The impact of linear covariance matrix shrinkage on mean-variance efficient frontier portfolio characteristics

Joseph Gregory Clayton

CLYJOS001

Minor dissertation submitted in partial fulfilment of the requirements for the degree of Master of Commerce specialising in Finance in Investment Management

Faculty of Commerce

Department of Finance and Tax

University of Cape Town

Supervisor: Professor Gary van Vuuren

Co-supervisor: Associate Professor Chun-Sung Huang

February 2025

Plagiarism Declaration

1. I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.
2. I have used the Harvard-UCT convention for citation and referencing. Each contribution to, and quotation in, this submission from the work(s) of other people has been attributed and has been cited and referenced.
3. This paper is my own work.
4. I have not allowed, and will not allow, anyone to copy my/our work with the intention of passing it off as his or her/their own work.
5. I acknowledge that copying someone else's assignment, essay or paper, or part of it, is wrong, and declare that this is my/our own work.

Signature: Signed by candidate

Date: February 12, 2025

Acknowledgements

I wish to express my gratitude to the following persons:

- Professor Gary van Vuuren, who introduced me to the topic of variance-covariance matrix shrinkage estimation and under whose supervision this research was conducted. His guidance and encouragement were very much appreciated.

- Professor Clarence C.Y. Kwan, for his guidance during the preparation of this dissertation. His assistance, particularly in providing information regarding the correct implementation of variance-covariance matrix shrinkage estimation, was instrumental in enhancing my understanding of the methodology. His generosity in sharing this expertise was highly valuable to me.

- Associate Professor Chun-Sung Huang, my co-supervisor, for his advice and support.

Acknowledgment is also due to the National Research Foundation, who provided financial support to me during my degree.

Abstract

In investment cases with many stocks under consideration compared to the available historical return observations, the sample variance-covariance (VCV) matrix is often estimated with considerable error, particularly during market turbulence. Errors arise due to extreme differences in VCV matrix eigenvalues. Traditional mean-variance portfolio optimisation naively employs these extreme values, leading to investment decisions placed on unreliable and unrealistic values. Shrinkage techniques (in which high eigenvalues are reduced, and low eigenvalues increased thereby "shrinking" the range of VCV eigenvalues) is one of the proposed suggestions to address this issue. This study extends previous work by applying linear shrinkage estimators to S&P 500 index-based stock portfolios over the sample period of 01-Jan-18 – 20-May-24. Findings indicate that no linear shrinkage estimator affects the VCV matrix sufficiently to repair the spurious risk-return characteristics of efficient portfolios under stressed market conditions. While some improvements in portfolio risk-adjusted return are observed, the instability of the efficient frontier under stressed market conditions largely remains. Future work could extend this research by considering more sophisticated (non-linear) VCV shrinkage estimators that may mitigate these observed instabilities and yield more stable and improved portfolio performance statistics.

Table of Contents

1. Introduction

1.1 Background and overview

Portfolio selection models, designed to capture institutional features and constraints, require input parameters such as a vector of expected returns and a symmetric covariance matrix of constituent returns. Estimating these parameters is essential, and a common method involves using a sample of past monthly or weekly return observations, assuming they are random draws from a stationary joint probability distribution. The choice of the sample period becomes a trade-off between maintaining stationarity and minimising estimation errors.

The seminal work of Markowitz (1952) introduced mean-variance optimisation, an efficient process for selecting portfolios with the maximum level of expected return for a given level of expected risk. The construction of the optimal portfolio in this framework requires the input of *accurate* parameter estimates. However, practical challenges in using past return observations exist. Chow et al. (1999) highlight that such portfolio theory parameters are often unreliable when estimated from small samples. Typical risk-estimation procedures, which often rely on equally weighting historical observations to estimate risk parameters, may also lead to unstable estimates. This is particularly problematic during periods of market turbulence or financial crises, where asset returns tend to become more volatile and correlated, effectively diminishing the diversification benefits that are assumed under normal market conditions (Chow et al., 1999).

As a result, the estimation of a stable and accurate covariance matrix may be prone to significant error, leading to unstable and misrepresentative risk attributes of portfolios during such periods. Jobson and Korkie (1980) and Michaud (1989) also show that the standard method of computing the full-sample covariance matrix can occasionally contain estimation error and thus provide inaccurate portfolio risk estimates, especially when the number of data points is small or comparable to the number of assets. This estimation error stems from when the most extreme coefficients in the sample covariance matrix tend to display extreme values. Subsequently, the mean-variance optimisation process overweights these most extreme and unreliable coefficients (Ledoit and Wolf, 2004b).

Empirical support for using the average of sample correlations as a predictor of individual correlations has been provided by Elton and Gruber (1973), Elton, Gruber, and Urich (1978) and Chan, Karceski and Lakonishok (1999). While this averaging method effectively reduces overall forecast errors, it may lead to the loss of idiosyncrasies in individual correlations. For example, securities from the same industry have higher sample correlations, in general, than the overall average. Shrinkage estimation, introduced to the finance profession by Ledoit and Wolf (2003, 2004a, 2004b), has been proposed as a remedy to balance reducing errors while retaining some individual correlation idiosyncrasies. Ledoit and Wolf (2012, 2015, 2017, 2020a, 2020b) have subsequently promoted various shrinkage estimators over the past two decades, which can be categorised into linear and nonlinear.

Linear shrinkage involves taking a convex linear combination of the sample covariance matrix and a target matrix of the same dimensions, which can be generic, such as an identity matrix, or incorporate structure (Ledoit and Wolf, 2003, 2004a, 2004b). This provides a bias-variance trade-off by optimally combining two extremes. The optimal shrinkage intensity (λ), which measures the weight that is given to the target matrix, can be derived analytically to minimise the expected loss between the true and estimated covariance matrices (Kwan, 2011).

Ledoit and Wolf (2012, 2015, 2017, 2020a, 2020b) have more recently proposed non-linear shrinkage estimators as alternatives to linear approaches. Nonlinear shrinkage operates on the eigenvalues of the sample covariance matrix, allowing an arbitrary, or nonlinear, transformation of the sample eigenvalues while keeping the eigenvectors (Ledoit and Wolf, 2020a). Nonlinear shrinkage estimators can also be combined with structured targets to incorporate known factor structure. This provides more flexibility than linear shrinkage and can further improve performance, albeit at the cost of being more computationally intensive and often more difficult to understand and implement (Ledoit and Wolf, 2020b).

1.2 Knowledge gap

Variance-covariance (VCV) matrix shrinkage estimation is crucial for portfolio asset allocation and risk management due to its ability to improve the estimation of covariance matrices, which are essential for understanding how asset returns move in relation to one another. Traditional methods of estimating the sample covariance matrix often lead to instability, especially in high-dimensional settings where the number of assets exceeds the number of historical observations. This can often result in poor portfolio performance due to overfitting and estimation errors. The various VCV matrix shrinkage methods, particularly those developed and applied to modern portfolio theory by Ledoit & Wolf (2003, 2004a, 2004b, 2012, 2015, 2017, 2020a, 2020b), aim to address these issues by "shrinking" the sample covariance matrix towards a more stable target matrix. These techniques can reduce the impact of noise in the data, leading to more reliable estimates that improve portfolio optimisation outcomes.

