UNIVERSITY OF LIVERPOOL Management School

Working Paper in Economics

202502

A Regularization Approach to Optimizing Large Portfolios Under Asymmetries in Returns and Risk Attitudes

Mohamed Doukali[1]
N'Golo Kone[2]
Abderrahim Taamouti[3]

[1] University of Liverpool Management School, m.doukali@liverpool.ac.uk
[2] Queen's University, ngk1@queensu.ca
[3] University of Liverpool Management School, abderrahim.taamouti@liverpool.ac.uk

https://www.liverpool.ac.uk/economics/

A Regularization Approach to Optimizing Large Portfolios Under Asymmetries in Returns and Risk Attitudes

Mohamed Doukali[†1], N'Golo Kone [‡2], and Abderrahim Taamouti[§3]

[1, 3]Department of Economics, University of Liverpool Management School.
[2]Department of Economics, Queen's University

Version: 3 April 2025

[†]Corresponding author. Address: Department of Economics, University of Liverpool Management School. Address: Chatham St, Liver- pool L69 7ZH. Email: m.doukali@liverpool.ac.uk

[‡]Department of Economics, Queen's University, 94 University Avenue Kingston, Ontario, Canada K7L 3N6. Email: ngk1@queensu.ca. Phone: 001 514 652-1005

[§]Address: Department of Economics, University of Liverpool Management School. Address: Chatham St, Liver- pool L69 7ZH. Email: abderrahim.taamouti@liverpool.ac.uk

A Regularization Approach to Optimizing Large Portfolios Under Asymmetries in Returns and Risk Attitudes

ABSTRACT

This paper introduces a methodology for selecting large portfolios in the presence of asymmetries in asset returns and risk attitudes. Within this framework, the optimal portfolio depends on inverting the covariance matrix of returns. However, traditional estimators of this matrix become nearly singular when the number of assets is significantly larger than the sample size. This results in a selected portfolio that deviates substantially from the optimal one. To address this challenge, we propose four regularization techniques aimed at stabilizing the inverse of the covariance matrix: Ridge, Spectral Cut-Off, Landweber-Fridman, and Lasso. These regularization techniques involve a tuning parameter that requires careful selection. To tackle this, we introduce a data-driven approach for choosing the optimal tuning parameter. Through extensive simulations and an empirical study, we demonstrate that by accounting for asymmetries and stabilizing the inverse of the covariance matrix, we substantially enhance the performance of the optimal portfolio compared to several benchmark portfolios.

Keywords: Portfolio selection, Generalized Disappointment Aversion, Asymmetric returns, Large portfolios, Regularization techniques, Monte Carlo simulations.

1 Introduction

1.1 Motivation and contributions

In the years subsequent to the publication of Markowitz's seminal work in 1952, his impact on modern portfolio theory has endured. Practitioners and numerous research papers still embrace his approach for portfolio selection. However, Markowitz (1952) portfolio theory has faced significant criticism, primarily due to the assumption of symmetric return distributions, which fails to capture the inherent asymmetry often observed in real financial markets [see Ang and Chen (2002), Hong et al. (2007), among others]. This oversight can lead to suboptimal portfolio allocations, especially in the presence of extreme events. Furthermore, as highlighted by Dahlquist et al. (2017), there is evidence that investors exhibit disparate attitudes towards risk by assigning greater significance to losses than gains when assessing the risk of their portfolios. Recently, various methods have emerged for choosing optimal portfolios under asymmetry in returns and/or risk attitudes; see Dahlquist et al. (2017), Tédongap and Tinang (2022), and references therein. Similar to Markowitz (1952), the inversion of the covariance matrix is a pivotal step in implementing these methods. However, with a large number of assets, the covariance matrix can become ill-conditioned. This poses a practical issue as it impedes the calculation of optimal portfolio weights, diminishing the applicability of these methods in scenarios with a diverse array of assets. This paper simultaneously addresses the above challenges by proposing an innovative approach to selecting large portfolios in the presence of asymmetries in assets returns and risk attitudes.

Dahlquist et al. (2017) provided a novel solution to the problem of portfolio selection under asymmetries in both assets returns and risk attitudes. They modelled the former asymmetry based on a normal-exponential model for returns, while the latter was modelled using generalized disappointment aversion (GDA) risk preferences, as introduced by Gul (1991) and Routledge and Zin (2010). GDA provides a flexible framework for modelling investor behavior, offering various advantages. In the context of portfolio choice, GDA risk preferences offer a sophisticated structure through which investors can tailor their responses to the financial market uncertainty, assigning distinct weights to downside losses and upside gains. Traditional utility functions [e.g. quadratic utility used in mean-variance portfolio] often assume symmetric risk aversion, whereas GDA departs from this by acknowledging that individuals may exhibit asymmetric attitudes toward risk.

Under the above asymmetries, Dahlquist et al. (2017) derived three-fund separation strategy empowering investors to allocate wealth among a risk-free asset, a standard mean-variance efficient fund, and an additional fund that accommodates return asymmetries. However, as mentioned previously, Dahlquist et al. (2017)'s optimal portfolio solution involves inverting the covariance matrix of asset returns, which is effective only in the presence of a small number of assets. In their empirical application, Dahlquist et al. (2017) considered only two risky assets.

In the dynamic landscape of current financial markets, the incorporation of large portfolios, encompassing a diverse array of assets, becomes imperative for a comprehensive understanding of risk and return dynamics. However, the wide range of securities introduces challenges, especially in estimating the covariance matrix required for optimal portfolio selection. As the number of assets under consideration grows, the commonly used estimator, the sample covariance matrix, becomes ill-conditioned or nearly singular/non-invertible, amplifying estimation errors and affecting the portfolio's performance.

To address the ill-conditioning of the estimator of the covariance matrix of asset returns, this paper proposes various regularization (stabilization) techniques that help to improve the performance of Dahlquist et al. (2017)'s optimal portfolio and extend its applicability to situations involving a substantial number of assets. Inverting a covariance matrix can be viewed as solving an inverse problem. To overcome this issue, we apply three widely used regularization techniques: (i) the Ridge, which involves adding a diagonal matrix to the covariance matrix; (ii) the Spectral cut-off (or principal component) method, which entails discarding eigenvectors associated with the smallest eigenvalues; and (iii) the Landweber-Fridman iterative method. Additionally, we consider a form of Lasso applied to the inverse covariance matrix; see Friedman et al. (2008).

These regularization schemes incorporate a tuning parameter that requires careful selection. For instance, in the principal components approach, the regularization parameter is the number of principal components. The choice of this parameter is crucial, and we propose a data-driven selection method based on generalized cross-validation (GCV); see Li (1986, 1987). It is worth noting that our criterion differs from that used in traditional factor models, where the number of factors is selected using Bai and Ng (2002)'s criterion, as we do not impose any factor structure on the data.

Furthermore, the weights assigned by the investor to the mean-variance efficient fund and the fund accommodating return asymmetries—depend not only on the preference parameters, such as the investor's risk aversion, degree of disappointment aversion, and the percentage of the investor's certainty equivalent, but also on endogenous coefficients. These coefficients include the optimal asset allocation itself. As a result, the optimal allocation and the endogenous coefficients must be solved simultaneously, as emphasized by Dahlquist et al. (2017). In this paper, we introduce a data-driven method for selecting these crucial parameters.

To assess the performance of our regularized optimal portfolio under asymmetry in returns and GDA risk preferences, we adopt the methodology inspired by Zakamouline and Koekebakker (2009) to derive a closed-form solution for the Generalized Sharpe Ratio (GSR). This ratio, which considers higher moments of the distribution of returns, serves as a more robust performance measure compared to the commonly used Sharpe ratio. The conventional Sharpe ratio, grounded in mean-variance theory, is valid only for normally distributed returns or quadratic preferences. In situations with non-normal return distributions, relying on the conventional Sharpe ratio can lead to potentially misleading conclusions; see Hodges (1998), Zakamouline and Koekebakker (2009), Bernardo and Ledoit (2000*a*), among others.

We conducted a series of Monte Carlo simulations to evaluate the performance of our regularized optimal portfolios in comparison to other benchmark portfolios. Specifically, we compared with Dahlquist et al.'s (2017) optimal portfolio without regularization, standard and regularized mean-variance portfolios that neglect asymmetries in returns and risk attitudes, and the equally weighted portfolio (referred to as the naive portfolio). By assessing portfolios' generalized Sharpe ratio and expected loss in utility, the simulation results demonstrate that the regularized optimal portfolios proposed in this paper significantly outperform the benchmark portfolios, particularly Dahlquist et al.'s (2017) optimal portfolio when the number of assets is large. Furthermore, we observe that considering asymmetries leads to a noteworthy improvement in the optimal portfolio's expected loss in utility.

