations research and finance. *Ann Oper Res* 176:191–220
24. Maillard S, Roncalli T, Teïletche J (2010) The properties of equally weighted risk contribution portfolios. *J Portf Manag* 36(4):60
25. Roncalli T (2013) Introduction to risk parity and budgeting. CRC Press, Boca Raton
26. Choueifaty Y, Coignard Y (2008) Toward maximum diversification. *J Portfolio Manag* 35(1):40–51
27. Meucci A (2009) Managing diversification. *Risk* 22(5):74–79
28. Yu J-R, Lee W-Y, Chiou W-JP (2014) Diversified portfolios with different entropy measures. *Appl Math Comput* 241:47–63
29. Petukhina A, Klochov Y, Härdle WK, Zhivotovskiy N (2024) Robustifying Markowitz. *J Econ* 239(2):105680
30. Pagnoncelli B, Canto F, Cifuentes A (2023) The effect of regularization in portfolio selection problems. *TOP An Official Journal of the Spanish Society of Statistics and Operations Research* 31(3):521–542
31. Kan R, Wang X, Zhou G (2022) Optimal portfolio choice with estimation risk: No risk-free asset case. *Manage Sci* 68(3):2047–2068
32. Avramov D, Cheng S, Metzker L (2023) Machine learning vs. economic restrictions: Evidence from stock return predictability. *Manage Sci* 69(5):2587–2619
33. Jagannathan R, Ma T (2003) Risk reduction in large portfolios: Why imposing the wrong constraints helps. *J Financ* 58(4):1651–1683
34. Michaud RO (1989) The markowitz optimization enigma: Is ‘optimized’ optimal? *Financ Anal J* 45(1):31–42
35. Fassino C, Torrente M-L, Uberti P (2022) A singular value decomposition based approach to handle ill-conditioning in optimization problems with applications to portfolio theory. *Chaos, Solitons & Fractals* 165:112746
36. Golub GH, Van Loan CF (2013) *Matrix Computations*, 4th edn. Johns Hopkins University Press, Baltimore
37. Golub GH, Hoffman A, Stewart GW (1987) A generalization of the Eckart -Young-Mirsky matrix approximation theorem. *Linear Algebra and its Applications* 88/89:317–327 (1987 88/89:317–327
38. Bini D, Pan VY (2012) *Polynomial and Matrix Computations: Fundamental Algorithms*. Springer, New York
39. Thompson RC (1972) Principal submatrices ix: interlacing inequalities for singular values of submatrices. *Linear Algebra Appl* 6:1–12
40. Ledoit O, Wolf M (2025) Markowitz portfolios under transaction costs. *Quarterly Review of Economics and Finance* 100:101962
41. Sharpe WF (1966) Mutual fund performance. *J Bus* 39(1):119–138