While the theoretical foundations of matrix shrinkage methods are well-established, there is often a gap in understanding their practical applications in real-world portfolio management scenarios. Research can provide empirical evidence on the effectiveness of these methods in various market conditions, thereby bridging the gap between theory and practice. For example, there arguably remains a knowledge gap regarding its application in the context of recent market volatility, particularly during and after the COVID-19 pandemic. The pandemic introduced unprecedented levels of market uncertainty and volatility, which may have altered the relationships between asset returns (Jiang, Liu & Tang, 2023).

There is a need for comparative studies that evaluate the performance of various shrinkage techniques against traditional estimation methods across different asset classes and market environments. Financial markets are dynamic, and the effectiveness of estimation techniques can vary over time. Research that examines the adaptability of matrix shrinkage methods in changing market conditions can fill a gap in the literature, providing insights into their robustness and reliability. This research may inform practitioners about the most effective approaches to use in specific contexts. Research can also explore how matrix shrinkage methods can be integrated with other risk management tools and frameworks. Understanding how these methods complement existing strategies can potentially enhance their adoption and effectiveness in portfolio management.

1.3 Problem statement

Traditional, unshrunk VCV matrices (used in portfolio risk estimation and the identification of optimal portfolios in mean-variance frameworks) often result in unstable values during particularly turbulent market periods. When these matrices are used to assign asset allocations and assess portfolio performance statistics, the instability translates into extreme and highly impractical values.

1.4 Research questions

By applying different shrinkage approaches, can the VCV matrix instability be reduced (or eliminated)? If so, how do the resulting adjusted matrices impact corresponding mean-variance efficient frontier portfolio outputs, such as risk-adjusted returns and volatility?

1.5 Research objectives

The research aims to

- Evaluate the performance of covariance matrix shrinkage techniques:
    - Assess the effectiveness of VCV matrix shrinkage techniques in improving the estimation of covariance matrices for USA equity shares during different market conditions. The focus is on the pre- and post-COVID-19 market environment and relatively volatile period of successive Federal Reserve rate hikes thereafter.
- Analyse mean-variance efficient portfolio optimisation outcomes:
    - Examine the impact of using shrinkage methods on portfolio optimisation outcomes, including annualised cumulative returns, annualised volatilities, and annualised Sharpe ratios. This involves comparing portfolios constructed using shrinkage-adjusted covariance matrices against those constructed with traditional sample covariance matrices.
- Investigate the stability and robustness of resulting portfolios:
    - Investigate the stability and robustness of portfolios constructed with shrinkage methods over the study period, particularly in response to market shocks and changes in economic conditions. This objective focuses on how well these portfolios perform during periods of market stress compared to those using conventional estimation techniques.

1.6 Relevance of the study and potential benefits

A study investigating the impact of covariance matrix shrinkage using data from 2018 to 2024, encompassing the volatile COVID-19 and Federal Reserve rate hike periods, may offer valuable insights for portfolio asset allocation and risk management. The significant market volatility and subsequent market reactions induced by the pandemic provide a robust setting to test the effectiveness of shrinkage methods, crucial for evaluating their robustness in real-world scenarios. By analysing data from both pre- and post-pandemic periods, the study aims to compare the performance of shrinkage methods across different market regimes, helping to identify the most effective techniques for specific conditions. For example, following the initial COVID-19 recovery, inflation surged in the US, prompting the Federal Reserve to implement aggressive interest rate hikes starting in 2022, which continued into 2023. This period of high inflation and rising interest rates presents a unique opportunity to assess the performance of shrinkage methods in contrasting market conditions.

The study also aims to assess the impact of shrinkage on portfolio mean-variance optimisation and risk management, comparing the performance of efficient frontier portfolios optimised with shrinkage methods against those using traditional estimation techniques. The comparison will quantify the benefits in terms of risk-adjusted returns and risk reduction. This will also entail assessing the suitability of the VCV matrix using sample and shrinkage-based estimators as to permit its use in the calculation of asset weights, portfolio risk, and other parameter estimates required for the efficient frontier generation. Such insights may be particularly useful during periods of high uncertainty, enabling practitioners to make more informed portfolio and risk management decisions. Lastly, the study may contribute to the growing body of research on covariance matrix shrinkage. By providing empirical evidence of its effectiveness (or lack thereof) during significant market events, it may help to further bridge the gap between academic research and practical applications in portfolio management.

2. Literature review

Mean-variance optimisation, introduced by Markowitz (1952), is a foundational concept in modern portfolio theory. This approach focuses on constructing investment portfolios that optimise expected return for a given level of risk, measured by the portfolio's variance. Markowitz (1952) demonstrated that diversification can significantly reduce portfolio risk, contrary to the prevailing belief that investors should concentrate their investments in securities with the highest expected returns. By considering the correlation between asset returns, investors can construct a portfolio that minimises risk for a given expected return or maximises return for a given level of risk. This trade-off between return and risk is encapsulated in the efficient frontier, a set of portfolios that offers the highest expected return for a given level of risk. The mean-variance optimisation process involves calculating the expected returns, variances, and covariances of all assets, and then determining the optimal asset weights that achieve the desired balance between return and risk.

Accurately estimating the covariance matrix of constituent asset returns is crucial in mean-variance optimisation as it fundamentally impacts the calculation of portfolio risk and the construction of the efficient frontier. In the mean-variance framework, the covariance matrix quantifies the variances and covariances of asset returns, providing a basis for assessing how asset combinations affect overall portfolio risk. Jobson and Korkie (1980) highlight that the covariance matrix is integral to deriving the weights of the Markowitz efficient portfolio. In their study, the sample covariance matrix $\boldsymbol{V}$ is estimated from 313 monthly return observations of 20 randomly selected NYSE listed stocks. Jobson and Korkie (1980) define the variance of a portfolio as $X'\Sigma X$, where X is the vector of asset weights. Therefore, the estimation of Σ directly affects the reliability of portfolio risk estimates in the mean-variance optimisation process.

A major issue with the estimation of the covariance matrix is sampling error, particularly in cases where there are many assets under consideration relative to the number of historical return observations available. As a result, the sample covariance matrix can become unstable and contain significant estimation error, leading to unreliable portfolio optimisation results. This phenomenon is well-documented in the literature. Michaud (1989) attributes this to the inherent nature of MV optimisers as "estimation-error maximisers." MV optimisation is highly sensitive to the input estimates of expected returns, variances, and covariances, which are inevitably subject to estimation errors. The optimisation process tends to significantly overweight (underweight) assets with high (low) expected returns, small (large) variances, and strong negative (positive) correlations, which are those most likely to have large estimation errors (Michaud, 1989).

Jobson and Korkie (1980) quantitatively depict this issue, demonstrating how estimation errors can severely degrade the performance of the "optimal" portfolio, defined as the portfolio on the efficient frontier providing the maximum Sharpe ratio. In their study, Monte Carlo simulations were used to optimise portfolios using estimated input data for a set of 20 stocks over a 60-month period. Their results showed a stark contrast between the true Sharpe ratios of the optimal portfolio derived from a known multivariate distribution, and that derived from the average of the simulated optimal portfolios. The true Sharpe ratio of the latter was significantly lower than that of both the true optimal portfolio and an equally weighted portfolio, suggesting that MV optimisation can indeed maximise estimation errors rather than minimise risk or maximise return in a meaningful way.