An empirical application is also considered to illustrate the practical usefulness of considering both regularization and asymmetries in portfolio choices. To do so, we apply our methods to several sets of portfolios from Kenneth French's website: the monthly 30-industry portfolios (FF30), the monthly 48-industry portfolios (FF48), and the monthly 100 portfolios formed on size and book-to-market (FF100). The empirical results show that by taking asymmetries into account and stabilizing the inverse of the covariance matrix, we considerably improve the performance of the optimal portfolio in terms of maximizing the generalized Sharpe ratio and minimizing the expected loss in utility.

1.2 Related literature

The literature on portfolio selection under asymmetric returns has attracted considerable attention for several years now. Researchers have explored various approaches to tackle the challenges posed by the observed asymmetry in financial market returns. Jondeau and Rockinger (2006) and Martellini and Ziemann (2010) investigate optimal portfolio allocation by considering higher moments of asset returns beyond the mean and variance. They incorporate skewness and kurtosis into the portfolio optimization process, recognizing the limitations of traditional mean-variance analysis. Guidolin and Timmermann (2008) explore the impact of regime-switching dynamics, skewness, and kurtosis preferences on international asset allocation. They consider investors with preferences beyond mean and variance, incorporating regime-switching models to account for changing market conditions. Mencía and Sentana (2009) focus on multivariate location–scale mixtures of normals and their application to mean–variance–skewness portfolio allocation. The authors explore portfolio optimization considering not only the mean and variance but also skewness, using a mixture model framework to capture the distributional characteristics of asset returns.

The portfolio selection approach considered in this paper is closely related to Dahlquist et al. (2017), who studied the impact of asymmetries in asset returns and risk attitudes on portfolio choice. They modelled asymmetric preferences using generalized disappointment aversion and asymmetric return distributions using a normal-exponential model. They provided insights into the role of skewness in shaping investors' decisions and risk preferences. However, Dahlquist et al.'s (2017) optimal portfolio solution works only for a small number of assets as it involves inverting an estimator of the covariance matrix of asset returns that can become nearly singular when the number of assets is sufficiently large. To extend Dahlquist et al. (2017) to portfolios with a large number of assets, we address the ill-conditioning of the estimator of the covariance matrix. We propose various regularization (stabilization) techniques, including Ridge, the Spectral cut-off (principal components) method, the Landweber-Fridman iterative method, and the Lasso applied to the inverse covariance matrix. We also introduce data-driven approaches for selecting tuning parameters in the regularization schemes and the weights assigned by the investor to the standard mean-variance efficient fund and the fund accommodating return asymmetries.

In the presence of a large number of assets, Ledoit and Wolf (2003) and Ledoit and Wolf (2004) aimed to enhance portfolio performance by introducing the linear shrinkage estimator for the covariance matrix. More recently, Ledoit and Wolf (2017) proposed a nonlinear shrinkage estimator and its factor-model-adjusted version. Along similar lines, Abadir et al. (2014) proposed a new way to estimate the covariance matrix of asset returns based on a consistent estimation of eigenvalues. Carrasco and Noumon (2011) and Carrasco and Doukali (2017) explored various regularization techniques to stabilize the inverse of the sample covariance matrix. However, it's noteworthy that these papers exclusively focus on mean-variance portfolios, overlooking higher moments in return distributions and thereby neglecting return asymmetries. In contrast, our work complements existing literature on large portfolios by specifically addressing asymmetries in asset returns within the context of investor portfolio choice.

Another approach to enhancing portfolio performance, as proposed in the context of the global minimum volatility (GMV) portfolio, involves imposing constraints on portfolio weights; see Jagannathan and Ma (2003), DeMiguel, Garlappi, Nogales and Uppal (2009), Brodie et al. (2009), Fastrich et al. (2015), and Fan et al. (2012). Furthermore, Ao et al. (2019) introduced a new approach for estimating the mean-variance portfolio, utilizing an unconstrained regression representation of the optimization problem combined with high-dimensional sparse regression methods. However, these works often assume that asset return distributions are either i.i.d or approximated by the multivariate normal distribution. Our contribution involves addressing high-dimensional estimation issues in portfolio selection under non-normal and asymmetric asset return distributions.

The rest of the paper is organized as follows. Section 2 presents the theoretical framework. Section 3 describes the four regularization techniques of the inverse of the covariance matrix. Section 4 derives the optimal portfolio and the optimal selection of the tuning parameter. Section 5 presents the generalized sharpe Ratio. Section 6 analyzes the performance of our methods through a Monte Carlo experiment. Section 7 illustrates the practical usefulness of our approach. Section 8 concludes.

2 Framework

This section is devoted to outlining the theoretical framework for the selection of portfolios in the context of asymmetries in asset returns and risk attitudes. To achieve this objective, we start by introducing the following notations. We assume the presence of N risky assets in the economy, characterized by a random returns vector denoted as r_{t+1}, along with a risk-free asset exhibiting a known rate r^f. The rate r^f is calibrated to correspond to the mean of the one-month Treasury-Bill (T-B) rate observed in the data.

Let $R_{t+1} = r_{t+1} - r^f$ denotes the vector of excess returns on the set of risky assets, and let 1_N represent the N-dimensional vector of ones. We define $\omega = (\omega_1, \ldots, \omega_N)'$ as the vector of portfolio weights, indicating the proportion of capital to be invested in the risky assets, while the remaining $1 - \omega' 1_N$ is allocated to the risk-free asset. Note that the short-selling is permissible in the financial market, implying that some of the weights can take on negative values. Using the above notations, we define the portfolio log return, say $r_{\omega,t+1}$, as follows:

$$r_{\omega,t+1} = \omega^{'} r_{t+1} + r_f \left(1 - \omega^{'} 1_N\right). \tag{2.1}$$

. Dahlquist et al. (2017) have recently explored the portfolio choices of an investor with generalized disappointment-aversion preferences, dealing with log returns characterized by a normal-exponential model. This innovative approach to portfolio selection accommodates asymmetries in both returns and risk attitudes, enabling the assignment of distinct weights to downside losses and upside gains. Formally, following Routledge and Zin (2010), Dahlquist et al. (2017) consider a portfolio optimization problem, which involves selecting the weights allocated to the N risky assets to maximize the utility of the certainty equivalent of terminal wealth W. This utility function is marked by the investor's level of risk aversion, their degree of disappointment aversion, and the percentage of the certainty equivalent below which outcomes are deemed disappointing.

To derive their optimal portfolio, Dahlquist et al. (2017) also assume that log returns on N risky assets are described by the model

$$r_t = \mu - \sigma \circ \delta + (\sigma \circ \delta)\varepsilon_{0,t} + \left(\sigma \circ \sqrt{1_N - \delta \circ \delta}\right) \circ \varepsilon_t, \tag{2.2}$$

where μ, σ, and δ are N-dimensional vectors of parameters that describe the mean, standard deviation, and asymmetry of the process of returns of the N risky assets, respectively. 1_N is a vector of ones, and $\circ$ denotes the element-wise product of vectors. The N-dimensional vector ε_t represents asset-specific shocks and follows a multivariate normal distribution with standard normal marginal densities and correlation matrix Ψ. The scalar $\varepsilon_{0,t}$ represents a common shock affecting all assets, following an exponential distribution with a rate parameter equal to one. The asymmetry in asset returns is ascribed to this shared source of risk. The exponential distribution is well-suited to model the concurrent happening of extreme events, such as substantial and infrequent losses. To illustrate how the vector δ characterizes the non-normality of returns, Dahlquist et al. (2017) derive the mean, variance, skewness, and excess kurtosis of each asset return $r_{i,t}$,

$$E(r_{i,t}) = \mu_i, \quad Var(r_{i,t}) = \sigma_i^2, \quad Skew(r_{i,t}) = 2\delta_i^3, \quad Xkurt(r_{i,t}) = 6\delta_i^4 \tag{2.3}$$

and the correlation and coskewness of the returns of asset i and asset j

$$Corr(r_{i,t}, r_{j,t}) = \Psi_{ij}\sqrt{1-\delta_i^2}\sqrt{1-\delta_i^2} + \delta_i\delta_j \tag{2.4}$$

$$Coskew(r_{i,t}, r_{j,t}) = \frac{E[(r_{i,t} - E(r_{i,t}))^2(r_{j,t} - E(r_{j,t}))]}{Var(r_{i,t})\sqrt{V}ar(r_{j,t})} = 2\delta_i^2\delta_j, \tag{2.5}$$

which reveals that δ plays a crucial role in characterizing the non-normality of returns, contributing to nonzero skewness, coskewness, and excess kurtosis. The aforementioned moments and cross-moments will serve as the basis for estimating the parameters of the returns process using the Generalized Method of Moments (GMM).