Michaud (1989) noted an important aspect of the Jobson-Korkie study, namely the absence of a short-selling constraint in their analysis, which might have exaggerated the error-maximisation effect. In practical investment management, where short-selling is often constrained, the impact of estimation errors might be somewhat mitigated. Michaud (1989) argues that the fundamental issue remains: the sensitivity of MV optimisation to estimation errors necessitates the careful consideration of imposing financially meaningful constraints on the procedure and the use of robust estimation techniques. Michaud (1989) cites "ill-conditioned" covariance matrices as a significant factor contributing to the instability of MV optimisations. Since MV optimisation requires the inversion of a covariance matrix, an ill-conditioned matrix – often due to input assumptions that do not reflect financially meaningful estimates or parameter estimates based on insufficient historical data – the resulting solutions can become highly unstable (Michaud, 1989). This instability reflects as large changes in portfolio allocations in response to small changes in input estimates, making the optimisation process unreliable. In contrast, highly structured estimators such as the single-factor model of Sharpe (1963) contain relatively little estimation error, but are often mis-specified and can be severely biased.

As a result, attempts have been made to find a compromise between the unbiased but noisy sample covariance matrix and a structured but potentially mis-specified estimator (Ledoit and Wolf, 2004b). One such method is termed "shrinkage estimation", which aims to provide a principled way to estimate large covariance matrices by optimally combining the sample covariance matrix with structured targets. First applied robustly to active portfolio management by Ledoit and Wolf (2003, 2004a, 2004b), shrinkage estimation originates from the seminal work of Stein (1956) and James and Stein (1961) whereby they demonstrated that shrinking the maximum likelihood estimator of a multivariate normal mean vector can reduce mean squared error.

Early efforts to incorporate shrinkage methods into portfolio selection were undertaken by Frost and Savarino (1986) and Jorion (1986).Their specific shrinkage techniques, however, struggled with situations where the number of stocks exceeded the number of historical return observations, a common occurrence in practice. Jagannathan and Ma (2003) demonstrated that mean-variance optimisers implicitly apply a form of shrinkage to the sample covariance matrix when short sales are restricted. This implicit shrinkage generally enhances the stability of portfolio weights. The explicit application of shrinkage techniques can potentially further improve the stability of portfolio weights, and risk and return parameters (Ledoit and Wolf, 2004b).

Ledoit and Wolf (2003, 2004a) first proposed shrinkage estimation by obtaining the best-weighted average of the sample covariance matrix of stock returns and a structured estimator, namely the single-index model covariance matrix, which is derived from Sharpe's (1963) single-index model. The key innovation lies in estimating $0 \leq \lambda \leq 1$, and determines the extent to which the sample covariance matrix is shrunk toward the structured estimator, or shrinkage target. This balances the trade-off between the unbiased but highly variable sample covariance matrix and the biased but less variable single-index model covariance matrix. This approach addresses two critical issues that plagued earlier methods: it avoids the problem of inverting an ill-conditioned covariance matrix and incorporates the correlation between estimation errors in both the sample covariance matrix and the shrinkage target, leading to a more accurate and stable estimator.

Using data from NYSE and AMEX stocks over the period 1972 to 1995, their findings indicate that portfolios constructed using their shrinkage estimator exhibit significantly lower out-of-sample variance compared to portfolios constructed using other established methods, including multifactor models. This improvement is particularly important, because it enhances the stability and reliability of the covariance matrix used in portfolio selection, a key component of Markowitz's (1952) mean-variance optimisation framework. The method developed by Ledoit and Wolf (2003, 2004a) reduces the dependence on multifactor models, which often suffer from unresolved issues regarding the number and identity of factors. Instead, their approach leverages the strong consensus around the market factor while providing a systematic way to account for extra-market covariances without imposing arbitrary factor structures.

This research was subsequently extended by Ledoit and Wolf (2004b), by substituting the single-factor matrix of Sharpe (1963) for a constant-correlation model as the shrinkage target. The constant-correlation model assumes that all pairwise correlations are the same, with the average of all the sample correlations serving as the estimate for the common constant correlation. This value, along with the vector of sample variances, determines the shrinkage target. The resulting shrinkage estimator is not only easy to implement but also performs comparably to the single-index model utilised in their previous studies. However, a notable caveat is that the constant-correlation model is not appropriate if the assets are from different asset classes, such as stocks and bonds.

Ledoit and Wolf's (2004b) empirical analysis, using historical U.S. stock market data, further demonstrates that the shrinkage estimator consistently outperforms the sample covariance matrix across various scenarios, resulting in higher information ratios and lower portfolio variances. Their method exhibits lower turnover compared to the sample covariance matrix, making it more attractive for active portfolio managers. Therefore, Ledoit and Wolf (2004b) assert that this shrinkage approach should replace the traditional sample covariance matrix in portfolio optimisation due to its superior performance and ease of implementation.

The intricacies of VCV matrix requirements for portfolio analysis, in the context of portfolio investment, have been discussed by Kwan (2010) who underscored the crucial role of a positive definite VCV matrix; a covariance matrix of asset returns must be positive definite to be suitable for portfolio analysis. A positive definite matrix is always invertible, although the reverse is not necessarily true. In portfolio investment, where variance is a measure of risk, Kwan (2010) asserts that an invertible covariance matrix is essential for portfolio selection models to generate portfolio allocation results. A positive definite covariance matrix guarantees strictly positive variances of portfolio returns, regardless of how funds are distributed among the assets. This characteristic not only ensures the presence of portfolio risk but also guarantees the uniqueness of efficient portfolio allocation outcomes (Kwan, 2010).

If the covariance matrix is estimated with insufficient data, it will not be positive definite. For instance, estimating a 100x100 covariance matrix of monthly asset returns requires more than 100 monthly return observations to ensure that the number of observed returns exceeds the number of unknown matrix elements (Kwan, 2010). Additional observations are thus necessary to minimise estimation errors and ensure the sample covariance matrix is a suitable input parameter for portfolio analysis. However, Kwan (2010) notes that using a longer sample period to obtain these additional observations raises concerns about the validity of the stationarity assumption of asset return distributions.

Kwan (2011) also asserts that relying on higher-frequency observations, such as weekly or daily asset returns within the same sample period, can mitigate concerns about non-stationary asset return distributions. However, this approach has drawbacks, as higher-frequency observations tend to be noisier, which can degrade the quality of the estimated covariance matrix. Although VCV matrix estimation accounting for time-varying volatility of asset returns can be achieved with multivariate time series models, such models are somewhat complicated (Silvennoinen and Teräsvirta, 2009).

Drawing parallels with life sciences and related fields, the idea that estimating VCV matrices with insufficient observations has been found to be a common challenge. As noted by Dobbin and Simon (2007) and Yao et al. (2008), biological studies often involve only a small number of observations due to limited experimental data or constraints related to budget and time. As a result, covariance matrices are often prone to estimation with insufficient observations. Schäfer and Strimmer (2005) also highlight the necessity of the estimated covariance matrix being invertible in gene association network studies using graphical Gaussian models (GGMs); the partial correlation between any two genes—which measures their degree of association after accounting for the effects of other genes—can be inferred from the inverse of a covariance matrix.

A cancer research study by Beerenwinkel et al. (2007) depicts the lack of a positive definite and invertible sample covariance matrix, pairwise partial correlations of 78 cancer genes (resulting in a total of 3003 partial correlations) being based on only 35 available tumour samples. The use of remedial measures by Schäfer and Strimmer (2005), Beerenwinkel et al. (2007), and Yao et al. (2008) to address the lack of experimental observations are based on the VCV shrinkage estimation approaches by Ledoit and Wolf (2003, 2004a, 2004b).

Kwan (2011) notes that these original Ledoit-Wolf shrinkage methods were not, however, designed to address issues arising from insufficient observations, prompting Schäfer and Strimmer (2005) to extend the Ledoit-Wolf approach to finite sample settings. In their study, they identified six potential shrinkage targets for covariance and correlation matrices, including an identity matrix, a covariance matrix based on the constant correlation model, a diagonal covariance matrix with sample variances on the diagonals, and three other related cases. Schäfer and Strimmer's shrinkage approach primarily focuses on the scenario where the target is a diagonal matrix. In this context, shrinkage estimation of the covariance matrix is relatively straightforward, both analytically and computationally; when variables are normalised to have unit variances, the same shrinkage approach becomes that for the correlation matrix instead (Kwan, 2011).

Complications arise since the normalisation relies on sample variances, which include estimation errors. To maintain the analytical integrity of the shrinkage estimation, Schäfer and Strimmer (2005) assume no estimation errors in the variances for normalised data. Building on this, Kwan (2008) extended the analytical framework by incorporating estimation errors in all variances when shrinking the sample correlation matrix with targets based on the average correlation of asset returns. This refinement implicitly improves the analytical expression of the Schäfer-Strimmer λ.

Kwan's (2011) subsequent study provides a foundational understanding of Ledoit and Wolf's (2003, 2004a, 2004b) shrinkage estimation techniques as applied to covariance matrices, emphasising their potential to enhance the reliability of sample covariance matrices. Kwan (2011) considers two primary versions of shrinkage estimation. The first version seeks an optimal weighted average between the sample covariance matrix and a diagonal matrix with elements corresponding to sample variances. The second involves the sample correlation matrix and an identity matrix, under certain simplifying assumptions. With a focus on the first approach, his study uses pedagogical tools, namely Excel-based demonstrations, to show how shrinkage estimation can improve the quality of the sample covariance matrix, especially when the number of historical observations is insufficient. These demonstrations help clarify how shrinkage estimation serves as a remedial measure for the instability often seen in sample covariance matrices subject to estimation error with limited data.

Kwan's (2017) paper extends the introduction to shrinkage estimation techniques from his earlier work (2011) by shifting the focus from a zero-correlation shrinkage target to a constant correlation shrinkage target in the context of the sample correlation matrix. This adjustment is practically and pedagogically significant, since a zero-correlation target consistently leads to attenuated individual correlations, resulting in a post-shrinkage average that is lower than the original sample average. The zero-correlation target approach lacks empirical support and is challenging to justify, especially from a pedagogic standpoint. Alternatively, a constant correlation target preserves the original sample average, thereby achieving a balance between reducing overall forecast errors and maintaining some of the unique characteristics of individual correlations (Kwan, 2017).

Kwan (2017) takes a pedagogical approach by simplifying the analysis, if any revisions to sample standard deviations of returns are based on the judgment of financial analysts. This assumption allows the focus to remain on shrinkage estimation of individual correlations, without the statistical complications associated with estimating variances and covariances as discussed in Kwan (2008). Kwan (2017) demonstrates an Excel-based computational approach that is designed to accommodate both the constant and zero correlation targets and addresses the statistical issue of bias in using sample correlations to estimate true but unknown correlations, as considered in Kwan (2008), offering users the option to either correct for this bias or ignore it. Given these choices in shrinkage target and bias correction, Kwan (2017) compares four sets of shrinkage results for a given set of Dow Jones stocks and their corresponding return observations. The results illustrate how the returns of US major stocks are correlated and how shrinkage estimation affects the sample correlation matrix of returns.

Ledoit and Wolf (2012) present an extension of their earlier work on covariance matrix estimation by introducing a nonlinear shrinkage estimator. The motivation for this approach stems from the oracle estimator of Ledoit and Péché (2011), which was derived by analysing the asymptotic behaviour of the optimal finite-sample rotation-equivariant estimator, particularly when measured using the Frobenius norm. In this study, Ledoit and Wolf (2012) apply a nonlinear transformation to the sample eigenvalues, a departure from the linear shrinkage approaches they previously developed in Ledoit and Wolf (2003, 2004a, 2004b).

The nonlinear shrinkage estimator they propose is designed to be asymptotically equivalent to the oracle estimator, and its development involves estimating the Stieltjes transform of the limiting spectral distribution of the sample eigenvalues. Through this approach, they construct a reliable nonlinear shrinkage estimator that demonstrates superior finite-sample properties in Monte Carlo simulations compared to both the sample covariance matrix and their earlier linear shrinkage estimator. The nonlinear estimator shows significant improvements, particularly in cases where the sample size is large relative to the matrix dimension or when the population eigenvalues exhibit high dispersion. However, the scope of the study is limited to cases where the matrix dimension is smaller than the sample size (Ledoit and Wolf, 2012).

Ledoit and Wolf (2015) develop an estimator for the population eigenvalues that remains consistent under large-dimensional asymptotics, regardless of whether these eigenvalues are clustered. This advancement is achieved by providing a more precise characterisation of the asymptotic behaviour of the sample eigenvalues. Unlike earlier results that only described the average behaviour of eigenvalues (i.e., how many fall within a given interval), this study determines the individual limits of the eigenvalues, leading to a more accurate and reliable estimation process. Ledoit and Wolf's (2015) consistent estimator improves covariance matrix estimation in scenarios where the dimension is large relative to the sample size. This extends the work of Ledoit and Wolf (2012), which was limited to cases where the matrix dimension was smaller than the sample size, allowing the authors to handle the more challenging case of high-dimensional settings, where the sample covariance matrix is singular.

Ledoit and Wolf (2020a) further advance the nonlinear shrinkage estimation framework by introducing a "direct", analytical method for estimating the oracle shrinkage function. This approach contrasts with their earlier "indirect" or numerical strategies from Ledoit and Wolf (2012, 2015), which required the estimation of the limiting distribution of population eigenvalues, a computationally intensive process involving large-scale optimisation. Ledoit and Wolf's (2020a) analytical method simplifies implementation, requiring minimal code and offering significant computational efficiency, making it feasible for large datasets with dimensions exceeding $N = 10\ 000$. In contrast, numerical strategies cannot accommodate dimensions more than around $N = 1\ 000$.