Using the second-order Taylor approximation a la Campbell and Viceira (2002) of the portfolio log return, Dahlquist et al. (2017) have shown that the optimal portfolio weights that maximize the utility of the certainty equivalent of investor's terminal wealth are given by:

$$\omega = \frac{1}{\tilde{\gamma}} \left(\omega^{MV} + \tilde{\chi} \omega^{AV} \right), \tag{2.6}$$

with

$$\omega^{MV} = \Sigma^{-1} \left(\mu - r_f 1_N + \frac{1}{2} \sigma^2 \right), \quad \omega^{AV} = \Sigma^{-1} (\sigma \circ \delta), \tag{2.7}$$

where σ^2 is the diagonal element of the variance-covariance matrix Σ. Note that the coefficients $\tilde{\gamma}$ and $\tilde{\chi}$ depend not only on the preference parameters, but also on the optimal asset allocation, ω. That is, the coefficients $\tilde{\gamma}$ and $\tilde{\chi}$ are endogenous to the model. For more details about the analytical expression of these parameters, the reader is refereed to Dahlquist et al. (2017).

In addition to the risk-free asset, the optimal solution given in (2.6) has two components. The first one, ω^{MV}, is called the "mean-variance" fund because it is the solution to the mean-variance optimal portfolio problem, and the second part, ω^{AV}, is the "asymmetry-variance" fund because its composition depends on the asymmetry vector δ, and the variance-covariance matrix of the risky asset returns.

One notable limitation of Dahlquist et al. (2017) approach lies in its implicit assumption that the number of assets N is small compared with the sample size T. While their analysis provides valuable insights into portfolio optimization with generalized disappointment-aversion preferences, the model's practical applicability may be limited when dealing with a larger set of assets. This implicit assumption can impact the robustness of the findings in scenarios where the portfolio comprises a more extensive array of risky assets. In contrast to their paper, and recognizing the inherent inclination of practitioners to explore patterns among as many assets as possible, we contribute to this literature by considering cases where the number of assets N is not restricted and may be smaller or larger than the sample size T.

The issue lies in the fact that the optimal portfolio, as proposed by Dahlquist et al. (2017), depends on the inverse of the covariance matrix of asset return distribution, a matrix that remains unknown and requires estimation. Conventionally, this estimation involves substituting the unknown quantities with their empirical counterparts. Specifically, this entails replacing the expected return μ with the sample average $\hat{\mu} = \frac{1}{T} \sum_{t=1}^{T} r_t$ and the covariance matrix Σ with the sample covariance $\hat{\Sigma} = (R - 1_T \hat{\mu})'(R -$

$1_T\hat{\mu})/T = \bar{R}'\bar{R}$, where R represents the $T \times N$ matrix with the tth row given by r_t'. However, relying on the sample covariance for forming the optimal portfolio may be problematic due to its potential near singularity. Consequently, inverting such a matrix might amplify estimation errors and significantly deteriorate the performance of the selected portfolio. To address this, in the subsequent sections, we introduce four regularization techniques aimed at stabilizing the inverse of the covariance matrix.

3 Regularization techniques

The stability of the optimal portfolio weights in (2.6) relies on the properties of the matrix $\hat{\Sigma} = \bar{R}'\bar{R}$. Two challenges may arise. The first one is associated with potential high correlation among assets; i.e., the population covariance matrix Σ is nearly singular. The second one is due to the situation where the number of assets under consideration is too large relative to the sample size, i.e., the sample covariance is (nearly) singular even though the population covariance is not. In such instances, $\hat{\Sigma}$ typically exhibits some singular values close to zero, rendering the problem ill-posed and complicating the optimization of the portfolio. These challenges are precisely captured by the condition number, representing the ratio of the maximal to minimal eigenvalue of $\hat{\Sigma}$. A large condition number results in an unstable inverse $(\bar{R}'\bar{R})^{-1}$. Formally, if we let v be an arbitrary $N \times 1$ vector, the solution z to the equation

$$v = \left(\frac{\bar{R}'\bar{R}}{T}\right) z \tag{3.1}$$

is unstable in the sense that it is not continuous in v. If v is replaced by a noisy observation $v + \Delta v$, the solution may be very far from the true value z. Hence, it is crucial to stabilize the inverse to ensure the stability of optimal portfolio weights.

The inverse problem literature, typically addressing infinite-dimensional problems, has put forth several regularization techniques to stabilize the solution to (3.1). For a comprehensive understanding of inverse problems, readers are referred to the papers Kress et al. (1989) and Kanwal (2013). In this section, we focus on four widely employed regularization techniques: Ridge, spectral cut-off, Landweber Fridman, and Lasso.

Before we delve into each regularization technique considered in this paper, let us first introduce some notations. At first, observe that the $T \times N$ matrix R can be perceived as an operator from $\mathbb{R}^N$ (equipped with the inner product $< v, z >= v'z$) into $\mathbb{R}^T$ (equipped with the inner product $< \phi, \rho >= \phi'\rho/T$). Denote the adjoint of $\bar{R}$ as $\bar{R}'/T$. Let $(\hat{\lambda}_j, \hat{\phi}_j, \hat{v}_j)$, for $j = 1, 2, ..., N$, be the singular system of R, i.e., $R\hat{\phi}_j = \hat{\lambda}_j \hat{v}_j$ and $\bar{R}\hat{v}_j/T = \hat{\lambda}_j \bar{\phi}_j$. Additionally, $(\hat{\lambda}_j^2, \hat{\phi}_j)$ represent the eigenvalues and orthonormal eigenvectors of $\bar{R}'\bar{R}/T$, while $(\hat{\lambda}_j^2, \hat{v}_j)$ represent the nonzero eigenvalues and orthonormal eigenvectors of $\bar{R}\bar{R}'/T$. We will now elaborate on the four regularization schemes under consideration in this paper:

Ridge (R) regularization

In the context of the inverse of the variance-covariance matrix, Ridge regularization is a technique that is used to enhance stability and address multicollinearity (very high correlation between assets' returns) when estimating the inverse of the variance-covariance matrix by introducing a regularization term to the optimization objective function. This term, proportional to the square of the matrix's elements, helps to prevent extreme values and ensures a well-conditioned and more reliable estimate. Formally, Ridge regularization consists in adding a diagonal matrix to $(\bar{R}'\bar{R}/T)^\tau$ or equivalently,

$$(\bar{R}'\bar{R}/T)^\tau z = \sum_{j=1}^{N} \frac{\hat{\lambda}_j}{\hat{\lambda}_j^2 + \tau} (z'\hat{\phi}_j)\hat{\phi}_j.$$

Thus, by applying Ridge regularization to the inverse of the variance-covariance matrix, we aim to strike a balance between fitting the data accurately and maintaining numerical stability in scenarios where multicollinearity or large number of assets might pose challenges to traditional estimation methods. This regularization has a Bayesian interpretation. For more details, the reader is referred to De Mol et al. (2008).

Landweber-Fridman (LF) regularization

Landweber-Fridman (LF) regularization also addresses numerical stability and multicollinearity concerns in the estimation of the inverse of the variance-covariance matrix by introducing a regularization term to the optimization process. However, unlike Ridge regularization, LF regularization technique incorporates an iterative algorithm, often resembling the Landweber iteration method. This iterative process helps stabilize the estimation of the inverse matrix, particularly in situations with highly correlated assets. Formally, let c be a constant such that $0 < c < 1/||\bar{R}||^2$ and $||\bar{R}||$ denotes the largest eigenvalue of $\bar{R}$. The LF regularized inverse can be computed as:

$$(\bar{R}'\bar{R}/T)^{\tau}z = \sum_{j=1}^{N} \frac{1}{\hat{\lambda}_j}(1 - \left(1 - c\hat{\lambda}_j^2\right)^{1/\tau})(z'\hat{\phi}_j)\hat{\phi}_j.$$

Here, the regularization parameter τ is such that $1/\tau - 1$ represents the number of iterations.

Spectral cut-off (SC) regularization

Spectral Cut-off regularization provides a mechanism to suppress the influence of smaller eigenvalues on the estimation of the variance-covariance matrix by truncating or setting a threshold. This effectively dampens the impact of multicollinearity and high-dimensionality, enhancing the stability of the inverse variance-covariance matrix estimator.