Ledoit and Wolf (2020a) highlight that while nonlinear shrinkage generally outperforms linear shrinkage, exceptions exist when the oracle shrinkage function is nearly linear or when both the matrix dimension N and sample size T are small. However, with N and T exceeding 50, nonlinear shrinkage becomes the preferred method, offering substantial performance gains over linear shrinkage. Ledoit and Wolf (2020a) effectively addresses previous limitations, making nonlinear shrinkage estimation of covariance matrices more accessible and practical for large-scale data applications.

Dom et al. (2024) critique the standard academic practice of evaluating variance-covariance (VCV) estimators using unconstrained global minimum variance (GMV) portfolios and their ex-post volatility. They argue that such portfolios are impractical due to high leverage, concentration, and turnover, which inflate transaction costs and diminish real-world applicability. The authors propose evaluating VCV estimators using more practical test portfolios, such as long-only GMV and risk-parity allocations, which better account for turnover, transaction costs, and diversification.

They reaffirm that constraints like long-only requirements inherently stabilise portfolio weights and reduce estimation errors, thereby diminishing the incremental benefits of complex shrinkage estimators (Jagannathan and Ma, 2003). While advanced methods, such as dynamic estimators (e.g., DCC-NLS) or factor models, can improve performance, simpler approaches like RiskMetrics combined with non-linear shrinkage achieve comparable results for constrained portfolios. The study concludes that practical portfolio designs and realistic evaluation metrics – balancing transaction costs, diversification, and estimation stability – are critical for accurately assessing the performance of VCV estimators, particularly in large asset universes.

3. Empirical design

We aim to evaluate select performance statistics of mean-variance efficient frontier portfolios constructed using sample variance-covariance matrices compared to those incorporating shrinkage-estimated matrices. These performance statistics involve portfolio annualised risk, annualised cumulative returns, and annualised Sharpe ratios. To evaluate the performance statistics of the shrinkage estimators under varying market conditions, a descriptive time series analysis is conducted. Portfolio performance statistics are back-tested over a recent out-of-sample period, capturing key economic environments, such as the post-COVID-19 period and the series of Federal Reserve rate hikes from Mar-22 – Jul-23. The observed relationships between shrinkage intensities and these measures will also be examined. Lastly, we illustrate the application of the different shrinkage estimators by comparing the eigenvalues of the shrunk and unshrunk (sample) covariance matrices. This demonstrates how these estimators fundamentally affect portfolio construction.

3.1 Construction of test portfolios for evaluating shrinkage estimators

Classical mean-variance portfolio optimisation, as introduced by Markowitz (1952), seeks to balance expected risk and return by leveraging the covariance matrix of asset returns to construct optimal portfolios. However, its practical application often suffers from estimation error in the sample covariance matrix, resulting in overly concentrated and unstable portfolios that may underperform in terms of risk-adjusted returns. Covariance matrix shrinkage estimation, which combines the sample covariance matrix with a structured target matrix, has been proposed as a remedy to reduce these errors and improve portfolio performance (Ledoit and Wolf, 2012, 2015, 2017, 2020a, 2020b).

In this study, we focus on two portfolios within the efficient frontier framework: the global minimum-variance (GMV) portfolio and the tangency (TG) portfolio. The GMV portfolio minimises total portfolio risk without considering expected returns, making it a suitable candidate for evaluating the effectiveness of shrinkage estimators in diversifying away risk. The TG portfolio, which lies tangent to the efficient frontier and aims to maximise the associated Sharpe ratio, is evaluated to assess whether shrinkage estimators can improve risk-adjusted returns. Lastly, to benchmark the performance of the efficient frontier portfolios, we follow Dom et al. (2024) and include two simple portfolio allocation strategies: the equally weighted (EW) portfolio and the inverse volatility (IV) portfolio.

It is important to note that we focus solely on simulating the performance of fully unconstrained portfolios. By limiting our analysis to an unconstrained setting, we aim to explore the effects of portfolio shrinkage in a simplified framework without adding complexities such as short-selling limitations, asset weighting constraints, and transaction costs. These elements are generally proprietary to each institution and, therefore, typically inaccessible. The use of an unconstrained framework may allow for a more foundational analysis of shrinkage behaviour and its impact on portfolio optimisation, by isolating these effects from constraints that could potentially obscure underlying patterns or relationships. Future work could extend these findings to constrained portfolio settings, incorporating practical considerations such as regulatory limits, long-only and other liquidity constraints, or position sizing. Such extensions would more closely reflect real-world portfolio management practices and would enable a more comprehensive evaluation of shrinkage techniques within applied contexts.

3.1.1 Efficient frontier portfolio construction

Following Jorion (2003), when constructing a mean-variance optimised portfolio comprising n stocks, the following variables are defined:

$\boldsymbol{q_P}$	=	vector of portfolio weights for the sample of n stocks
$\mathbf{E}$	=	vector of expected returns
$\mathbf{V}$	=	covariance matrix of asset returns

It is assumed that net short sales *are* allowed to preserve linearity.

Jorion (2003) states expected returns and variances in matrix notation as:

$\mu_P = \boldsymbol{q'_P}\mathbf{E}$	=	expected return on the portfolio	(1)
$\sigma_P^2 = \boldsymbol{q'_P}\mathbf{V}\boldsymbol{q_P}$	=	variance of portfolio return	(2)

The portfolio is also subject to a full-investment constraint, where the total portfolio weights must sum to unity:

$$\boldsymbol{q'_P}\mathbf{1} = 1$$

where $\mathbf{1}$ represents a vector of 1s.

Jorion (2003) derives the conventional mean-variance efficient frontier, in which there is no risk-free asset, where G is the target expected return. The portfolio allocation problem involves a constrained minimisation of the portfolio variance over the weights q_P:

Minimise $\boldsymbol{q'_P}\mathbf{V}\boldsymbol{q_P}$ subject to:

$\boldsymbol{q'_P}\mathbf{1} = 1$ = portfolio full-investment constraint

$\boldsymbol{q'_P}\mathbf{E} = G$ = target expected return

Following Merton (1972) and Jorion (2003), the efficient set constants are defined as:

$a = \mathbf{E}'\mathbf{V}^{-1}\mathbf{E}$;

$b = \mathbf{E}'\mathbf{V}^{-1}\mathbf{1}$;

$c = \mathbf{1}'\mathbf{V}^{-1}\mathbf{1}$;

$$d = a - \frac{b^2}{c}$$

The efficient frontier can then be fully defined by two portfolios, the GMV portfolio and the TG portfolio that maximises the return-to-risk ratio (Sharpe ratio). The respective weights are:

$$\boldsymbol{q_{GMV}} = \mathbf{V}^{-1}\frac{\mathbf{1}}{c}; \tag{3}$$

$$\boldsymbol{q_{TG}} = \mathbf{V}^{-1}\frac{\boldsymbol{E}}{b}; \tag{4}$$

The expected return, E, and variance, V, of these two portfolios are:

$$E_{GMV} = \frac{b}{c}; \tag{5}$$

$$V_{GMV} = \frac{1}{c}; \tag{6}$$

and

$$E_{TG} = \frac{a}{b}; \tag{7}$$

$$V_{TG} = \frac{a}{b^2}; \tag{8}$$

The efficient set is given by:

$$\sigma_P^2 = \frac{1}{d}(\mu_P - \mu_{MV})^2 + \sigma_{MV}^2 \tag{9}$$

which represents a parabola in the (σ^2, μ) space, or a hyperbola in the (σ, μ) space with asymptotes having a slope of $\pm\sqrt{d}$; this slope represents the best return-to-risk ratio for this set of stocks (Jorion, 2003).

3.1.2 Benchmark portfolios

The equally-weighted (EW) portfolio is a simple allocation strategy where each of the n assets in the portfolio receives an identical weight, regardless of its risk or return characteristics. The weights for each asset i are calculated as:

$$w_i^{EW} = \frac{1}{n} \qquad \forall i, \qquad \sum_{i=1}^{n} w_i^{EW} = 1$$

In addition to its simplicity and ease of implementation, the EW portfolio has been shown to exhibit surprisingly robust performance. As DeMiguel, Garlappi and Uppal (2009) demonstrate, this strategy can be difficult to outperform out-of-sample.

The inverse volatility (IV) portfolio assigns weights inversely proportional to individual asset volatilities, with lower weights assigned to riskier assets and higher weights to less volatile ones. For an asset i with standard deviation (volatility) σ_i, the weight is given by:

$$w_i^{IV} = \frac{\sigma_i^{-1}}{\sum_{i=1}^{n} \sigma_i^{-1}} \qquad \forall i, \qquad \sum_{i=1}^{n} w_i^{IV} = 1$$

This allocation strategy does not account for risk diversification benefits and can be viewed as a variant of a fully-invested GMV portfolio where all off-diagonal elements of the VCV matrix are equal to zero (Dom et al., 2024). As both EW and IV portfolios do not consider asset return correlations in formulating asset weights, they are unaffected by shrinkage estimators and thus serve only as benchmark portfolios.

3.2 Shrinkage estimation of the sample correlation matrix

Our study explores the approach of Kwan (2017), which follows the work of Ledoit and Wolf (2003, 2004a, 2004b). Four methods of linear shrinkage estimation are implemented. These involve shrinking the sample correlation matrix of returns towards both a zero-correlation target matrix and a constant correlation target matrix, with the statistical issue of bias either being ignored or corrected; this bias stems from the use of the sample correlations to estimate the corresponding true but unknown correlations.

Ledoit and Wolf (2020a) have shown that nonlinear shrinkage techniques generally dominate linear approaches, in which they capture a greater measure of the potential for variance reduction. However, the linear shrinkage methods employed by Kwan (2017) may still provide a foundation to investigate the general impact and effectiveness of VCV matrix shrinkage on mean-variance optimal portfolio performance statistics. This analysis may be extended in future research to include nonlinear approaches for a more comprehensive insight into the practical application of VCV matrix shrinkage techniques.

3.2.1 Shrinkage towards a constant correlation target

Following Kwan (2017), for an n-security portfolio, let $\boldsymbol{R_1}, \boldsymbol{R_2}, \dots, \boldsymbol{R_n}$ be the random returns of the individual securities. Let also $\boldsymbol{R_{i1}}, \boldsymbol{R_{i2}}, \dots, \boldsymbol{R_{iT}}$ be the returns observed at times $1,2,\dots,T$, respectively, for $i = 1,2,\dots,n$. Let the sample correlation of returns between securities $\boldsymbol{i}$ and $\boldsymbol{j}$, for $i, j = 1,2,\dots,n$ be denoted by $\boldsymbol{r_{ij}}$; the average thereof is denoted by $\boldsymbol{\bar{r}}$. The true correlation between $\boldsymbol{R_i}$ and $\boldsymbol{R_j}$, for $i, j = 1,2,\dots,n$ and $i \neq j$, denoted by $\boldsymbol{\rho_{ij}}$, is unknown.

The constant-correlation model states that all the (pairwise) correlations are identical (Ledoit and Wolf, 2004b). For a constant correlation shrinkage target, the constant correlation is provided by the average of the individual sample correlations, $\boldsymbol{\bar{r}}$. Let the shrinkage intensity (constant) applied to the sample correlation matrix be denoted as $\boldsymbol{\lambda}$, which is the weight assigned to $\boldsymbol{\bar{r}}$. As a result, the departure of the resulting correlation between $\boldsymbol{R_i}$ and $\boldsymbol{R_j}$ from the corresponding true correlation is $\boldsymbol{\lambda\bar{r} + (1-\lambda)r_{ij} - \rho_{ij}}$ (Kwan, 2017).

Following Ledoit and Wolf (2004b) and Kwan (2017), we find the value of $\boldsymbol{\lambda}$ that minimises the expected value of a quadratic loss function, given by (10):

$$L(\lambda) = E\left\{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left[\lambda\bar{r} + (1-\lambda)r_{ij} - \rho_{ij}\right]^2\right\} \tag{10}$$

where $\boldsymbol{E}(\cdot)$ is the expected value operator. After simplification of (10), the optimal shrinkage intensity is given by (11):

$$\lambda = \frac{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left\{\alpha_{ij} - E(\bar{r} - r_{ij})\left[E(r_{ij}) - \rho_{ij}\right]\right\}}{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left\{\alpha_{ij} + \left[E(\bar{r} - r_{ij})\right]^2\right\}} \tag{11}$$

where

$$\alpha_{ij} = Var(r_{ij}) - \frac{2}{n(n-1)}\sum_{k=1}^{n-1}\sum_{\ell=k+1}^{n} Cov(r_{ij}, r_{k\ell}) \tag{12}$$

In estimating $\boldsymbol{\lambda}$, let $\boldsymbol{E}(\boldsymbol{r_{ij}})$ be an estimator of $\boldsymbol{\rho_{ij}}$ for analytical convenience. Therefore, with $\boldsymbol{E}(\boldsymbol{r_{ij}}) - \boldsymbol{\rho_{ij}}$ being zero, and using $\boldsymbol{r_{ij}}$ and $\bar{\boldsymbol{r}}$ for $\boldsymbol{E}(\boldsymbol{r_{ij}})$ and $\boldsymbol{E}(\bar{\boldsymbol{r}})$, respectively, the estimated $\boldsymbol{\lambda}$ is given by (13):

$$\hat{\lambda} = \frac{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\hat{\alpha}_{ij}}{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left[\hat{\alpha}_{ij} + \left(\bar{r} - r_{ij}\right)^2\right]} \tag{13}$$

where

$$\hat{\alpha}_{ij} = \widehat{Var}\left(r_{ij}\right) - \frac{2}{n(n-1)}\sum_{k=1}^{n-1}\sum_{\ell=k+1}^{n}\widehat{Cov}\left(r_{ij}, r_{k\ell}\right) \tag{14}$$

For a finite sample, $\boldsymbol{E}(\boldsymbol{r_{ij}})$ is inherently a biased estimator of $\boldsymbol{\rho_{ij}}$; the sample correlation tends to understate the true correlation in magnitude, with the bias attenuating as the sample periods lengthens. However, utilising a longer sample period is susceptible to violations of the stationary assumption of security return distributions (Kwan, 2017).