By mitigating the impact of small eigenvalues, SC regularization aims to produce a more robust estimate, particularly in situations where the variance-covariance matrix may be ill-conditioned. Formally, SC regularization consists in selecting the eigenvectors associated with the eigenvalues greater than some threshold:

$$(\bar{R}'\bar{R}/T)^{\tau}z = \sum_{\hat{\lambda}_j^2 > \tau} \frac{1}{\hat{\lambda}_j}(z'\hat{\phi}_j)\hat{\phi}_j,$$

for $\tau > 0$. Since the $\hat{\phi}_j$ are related to the principal components of $(\bar{R}'\bar{R}/T)$, this technique is also called principal components (PC) regularization.

Notice that the above three regularization techniques (R, LF, and SC) can be rewritten using a common notation as:

$$(\bar{R}'\bar{R}/T)^{\tau}z = \sum_{j=1}^{N} \frac{q\left(\tau, \hat{\lambda}_j^2\right)}{\hat{\lambda}_j}(z'\hat{\phi}_j)\hat{\phi}_j,$$

where $q\left(\tau, \hat{\lambda}_j^2\right)$ is a weight that takes different forms depending on the regularization scheme:

- $q\left(\tau, \hat{\lambda}_j^2\right) = \hat{\lambda}_j^2 / \left(\tau + \hat{\lambda}_j^2\right)$ for R regularization,
- $q\left(\tau, \hat{\lambda}_j^2\right) = 1 - \left(1 - c\hat{\lambda}_j^2\right)^{1/\tau}$ for LF regularization.
- $q\left(\tau, \hat{\lambda}_j^2\right) = I\left(\hat{\lambda}_j^2 \geq \tau\right)$ for SC regularization.

All these regularization techniques involve a tuning parameter denoted as τ. The case where $\tau = 0$ corresponds to absence of regularization, resulting in $q(\tau, \hat{\lambda}_j^2) = 1$. In Section 4.1, we present a data-driven approach for selecting the tuning parameter τ within the context of the portfolio optimization problem discussed in this paper.

Least Absolute Shrinkage and Selection Operator (LASSO) regularization

For linear regression models, Least Absolute Shrinkage and Selection Operator (LASSO) regularization introduces a penalty term to the ordinary least squares objective function, reinforcing sparsity in the estimated coefficients. When applied to the inverse of the variance-covariance matrix, LASSO regularization facilitates variable selection by driving certain elements of the matrix to zero, effectively shrinking less relevant parameters. This regularization method can be particularly valuable in situations where the number of assets is large relative to the sample size, helping to improve the stability and interpretability of the estimated covariance matrix.

Formally, let $\Omega = E\left(r_t r_t'\right) = E\left(R'R\right)/T$, and denote $\beta = \Omega^{-1}\mu = E\left(R'R\right)^{-1} E\left(R' 1_T\right)$. For a mean-variance portfolio, we have

$$\Sigma^{-1}\mu \quad = \quad (\Omega - \mu\mu')^{-1}\mu = (\Omega^{-1} + \frac{\Omega^{-1}\mu\mu'\Omega^{-1}}{1-\mu'\Omega^{-1}\mu})\mu = \frac{\Omega^{-1}\mu}{1-\mu'\Omega^{-1}\mu} = \frac{\beta}{1-\mu'\beta}.$$

Notice that the Lasso regularization technique introduced by Tibshirani (1996) is the l_1-penalized version of $\hat{\beta} = argmin||1_T - R\beta||_2^2$. Hence, the Lasso regularized solution is obtained by solving:

$$\hat{\beta} = argmin||1_T - R\beta||_2^2 + \tau||\beta||_1.$$

However, when dealing with a portfolio that incorporates asymmetries in preferences and returns, obtaining the closed form of the optimal solution becomes more complicated compared to the standard mean-variance problem.

In the presence of asymmetries, using equations (2.6), and (2.7), we obtain:

$$\omega = \frac{1}{\tilde{\gamma}}\left(\Sigma^{-1}\mu_1 + \tilde{\chi}\Sigma^{-1}\mu_2\right) \tag{3.2}$$

$$\omega = \frac{1}{\tilde{\gamma}}\left(\frac{\beta_1}{1-\mu_1'\beta_1} + \tilde{\chi}\frac{\beta_2}{1-\mu_2'\beta_2}\right),$$

where

$$\mu_1 = \left(\mu - r_f\iota + \frac{1}{2}\sigma^2\right), \beta_1 = \Omega^{-1}\mu_1, \quad \mu_2 = (\sigma \circ \delta), \beta_2 = \Omega^{-1}\mu_2. \tag{3.3}$$

We can now substitute the unknown quantities μ_1 and μ_2 with their sample counterparts and derive the following sample-based optimal allocation:

$$\hat{\omega} = \frac{1}{\tilde{\gamma}}\left(\frac{\hat{\beta}_1}{1-\hat{\mu}_1'\hat{\beta}_1} + \tilde{\chi}\frac{\hat{\beta}_2}{1-\hat{\mu}_2'\hat{\beta}_2}\right).$$

We note that adopting the Lasso approach requires the regularization of two distinct regression coefficients, namely β_1 and β_2. Instead, in this paper, we focus on directly regularizing the covariance matrix of asset returns. We recognize the utility of Lasso regularization in introducing sparsity when inverting the covariance matrix. However, similar to the other regularization procedures, our regularized inverse of the sample covariance matrix obtained through the Lasso method, denoted as $\hat{\Sigma}^\tau$, depends on an unknown tuning parameter. We will optimally select this parameter using a data-driven selection method based on generalized cross-validation (GCV) as shown in the next section.

4 Data-driven method for selecting the tuning parameter

This section aims to present a data-driven method for selecting the tuning parameter τ required for implementing the regularization techniques proposed in Section 3.

As shown earlier, the optimal portfolio weights under asymmetries are given by:

$$\omega = \alpha_{mv}\omega^{MV} + \alpha_{av}\omega^{AV},$$

where $\alpha_{mv} = 1/\tilde{\gamma}$ and $\alpha_{av} = \tilde{\chi}/\tilde{\gamma}$. As previously mentioned, the weights ω^{MV} and ω^{AV} are estimated as follows

$$\hat{\omega}_\tau^{MV} = \hat{\Sigma}^\tau\left(\hat{\mu} - r_f\iota + \frac{1}{2}\hat{\sigma}^2\right), \quad \hat{\omega}_\tau^{AV} = (\hat{\Sigma}^\tau)^{-1}(\hat{\sigma} \circ \hat{\delta}). \tag{4.1}$$

Consequently, the optimal portfolio weights ω can be estimated by replacing the unknown parameters by their estimates from the sample:

$$\hat{\omega}_\tau = \alpha_{mv}\hat{\omega}_\tau^{MV} + \alpha_{av}\hat{\omega}_\tau^{AV}. \tag{4.2}$$

The estimator of the optimal weights in (4.2), however, cannot be implemented in practice as it depends on the unknown parameters α_{mv}, α_{av}, and τ. On the one hand, as previously discussed, the coefficients α_{mv} and α_{av} depend not only on the preference parameters but also on the optimal asset allocation, ω, making them endogenous to the model. Nonetheless, Dahlquist et al. (2017) provide analytical expressions for these coefficients, which we utilize in this paper for their estimation. On the other hand, our regularized portfolio relies on the tuning parameter τ, which we select using a data-driven method, as illustrated below.

Now to select the tuning parameter τ, we first construct a loss function that minimizes the quadratic distance between the regularized inverse of the estimated covariance matrix and the inverse of the true covariance matrix (e.g. Kan and Zhou (2007), and Frost and Savarino (1986)). Thanks to this loss function, the regularized inverse of the covariance matrix is always invertible, even when $N \geq T$, meaning that this loss function exists for $N \geq T$. Formally, we consider the following loss function:

$$\mu' \left[\left(\hat{\Sigma}^\tau - \Sigma^{-1}\right)' \Sigma \left(\hat{\Sigma}^\tau - \Sigma^{-1}\right) \right] \mu, \tag{4.3}$$

where μ is the expected excess return and $\hat{\Sigma}^\tau$ is the regularized inverse of the sample covariance matrix $\hat{\Sigma}$. The choice of this specific quadratic distance is beneficial for deriving a criterion that can be easily approximated through a generalized cross-validation approach. Using the above loss function, the tuning parameter is selected as follows:

$$\hat{\tau} = arg \min_{\tau \in H_T} E \left\{ \mu' \left[\left(\hat{\Sigma}^\tau - \Sigma^{-1}\right)' \Sigma \left(\hat{\Sigma}^\tau - \Sigma^{-1}\right) \right] \mu \right\}. \tag{4.4}$$

In the following, we present an approximation to the objective function of the above optimization problem, which helps in selecting the tuning parameter based on a generalized cross-validation criterion. For this purpose, we require the following assumption:

Assumption A:

(i) For some $\nu > 0$, we assume that

$$\sum_{j=1}^{N} \frac{<\beta,\phi_j>^2}{\eta_j^{2\nu}} < \infty,$$

where ϕ_j and η_j^2 denote the eigenvectors and eigenvalues of $\frac{\Omega}{N}$, respectively, with $\Omega = E\left(r_t r_t'\right) = E\left(R'R\right)/T$, and $\beta = \Omega^{-1}\mu = E\left(R'R\right)^{-1} E\left(R'1_T\right)$.