Following Zimmerman, Zumbo and Williams (2003), (15) gives an estimator of $\boldsymbol{\rho_{ij}}$, proposed by Olkin and Pratt (1958), which can eliminate most of the bias in the sample correlation $\boldsymbol{r_{ij}}$ of normally distributed variables i and j:

$$\hat{\rho}_{ij} = r_{ij}\left[1 + \frac{1 - r_{ij}^2}{2(T-3)}\right] \tag{15}$$

Subsequently, (16) accounts for the bias in the estimation of $\boldsymbol{\rho_{ij}}$:

$$\hat{\lambda} = \frac{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left(\hat{\alpha}_{ij} - \hat{\gamma}_{ij}\right)}{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left[\hat{\alpha}_{ij} + \left(\bar{r} - r_{ij}\right)^2\right]} \tag{16}$$

where

$$\hat{\gamma}_{ij} = \left(r_{ij} - \bar{r}\right)\left[\frac{r_{ij}\left(1 - r_{ij}^2\right)}{2(T-3)}\right] \tag{17}$$

3.2.2 Shrinkage towards a zero-correlation target

When considering a zero-correlation shrinkage target, namely an identity matrix, $\bar{\boldsymbol{r}}$ is substituted by zero (Kwan, 2017). As a result, if the bias in the estimation of $\boldsymbol{\rho_{ij}}$ with $\boldsymbol{E}(\boldsymbol{r_{ij}})$ is ignored, the estimated optimal λ is given by (18):

$$\hat{\lambda} = \frac{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\widehat{Var}(r_{ij})}{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left[\widehat{Var}\left(r_{ij}\right) + r_{ij}^2\right]} \tag{18}$$

Alternatively, if bias is corrected for, $\hat{\boldsymbol{\lambda}}$ is given by (10):

$$\hat{\lambda} = \frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left[\widehat{Var}(r_{ij}) - \hat{\gamma}_{ij}\right]}{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left[\widehat{Var}(r_{ij}) + r_{ij}^2\right]} \tag{19}$$

Overall, four linear shrinkage estimators are considered, which are summarised in Table 1.

Table 1: Covariance matrix linear shrinkage estimators.

Shrinkage target	Bias	Optimal $\hat{\lambda}$ estimate
Constant correlation target	Ignored	$\frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \hat{\alpha}_{ij}}{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left[\hat{\alpha}_{ij} + (\bar{r} - r_{ij})^2\right]}$
Constant correlation target	Corrected	$\frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (\hat{\alpha}_{ij} - \hat{\gamma}_{ij})}{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left[\hat{\alpha}_{ij} + (\bar{r} - r_{ij})^2\right]}$
Zero correlation target	Ignored	$\frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \widehat{Var}(r_{ij})}{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left[\widehat{Var}(r_{ij}) + r_{ij}^2\right]}$
Zero correlation target	Corrected	$\frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left[\widehat{Var}(r_{ij}) - \hat{\gamma}_{ij}\right]}{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left[\widehat{Var}(r_{ij}) + r_{ij}^2\right]}$

Ledoit and Wolf (2004b) define the best weighted average of the sample covariance matrix $\boldsymbol{S}$ and a highly structured estimator, or shrinkage target matrix, $\boldsymbol{F}$, as:

$$\lambda F + (1 - \lambda) S \tag{20}$$

Following Kwan (2017), after linearly transforming the return observations to make the resulting sample covariance and correlation matrices indistinguishable from each other, (20) may be rewritten as (21), which gives the post-shrinkage correlation of returns between securities $\boldsymbol{i}$ and $\boldsymbol{j}$:

$$\lambda \bar{r} + (1 - \lambda) r_{ij} \tag{21}$$

In addition, we also include the average of sample correlations as a predictor of individual asset correlations. This approach serves as an additional benchmark for the four linear shrinkage estimators. Empirical support for this averaging method has been provided by Elton and Gruber (1973), Elton, Gruber, and Urich (1978) and Chan, Karceski and Lakonishok (1999), who demonstrate that averaging can effectively reduce overall forecast errors and provide a more stable correlation matrix. However, it may lead to the loss of idiosyncrasies in individual asset correlations.

3.3 Data and descriptive statistics

The data for this study consist of weekly adjusted closing prices of selected equities, and US 3-month Treasury bill rates, sourced from Bloomberg. Treasury bill rates are used as the risk-free rate to calculate excess returns. The focus is on the US, a developed economy, and the full sample period spans from 1-Jan-18 - 20-May-24. This period captures approximately one business cycle and includes significant economic events, notably the COVID-19 market turbulence of 2020-2021, which resulted in heightened volatility and uncertainty, and the subsequent period of aggressive interest rate hikes aimed at controlling inflation (vom Lehn and Winberry, 2021). These economic shocks are critical as they provide a comprehensive timeframe for analysing portfolio performance under varied market conditions.

For this analysis, weekly price data were selected to align with the practicalities of portfolio management. In real-world settings, portfolio managers generally avoid the costs and complexity associated with daily rebalancing, making weekly data a more realistic reflection of their decision-making frequency. Weekly data capture sufficient detail to observe meaningful price changes and correlations without the excessive noise and transaction costs associated with daily frequency.

The study also analyses hypothetical portfolios consisting only of stocks. This is due to the use of the constant correlation model as a shrinkage target, which assumes that all pairwise correlations among the assets are identical. This assumption limits the constant correlation model's applicability, as it would not be appropriate for portfolios containing assets from different asset classes, such as stocks and bonds, which typically exhibit varied correlation patterns. For portfolios that span multiple asset classes, more general shrinkage targets are available to account for these differences in inter-asset correlations.

To create a diversified portfolio reflective of the broader US market, we select 25 large-cap, highly liquid stocks from each of the 11 sectors of the S&P 500 index (Table 2). This selection ensures that the portfolio benefits from sectoral diversification, which is essential for accurately assessing the impact of covariance shrinkage on portfolio risk. Large-cap, liquid stocks are chosen as these securities typically exhibit more stable price behaviours and are less susceptible to the liquidity constraints that might affect smaller-cap stocks.

VCV matrices are estimated using a moving window approach. Given the use of weekly data in our study, we employ a two-year rolling window, corresponding to 104 weekly observations, for each estimation. The in-sample period is thus defined by the period from 01-Jan-18 – 23-Dec-19. While a shorter one-year window (52 weeks) was considered, it was deemed too limited to generate meaningful efficient frontier results. Conversely, a three-year window (156 weeks) would provide a more stable covariance matrix but would significantly reduce the size of the usable dataset, making it impractical for the purposes of this study. For each week in the out-of-sample period from 30-Dec-19 – 20-May-24, the VCV matrix is estimated based on the 104 weeks preceding that time point. Portfolios are rebalanced weekly, with efficient frontiers and corresponding portfolio weights, risk, and returns recalculated for each week. Shrinkage estimators are then applied to the same two-year rolling window to assess how they would alter out-of-sample portfolio characteristics compared to those derived from the sample covariance matrix.

Table 2 illustrates data used in the study, whereas Table 3 outlines the individual stocks within each constituent sector of the S&P 500 Index.

Table 2: Weekly price data used, sourced from Bloomberg™.