(ii) $\frac{\Sigma}{N}$ and $\frac{\Omega}{N}$ are trace class operators.

Notice that the regularity condition A(i) is utilized in several papers, including Carrasco and Doukali (2017). Regarding A(ii), it's worth noting that the trace-class operator K is defined as a compact operator with a finite trace, denoted as $Tr(K) = O(1)$. This assumption A(ii) is considered more realistic than assuming that Σ is a Hilbert-Schmidt operator.

From Equation (4.4), under assumption A and building on the findings from Proposition 1 in Carrasco and Noumon (2011), using a generalized cross-validation criterion (see Wahba (1975), Li (1986, 1987), and Andrews (1991) among others), the optimal value of τ is determined by the following optimization problem:

$$\hat{\tau} = \arg\min_{\tau \in H_T} \left\{ GCV\left(\tau\right) + \frac{\left(1_T'\left(M_T\left(\tau\right) - I_T\right) X \hat{\beta}_{\widetilde{\tau}}\right)^2}{T^2\left(1 - \hat{\mu}'\hat{\beta}_{\widetilde{\tau}}\right)} \right\},$$

where $H_T = \{1, 2, ..., T\}$ for spectral cut-off and Landweber Fridman and $H_T = (0, 1)$ for Ridge. $GCV\left(\tau\right)$ is the generalized cross validation criterion:

$$GCV\left(\tau\right) = \frac{1}{T} \frac{\left\|\left(I_T - M_T\left(\tau\right)\right) 1_T\right\|^2}{\left(1 - tr\left(M_T\left(\tau\right)\right)/T\right)^2},$$

and the operator $M_T\left(\tau\right)$ is defined by the following identity:

$$\hat{\mu}'\left(\beta_\tau - \beta\right) = \frac{1_T'}{T}\left(M_T\left(\tau\right) - I_T\right) X\beta.$$

Notice that $\hat{\beta}_{\widetilde{\tau}}$ is an estimator of β which we obtain using some initial consistent estimator $\widetilde{\tau}$ of τ (e.g. $\widetilde{\tau}$ can be obtained by minimizing $GCV(\tau)$). In the following, we show how one can estimate β using Lasso method; see Tibshirani (1996).

Let $\tilde{\beta}(\tau)$ be the Lasso estimator of β for a given value of τ. By replacing the Lasso's penalty term $\sum|\beta_j|$ by the term $\sum\beta_j^2/|\beta_j|$, we can show that $\tilde{\beta}(\tau)$ can be approximated by:

$$\beta^* = \left(R'R + \tau(c)W^-(\tau)\right)^{-1} R' 1_T,$$

where c is the upper bound $\sum|\beta_j|$ in the constrained problem equivalent to the penalized Lasso, $W(\tau)^-$ is the generalized inverse of the diagonal matrix $W(\tau)$ with diagonal elements $\left|\tilde{\beta}_j(\tau)\right|$, and $\tau(c)$ is chosen so that $\sum_j\left|\beta_j^*\right| = c$. The term $\tau(c)$ represents the Lagrangian multiplier for the constraint $\sum_j\left|\beta_j^*\right| \leq c$. This constraint is always binding for $\tau(c) \neq 0$ (ill-posed cases). Now, for

$$p(\tau) = tr\left\{R\left(R'R + \tau(c)W^-(\tau)\right)^{-1} R'\right\},$$

The generalized cross-validation criterion using Lasso method is given by:

$$GCV(\tau) = \frac{1}{T}\frac{\left\|1_T - R\widetilde{\beta}(\tau)\right\|^2}{\left(1 - p(\tau)/T\right)^2}.$$

Using Monte Carlo simulations, Tibshirani (1996) shows that the above criterion leads to satisfactory results in terms of estimating β.

5 Generalized Sharpe Ratio

The Sharpe ratio is a widely recognized metric for evaluating the performance of an investment strategy. Nevertheless, it is only valid in cases where returns are normally distributed or when quadratic preferences are involved. When returns are not normally distributed, as is the case in this paper, the Sharpe ratio can yield misleading conclusions; see Hodges (1998), Bernardo and Ledoit (2000*b*), and Kadan and Liu (2014)). To address this limitation, Zakamouline and Koekebakker (2009) introduced a portfolio performance measure—the generalized Sharpe ratio (GSR)—which accounts for higher moments of the distribution and captures asymmetry in returns.

Expanding on the work of Zakamouline and Koekebakker (2009), hereafter we derive a closed-form solution for the GSR that integrates the first four moments of the normal-exponential distribution of returns. As outlined by Zakamouline and Koekebakker (2009) (page 1247, Equation 26), the GSR can be calculated using the following formula:

$$\frac{1}{2}GSR^2 = -log(-E[U(\omega)]), \tag{5.1}$$

where ω is the investor's optimal portfolio weights that should take into account asymmetries in returns and risk attitudes.

We now need to derive the expression of $E[U(\omega)]$ in terms of the returns distribution, preference parameters, and other relevant factors. This expression can be obtained by addressing the following mean-variance-asymmetry investment problem:

$$max_{\omega}E(U(\omega)) = max_{\omega}\mu_{\omega} - r_f - \frac{\tilde{\gamma}-1}{2}\sigma_{\omega}^2 + \tilde{\chi}\sigma_{\omega}\delta_{\omega}. \tag{5.2}$$

After solving the above problem for the optimal allocation as in Equation (2.7), we can substitute this solution into Equation (5.2) to obtain the desired expression of the GSR under asymmetries in returns and risk attitudes, which we provide in the following proposition.

Proposition 1. *The generalized Sharpe ratio of a portfolio of N risky assets generated by a normal-exponential model is given by:*

$$GSR(\omega) = \sqrt{-2log(-\mu_{\omega} + r_f + \frac{\tilde{\gamma}-1}{2}\sigma_{\omega}^2 - \tilde{\chi}\sigma_{\omega}\hat{\delta}_{\omega})}, \tag{5.3}$$

where ω is the portfolio weights, μ_{ω}, σ_{ω}^2, and δ_{ω} are the mean, standard deviation, and asymmetry coefficient of the portfolio returns, respectively, $\tilde{\gamma}$ and $\tilde{\chi}$ are the preference parameters, and r_f is the risk free asset.

It is worth noting that the GSR is computed subsequent to resolving the optimal portfolio choice problem after incorporating asymmetries and addressing the ill-posed inverse of the covariance matrix of returns.

6 Monte Carlo simulations

We conduct extensive Monte Carlo simulations to evaluate the performance of our methodology for selecting large portfolios, taking into account asymmetries in asset returns and risk attitudes, and the challenges posed by the ill-conditioned inverse of the covariance matrix of returns. In particular, we assess and compare the generalized Sharpe ratio and the expected loss in utility of our portfolios with those of several benchmark portfolios. Specifically, we compare our regulated portfolios (hereafter R, LF, SC, and Lasso based portfolios) with the following portfolios: the equally weighted portfolio (hereafter naive portfolio), the regularized mean-variance (hereafter MvP) portfolios, the ridge-regularized portfolio (hereafter R), the Spectral cut-off (principal component) portfolio (hereafter SC), the Landweber-Fridman portfolio (hereafter LF), and the Lasso portfolio (Lasso). Additionally, we report results for the portfolio proposed by Dahlquist et al. (2017) (hereafter DFT).

We use Equation (2.2) to simulate returns under asymmetry. In our simulation study, we generate three different sizes for the vector of risky assets, representing small, moderate, and large portfolios: $N = \{30, 50, 100\}$. We use various values of N to investigate how the size of the portfolio affects its performance. The case $N = 100$ presents the most challenging ill-posed problem, while the others can be viewed as less severe cases.

The unknown parameters μ, δ, and Σ (or σ^2) are calibrated using the monthly 30-industry portfolios (FF30), the monthly 48-industry portfolios (FF48), and the monthly 100 portfolios formed on size and book-to-market (FF100) from Fama and French data library. The return on the risk-free asset R_f is calibrated to match the mean of the one-month Treasury-Bill (T-Bill) observed in the data from July 1980 to June 2019. We assume the covariance matrix of the residual vector to be diagonal, denoted as Σ_ε=diag$\left(\sigma_1^2, ..., \sigma_N^2\right)$, where the diagonal elements are drawn from a uniform distribution between 0.10 and 0.30, resulting in an average cross-sectional volatility of 20%.

We assess the performance of each investment portfolio across various levels of the investor's degree of disappointment aversion, $l = 1, 2, 5$. The investor's risk aversion γ is set to 2, with the threshold for outcomes considered disappointing κ at 0.95. Additional results, not included in this paper but available upon request, were obtained by exploring alternative values for the coefficients of γ, l, and κ. All results are based on $C = 1000$ replications and sample size $T = 120$.

As we mentioned previously, one of the performance measures we use to evaluate and compare portfolios is the generalized Sharpe ratio given in Proposition 1. Formally, at each simulation j, for $j = 1, \ldots, C$, this generalized Sharpe ratio is estimated as follows:

$$GSR(\hat{\omega}_{\hat{\tau}}) = \sqrt{-2log\left(-\mu_{\hat{\omega}_{\hat{\tau}_j,j}} + r_f + \frac{\tilde{\gamma}-1}{2}\sigma^2_{\hat{\omega}_{\hat{\tau}_j,j}} - \tilde{\chi}\sigma_{\hat{\omega}_{\hat{\tau}_j,j}}\delta_{\hat{\omega}_{\hat{\tau}_j,j}}\right)} \tag{6.1}$$

where

$$\begin{aligned}
\mu_{\hat{\omega}_{\hat{\tau}_j,j}} &= r_f + \hat{\omega}'_{\hat{\tau}_j,j}\left(\mu - r_f\iota + \frac{1}{2}\sigma^2\right) - \frac{1}{2}\hat{\omega}'_{\hat{\tau}_j,j}\Sigma\hat{\omega}_{\hat{\tau}_j,j},\\
\sigma^2_{\hat{\omega}_{\hat{\tau}_j,j}} &= \hat{\omega}'_{\hat{\tau}_j,j}\Sigma\hat{\omega}_{\hat{\tau}_j,j},\\
\delta_{\hat{\omega}_{\hat{\tau}_j,j}} &= \frac{\hat{\omega}'_{\hat{\tau}_j,j}(\sigma\circ\delta)}{\sigma_{\hat{\omega}_{\hat{\tau}_j,j}}}.
\end{aligned}$$

For the C replications, the performance measure $GSR(\hat{\omega}_{\hat{\tau}})$ is estimated as follows:

$$G\hat{S}R(\hat{\omega}_{\hat{\tau}}) = \frac{1}{C}\sum_{j=1}^{C}\sqrt{-2log\left(-\mu_{\hat{\omega}_{\hat{\tau}_j,j}} + r_f + \frac{\tilde{\gamma}-1}{2}\sigma^2_{\hat{\omega}_{\hat{\tau}_j,j}} - \tilde{\chi}\sigma_{\hat{\omega}_{\hat{\tau}_j,j}}\delta_{\hat{\omega}_{\hat{\tau}_j,j}}\right)}.$$

Tables 1, 2, and 3 present the simulation results for the Generalized Sharpe Ratio (GSR) of both our regularized portfolios and benchmark portfolios for $N = 30$, $N = 50$, and $N = 100$, respectively. A higher GSR signifies better portfolio performance. Notably, we observe a significant improvement in maximizing the GSR of our regularized portfolios compared to the benchmark portfolios, which include the portfolio proposed by Dahlquist et al. (2017). The latter portfolio incorporates asymmetries but does not address the ill-posed inverse of the covariance matrix of returns. Interestingly, we observe that the regularized mean-variance portfolios perform much better than the Dahlquist et al. (2017) portfolio. Furthermore, our regularized portfolios consistently outperform all benchmark portfolios across different values of the degree of disappointment aversion, l, and portfolio sizes, N. Particularly noteworthy is the superior performance of our portfolios compared to the benchmark portfolios, especially when N is large and for small degrees of disappointment aversion. These results clearly illustrate the importance of incorporating asymmetries and addressing the ill-posed inverse of the covariance matrix of returns when optimizing portfolios.

Table 1: Generalized Sharpe ratio for $N = 30$, $\gamma = 2$, and $\kappa = 0.95$

	MvP				DFT	Optimal portfolio				Naive
	R	SC	LF	Lasso		R	SC	LF	Lasso	Portfolio
$l = 1$	0.2054	0.1813	0.2082	0.1625	0.1305	0.2281	0.2146	0.2258	0.2214	0.2053
$l = 2$	0.2132	0.1837	0.2121	0.2050	0.1208	0.2339	0.2188	0.2304	0.2258	0.2053
$l = 5$	0.2153	0.2019	0.2121	0.2108	0.1582	0.2359	0.2234	0.2319	0.2318	0.2053

Table 2: Generalized Sharpe ratio for $N = 50$, $\gamma = 2$, and $\kappa = 0.95$

	MvP				DFT	Optimal portfolio				Naive
	R	SC	LF	Lasso		R	SC	LF	Lasso	Portfolio
$l = 1$	0.2274	0.1528	0.2194	0.2306	0.1126	0.2446	0.1876	0.2329	0.2485	0.2016
$l = 2$	0.2277	0.1617	0.2176	0.2314	0.1186	0.2467	0.1939	0.2341	0.2521	0.2016
$l = 5$	0.2320	0.1635	0.2257	0.2363	0.1598	0.2516	0.1948	0.2422	0.2566	0.2016

When comparing the performance of our four regularized portfolios, Tables 1-3 indicate that for small and moderate-sized portfolios, Ridge regularization yields better performance compared to the other regularization-based portfolios. However, for large-sized portfolios, the Landweber-Fridman portfolio outperforms the others. Additionally, we observe that the spectral cut-off (or principal component) portfolio is generally outperformed by the other regularized portfolios. This result is plausible, as the returns were generated using a model that does not exhibit a factor structure. In other words, the SC-based portfolio is expected to perform better when the returns exhibit a factor structure.

Table 3: Generalized Sharpe ratio for $N = 100$, $\gamma = 2$, and $\kappa = 0.95$

	MvP				DFT	Optimal portfolio				Naive
	R	SC	LF	Lasso		R	SC	LF	Lasso	Portfolio
$l = 1$	0.1774	0.1398	0.1711	0.1778	0.1323	0.2405	0.2032	0.2503	0.2400	0.1982
$l = 2$	0.1775	0.1547	0.1717	0.1835	0.1459	0.2446	0.2028	0.2586	0.2421	0.1982
$l = 5$	0.1787	0.1653	0.1754	0.1852	0.1616	0.2538	0.2134	0.2653	0.2471	0.1982

We next conducted additional simulations to investigate and compare the performances of our regularized portfolios with those of benchmark portfolios in terms of the expected loss in the utility function. Note that the optimal solution in Equation (2.6) can be obtained by solving the following mean-variance-asymmetry investment problem:

$$\max_{\omega} \mu_{\omega} - r_f - \frac{\tilde{\gamma} - 1}{2} \sigma_{\omega}^2 + \tilde{\chi} \sigma_{\omega} \delta_{\omega}. \tag{6.2}$$

Thereafter, we use the above portfolio problem to define the following loss in utility function:

$$L_T\left(\hat{\omega}_{\hat{\tau}}\right) = U\left(\omega\right) - U\left(\hat{\omega}_{\hat{\tau}}\right), \tag{6.3}$$

where $U\left(\omega\right) = \mu_\omega - r_f - \frac{\tilde{\gamma}-1}{2}\sigma_\omega^2 + \tilde{\chi}\sigma_\omega\delta_\omega$ is the true utility function and $U(\hat{\omega}_{\hat{\tau}})$ is the estimated one based on one of the portfolios under consideration. Thus, the expected loss in utility for a given optimal portfolio is $E\left[L_T\left(\hat{\omega}_{\hat{\tau}}\right)\right]$. In our simulation, this quantity is approximated by:

$$\hat{E}\left[L_T\left(\hat{\omega}_{\hat{\tau}}\right)\right] = \frac{1}{C}\sum_{j=1}^{C} L_T\left(\hat{\omega}_{\hat{\tau}_j,j}\right). \tag{6.4}$$

Table 4: Expected loss in utility for $N = 30$, $\gamma = 2$, and $\kappa = 0.95$

	MvP				DFT	Optimal portfolio				Naive
	R	SC	LF	Lasso		R	SC	LF	Lasso	Portfolio
$l = 1$	0.04886	0.08709	0.0186	0.0692	0.07321	0.00771	0.01464	0.00851	0.00998	0.07348
$l = 2$	0.04097	0.08266	0.01716	0.0588	0.0789	0.00721	0.01401	0.00672	0.00733	0.07348
$l = 5$	0.03388	0.03252	0.01246	0.04563	0.0601	0.00644	0.00982	0.00564	0.00660	0.07348

Tables 4, 5, and 6 present the results of the expected loss in utility for portfolio sizes $N = 30$, $N = 50$, and $N = 100$, respectively. A portfolio is considered to perform well in terms of expected loss in utility if it leads to a minimal utility loss. Our analysis reveals a significant improvement in minimizing the expected loss in utility of our regularized portfolios compared to the benchmark portfolios. Our findings indicate that the proposed optimal regularized portfolios consistently outperform the naïve portfolio, the regularized mean-variance portfolios, and Dahlquist et al. (2017)'s portfolio across all tested sizes ($N = 30$, $N = 50$, and $N = 100$). The gap in expected utility loss between our regularized portfolios and the regularized mean-variance portfolio is notably large. This underscores the importance of considering asymmetries in returns and risk attitudes, as well as addressing the ill-posed inverse of the covariance matrix of returns, when optimizing portfolios.

Table 5: Expected loss in utility for $N = 50$, $\gamma = 2$, and $\kappa = 0.95$

	MvP				DFT	Optimal portfolio				Naive
	R	SC	LF	Lasso		R	SC	LF	Lasso	Portfolio
$l = 1$	0.0368	0.0689	0.0370	0.0387	0.1033	0.0047	0.0158	0.0054	0.0046	0.0784
$l = 2$	0.0320	0.0550	0.0337	0.0318	0.1172	0.0041	0.0117	0.0050	0.0042	0.0784
$l = 5$	0.0025	0.0487	0.0308	0.0227	0.0944	0.00328	0.0097	0.0042	0.0035	0.0784

Table 6: Expected loss in utility for $N = 100$, $\gamma = 2$, and $\kappa = 0.95$

	MvP				DFT	Optimal portfolio				Naive
	R	SC	LF	Lasso		R	SC	LF	Lasso	Portfolio
$l = 1$	0.01065	0.0915	0.0559	0.0156	0.1480	0.00406	0.0038	0.0040	0.00407	0.0805
$l = 2$	0.0095	0.0317	0.0467	0.0103	0.1590	0.00403	0.0036	0.0039	0.00403	0.0805
$l = 5$	0.0084	0.0208	0.0453	0.009	0.1289	0.00398	0.0034	0.0039	0.00398	0.0805

Table 7 presents descriptive statistics on the optimal tuning parameter selected for each stabilization method, computed for three different values of N with $l = 2$. The tuning parameter for the SC method reflects the number of factors that can explain the distribution of excess returns. We find that this number of factors increases with N. For the LF method, the tuning parameter corresponds to the number of iterations required for convergence of the portfolio optimization, with results indicating that the average number of iterations is notably higher for $N = 100$ compared to the other values $N = 30$ and $N = 50$. This suggests that a larger number of assets requires more iterations for the algorithm to converge. For the Ridge method, the average selected tuning parameter also increases with the size of the portfolio. Essentially, a larger portfolio requires greater shrinkage of the covariance matrix to improve the optimal portfolio's performance.

Table 7: Distribution characteristics of the optimal regularization parameters. The number of iterations performed is 1000 for $T = 120$ observations

Rule	R			SC			LF		
N	30	50	100	30	50	100	30	50	100
Mean	0.0869	0.089	0.0947	6.925	5.953	7.558	222.8	212.037	281.336
Std	0.0055	0.0054	0.0066	2.3552	2.6038	2.3766	56.6342	64.1889	57.0728
q1	0.0458	0.0658	0.0784	5	4	6	179	162	234
Median	0.0898	0.0897	0.0982	7	6	8	212	199	283
q3	0.0987	0.102	0.106	9	8	10	257	250	342.5

7 Empirical study

In this section, we empirically assess the performance of our regularized optimal portfolios using the rolling sample approach introduced by Kan and Smith (2008). We also conduct a comparative analysis with benchmark portfolios employed in the Monte Carlo simulation section. The rolling sample approach involves recursively generating $T - M$ out-of-sample returns from a dataset of size T, utilizing M previous returns. Subsequently, for each rolling window, the regularized optimal portfolios are hold for one year. The resulting time series of out-of-sample optimal portfolio returns are then used to compute the out-of-sample performance for both our regularized portfolios and the benchmark portfolios.

We implement our regularized portfolios using three sets of monthly portfolios obtained from Kenneth French's website, covering the period spanning from July 1980 to June 2017: the monthly 30-industry portfolios (FF30), the 48 industry portfolios (FF48), and the monthly 100 portfolios formed based on size and book-to-market (FF100). Following the approach outlined in Barberis (2000), investors are allowed to rebalance their portfolios annually. This involves constructing the optimal portfolios at the end of June of each year, covering the period from July 1980 to June 2017, using a designated estimation window (EW). In our study, we consider two different estimation windows: 60 and 120. Subsequently, investors hold these optimal portfolios for one year, realizing gains and losses, updating information, and recalculating optimal portfolio weights for the subsequent year. As shown by DeMiguel, Garlappi and Uppal (2009); DeMiguel, Garlappi, Nogales and Uppal (2009), this procedure yields results consistent with those obtained when rebalancing optimal portfolios on a monthly basis.

The procedure outlined above is iterated for each year within the dataset, resulting in a time series of out-of-sample portfolio returns. This time series is subsequently used to evaluate and compare the out-of-sample performance of both our regularized portfolios and benchmark portfolios based on the Generalized Sharpe ratio. We consider an appropriate range for the tuning parameter to conduct our portfolio optimization, which varies depending on the regularization approach employed. Specifically, for the Ridge approach, we employ a grid spanning the interval $[0,1]$, with a precision set at 10^{-2}. In the case of the Spectral Cut-off (SC) approach, the parameter ranges from 1 to a maximum value $N-1$, where N represents the number of assets. For the Landweber-Fridman (LF) approach, we set a maximum number of iterations at 300.

Subsequently, we evaluate the performance of each optimal portfolio setting the investor's degree of disappointment aversion at $l = 2$. The investor's risk aversion γ is held constant at 2, and the threshold for the investor's certainty equivalent, below which outcomes are deemed disappointing, is set at $\kappa = 0.95$. Additional results, not included in this report but available upon request, were obtained by exploring alternative values for the coefficients γ, l, and κ.

Table 8: Out-of-sample generalized Sharpe (GSR) ratio for standard mean-variance portfolio

	MvP					
P	FF30		FF48		FF100	
EW	60	120	120	240	120	240
GSR	0.069	0.095	0.076	0.099	0.078	0.106

Table 9: Out-of-sample generalized Sharpe ratio of our regularized portfolios, regularized mean-variance portfolios, and naive portfolio

Portfolios	EW	MvP				Optimal portfolio				Naive
		R	SC	LF	Lasso	R	SC	LF	Lasso	Portfolio
FF30	60	0.1001	0.1862	0.1167	0.1293	0.3419	0.3186	0.3478	0.2541	0.2385
	120	0.1545	0.2569	0.1845	0.1857	0.3803	0.3191	0.3744	0.2694	0.2335
FF48	120	0.1201	0.1985	0.1858	0.1698	0.1568	0.3120	0.3519	0.3597	0.2389
	240	0.1865	0.2754	0.2098	0.2189	0.3840	0.3350	0.3745	0.3687	0.2468
FF100	120	0.1432	0.2120	0.1858	0.2167	0.3518	0.3119	0.3597	0.3467	0.2415
	240	0.2619	0.2966	0.2930	0.2572	0.3880	0.3463	0.3741	0.3815	0.2690

Tables 8 and 9 present the results of the out-of-sample analysis using the three data sets FF30, FF48, and FF100. Our results show that our regularized portfolios significantly outperform both the standard and regularized mean-variance portfolios, as well as the naïve portfolios, in maximizing the out-of-sample generalized Sharpe ratio. This empirical evidence underscores the importance of incorporating asymmetries and addressing the ill-posed inverse of the covariance matrix of returns when optimizing portfolios. Additionally, we observe that the performance of all portfolios improves as the estimation window (EW) increases. This improvement might be attributed to the reduction in estimation error associated with unknown parameters, thereby positively impacting the performance of the selected portfolio.

8 Conclusion

We investigated the portfolio choice problem for an investor dealing with asymmetries in returns and risk attitudes in an economy with a large number of assets. This selection problem requires estimating the covariance matrix of returns and computing its inverse. However, inverting the covariance matrix can be challenging, as it may become ill-conditioned when dealing with a substantial number of assets.

We proposed addressing the portfolio choice problem in the presence of asymmetries and high-dimensional covariance matrices of asset returns by employing four regularization techniques: Ridge, spectral cut-off, Landweber-Fridman, and Lasso. These regularization approaches involve a tuning parameter that must be carefully selected. To this end, we introduced a data-driven approach for selecting the optimal tuning parameter by minimizing the distance between the inverse of the estimated covariance matrix and the inverse of the population covariance matrix. To assess the performance of our chosen portfolio, we derived the generalized Sharpe ratio (GSR), which accounts for asymmetries through higher moments of the return distribution.

Furthermore, we conducted an extensive Monte Carlo simulation to assess the performance of our regularized optimal portfolios in the presence of asymmetries and large number of assets. We compared the performance of our portfolios with those of several benchmark portfolios including the portfolio proposed by Dahlquist et Dahlquist et al. (2017), the regularized mean-variance portfolios, and the naïve portfolio. The simulation results demonstrated that our regularized portfolios outperformed all benchmark portfolios in terms of the generalized Sharpe ratio and the expected loss in utility.

Finally, we applied our portfolio approaches to three sets of portfolios: the 30 industry portfolios, 48 industry portfolios, and 100 portfolios formed based on size and book-to-market. The empirical findings indicate that by accounting for asymmetries and stabilizing the inverse of the covariance matrix, we significantly enhance the performance of the optimal portfolio in maximizing the generalized Sharpe ratio.

References

Abadir, K. M., Distaso, W. and Žikeš, F. (2014), 'Design-free estimation of variance matrices', *Journal of Econometrics* **181**(2), 165–180.

Andrews, D. W. (1991), 'Asymptotic optimality of generalized cl, cross-validation, and generalized cross-validation in regression with heteroskedastic errors', *Journal of Econometrics* **47**(2-3), 359–377.

Ang, A. and Chen, J. (2002), 'Asymmetric correlations of equity portfolios', *Journal of financial Economics* **63**(3), 443–494.

Ao, M., Yingying, L. and Zheng, X. (2019), 'Approaching mean-variance efficiency for large portfolios', *The Review of Financial Studies* **32**(7), 2890–2919.

Bai, J. and Ng, S. (2002), 'Determining the number of factors in approximate factor models', *Econometrica* **70**(1), 191–221.

Barberis, N. (2000), 'Investing for the long run when returns are predictable', *The Journal of Finance* **55**(1), 225–264.

Bernardo, A. E. and Ledoit, O. (2000*a*), 'Gain, loss and asset pricing', *Journal of Political Economy* **108**(1), 144–172.

Bernardo, A. E. and Ledoit, O. (2000*b*), 'Gain, loss, and asset pricing', *Journal of political economy* **108**(1), 144–172.

Brodie, J., Daubechies, I., De Mol, C., Giannone, D. and Loris, I. (2009), 'Sparse and stable markowitz portfolios', *Proceedings of the National Academy of Sciences* **106**(30), 12267–12272.

Campbell, J. and Viceira, L. (2002), *Strategic asset allocation: Portfolio choice for long-term investors*, Oxford: Oxford University Press.

Carrasco, M. and Doukali, M. (2017), 'Efficient estimation using regularized jackknife iv estimator', *Annals of Economics and Statistics/Annales d'Économie et de Statistique* (128), 109–149.

Carrasco, M. and Noumon, N. (2011), ‘Optimal portfolio selection using regularization’, *Citeseer, Tech. Rep.* .

Dahlquist, M., Farago, A. and Tédongap, R. (2017), ‘Asymmetries and portfolio choice’, *The Review of Financial Studies* **30**(2), 667–702.

De Mol, C., Giannone, D. and Reichlin, L. (2008), ‘Forecasting using a large number of predictors: Is bayesian shrinkage a valid alternative to principal components?’, *Journal of Econometrics* **146**(2), 318–328.

DeMiguel, V., Garlappi, L., Nogales, F. J. and Uppal, R. (2009), ‘A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms’, *Management science* **55**(5), 798–812.

DeMiguel, V., Garlappi, L. and Uppal, R. (2009), ‘Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy?’, *Review of Financial Studies* **22**(5), 1915–1953.

Fan, J., Li, Y. and Yu, K. (2012), ‘Vast volatility matrix estimation using high-frequency data for portfolio selection’, *Journal of the American Statistical Association* **107**(497), 412–428.

Fastrich, B., Paterlini, S. and Winker, P. (2015), ‘Constructing optimal sparse portfolios using regularization methods’, *Computational Management Science* **12**(3), 417–434.

Friedman, J., Hastie, T. and Tibshirani, R. (2008), ‘Sparse inverse covariance estimation with the graphical lasso’, *Biostatistics* **9**(3), 432–441.

Frost, P. A. and Savarino, J. E. (1986), ‘An empirical bayes approach to efficient portfolio selection’, *Journal of Financial and Quantitative Analysis* **21**(3), 293–305.

Guidolin, M. and Timmermann, A. (2008), ‘International asset allocation under regime switching, skew, and kurtosis preferences’, *Review of Financial Studies* **21**(2), 889–935.

Gul, F. (1991), ‘A theory of disappointment aversion’, *Econometrica* **59**(3), 667–86.

Hodges, S. (1998), *A generalization of the Sharpe ratio and its applications to valuation bounds and risk measures*, Financial Options Research Centre, Warwick Business School, University of Warwick.

Hong, Y., Tu, J. and Zhou, G. (2007), 'Asymmetries in stock returns: Statistical tests and economic evaluation', *The Review of Financial Studies* **20**(5), 1547–1581.

Jagannathan, R. and Ma, T. (2003), 'Risk reduction in large portfolios: Why imposing the wrong constraints helps', *The Journal of Finance* **58**(4), 1651–1683.

Jondeau, E. and Rockinger, M. (2006), 'Optimal portfolio allocation under higher moments', *European Financial Management* **12**(1), 29–55.

Kadan, O. and Liu, F. (2014), 'Performance evaluation with high moments and disaster risk', *Journal of Financial Economics* **113**(1), 131–155.

Kan, R. and Smith, D. R. (2008), 'The distribution of the sample minimum-variance frontier', *Management Science* **54**(7), 1364–1380.

Kan, R. and Zhou, G. (2007), 'Optimal portfolio choice with parameter uncertainty', *Journal of Financial and Quantitative Analysis* **42**(3), 621–656.

Kanwal, R. P. (2013), *Linear integral equations*, Springer Science & Business Media.

Kress, R., Maz'ya, V. and Kozlov, V. (1989), *Linear integral equations*, Vol. 82, Springer.

Ledoit, O. and Wolf, M. (2003), 'Improved estimation of the covariance matrix of stock returns with an application to portfolio selection', *Journal of empirical finance* **10**(5), 603–621.

Ledoit, O. and Wolf, M. (2004), 'Honey, i shrunk the sample covariance matrix', *The Journal of Portfolio Management* **30**(4), 110–119.

Ledoit, O. and Wolf, M. (2017), 'Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks', *The Review of Financial Studies* **30**(12), 4349–4388.

Li, K.-C. (1986), 'Asymptotic optimality of cl and generalized cross-validation in ridge regression with application to spline smoothing', *The Annals of Statistics* pp. 1101–1112.

Li, K.-C. (1987), ‘Asymptotic optimality for cp, cl, cross-validation and generalized cross-validation: discrete index set’, *The Annals of Statistics* pp. 958–975.

Markowitz, H. (1952), ‘Portfolio selection’, *The Journal of Finance* pp. 77–91.

Martellini, L. and Ziemann, V. (2010), ‘Improved estimates of higher-order comoments and implications for portfolio selection’, *Review of Financial Studies* **23**(4), 1467–502.

Mencía, J. and Sentana, E. (2009), ‘Multivariate location–scale mixtures of normals and mean–variance–skewness portfolio allocation’, *Finance* **153**(2), 105–121.

Routledge, B. R. and Zin, S. E. (2010), ‘Generalized disappointment aversion and asset prices’, *The Journal of Finance* **65**(4), 1303–1332.

Tédongap, R. and Tinang, J. (2022), ‘Portfolio optimization and asset pricing implications under returns non-normality concerns’, *Finance* **43**(1), 47–94.

Tibshirani, R. (1996), ‘Regression shrinkage and selection via the lasso’, *Journal of the Royal Statistical Society: Series B (Methodological)* **58**(1), 267–288.

Wahba, G. (1975), ‘Smoothing noisy data with spline functions’, *Numerische mathematik* **24**(5), 383–393.

Zakamouline, V. and Koekebakker, S. (2009), ‘Portfolio performance evaluation with generalized sharpe ratios: Beyond the mean and variance’, *Journal of Banking & Finance* **33**(7), 1242–1254.