Data	Quantity
S&P 500 index listed stocks, represented by constituent sectors	**25**
Financial Services	4
Technology	3
Communication Services	2
Consumer Defensive	2
Consumer Cyclical	2
Healthcare	2
Utilities	2
Industrials	2
Energy	2
Basic Materials	2
Real Estate	2
US 3-month Treasury Yields	-

Table 3: List of individual stocks and respective constituent sectors of S&P 500 Index.

Stock	Description	Sector
JPM	JPMorgan Chase & Co.	Financial Services
GS	The Goldman Sachs Group, Inc.	Financial Services
BAC	Bank of America Corporation	Financial Services
WFC	Wells Fargo & Company	Financial Services
AAPL	Apple Inc.	Technology
MSFT	Microsoft Corporation	Technology
NVDA	Nvidia Corporation	Technology
GOOGL	Alphabet Inc.	Communication Services
NFLX	Netflix, Inc.	Communication Services
WMT	Walmart Inc.	Consumer Defensive
PG	The Procter & Gamble Company	Consumer Defensive
AMZN	Amazon.com, Inc.	Consumer Cyclical
HD	The Home Depot, Inc.	Consumer Cyclical
UNH	UnitedHealth Group Incorporated	Healthcare
JNJ	Johnson & Johnson	Healthcare
NEE	NextEra Energy, Inc.	Utilities
AWK	American Water Works Company, Inc.	Utilities
CAT	Caterpillar Inc.	Industrials
RTX	RTX Corporation	Industrials
XOM	Exxon Mobil Corporation	Energy
SLB	Schlumberger Limited	Energy
SHW	The Sherwin-Williams Company	Basic Materials
SCCO	Southern Copper Corporation	Basic Materials
PLD	Prologis, Inc.	Real Estate
SPG	Simon Property Group, Inc.	Real Estate

Figures 1-6 below illustrate the time series of weekly adjusted closing prices, rebased at $100, for each of the 25 stocks under consideration from 01-Jan-18 – 20-May-24, grouped by their respective S&P 500 Index sectors.

Figure 1: Time series of weekly price data for Financial Services stocks.

The four financial services stocks (JPM, GS, BAC, and WFC) demonstrate relatively similar patterns in response to economic events over the observed period. By rebasing at $100, we can more directly observe differences in growth trajectories and volatility.

Most notably, all four stocks show a significant dip in early 2020, correlating with the COVID-19 pandemic's economic impact, which disrupted financial markets globally. Pre-COVID, all four stocks had similar trajectories with moderate fluctuations, though Goldman Sachs (GS) experienced more pronounced dips, suggesting higher sensitivity to market fluctuations relative to the other banks. The onset of the pandemic in early 2020 induced a sharp decline across all stocks, particularly for Wells Fargo (WFC), which fell significantly lower than its peers. This discrepancy suggests WFC had greater vulnerability to the economic shocks induced by the crisis, potentially due to its specific business structure or exposure to riskier assets.

As interest rates hikes began post-2021, banks generally benefited due to improved interest margins. While all four stocks rebounded post-COVID, JPMorgan (JPM) and GS displayed the strongest recovery and growth, diverging from Bank of America (BAC) and WFC, which followed more volatile paths. This trend might indicate that JPM and GS have business models or asset structures that made them more resilient in the post-pandemic market, while BAC and WFC were slower to stabilise.

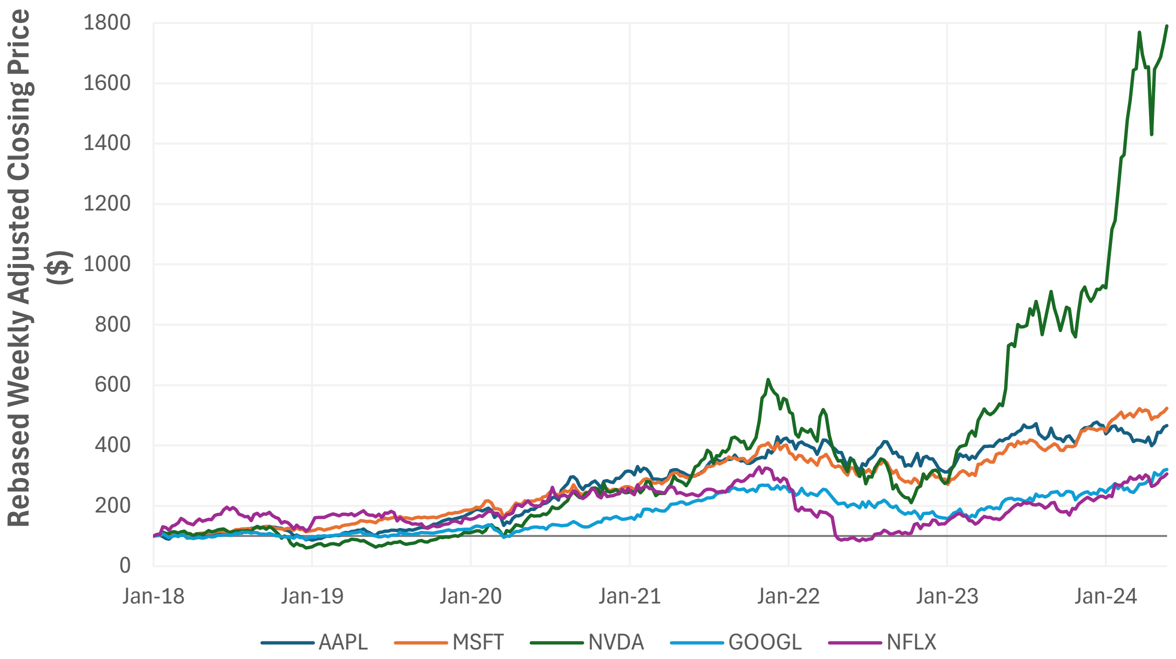


Figure 2: Time series of weekly price data for Technology and Communication Services stocks.

Like the financial services stocks, technology (AAPL, MSFT, and NVDA) and communication services (GOOGL and NFLX) stocks experienced a downturn during COVID-19, but the rebound was stronger and quicker, showing resilience and high growth potential.

Nvidia (NVDA) stands out as the most volatile and highest-growth stock in this group, especially from 2021 onward, largely driven by its focus on semiconductors and artificial intelligence (AI), areas that saw increased demand and investor interest. This high growth, however, also came with substantial volatility, indicating a higher risk profile compared to more stable giants like Microsoft (MSFT) and Apple (AAPL).

AAPL, MSFT, and Alphabet (GOOGL) followed relatively similar trajectories, recovering quickly after the COVID-induced dip and showing steady growth. Netflix (NFLX), however, shows a more distinct path with pronounced peaks and troughs, reflecting its cyclical revenue model tied to consumer demand for streaming services, which faced more industry competition and volatility post-pandemic. NVDA's dramatic surge from 2023 aligns with heightened interest in AI technologies and advanced computing solutions, which has driven its stock price up much faster than the others. This level of growth, while beneficial for portfolio returns, also suggests increased exposure to sector-specific risks, as NVDA's performance is highly sensitive to shifts in technology demand and market speculation.

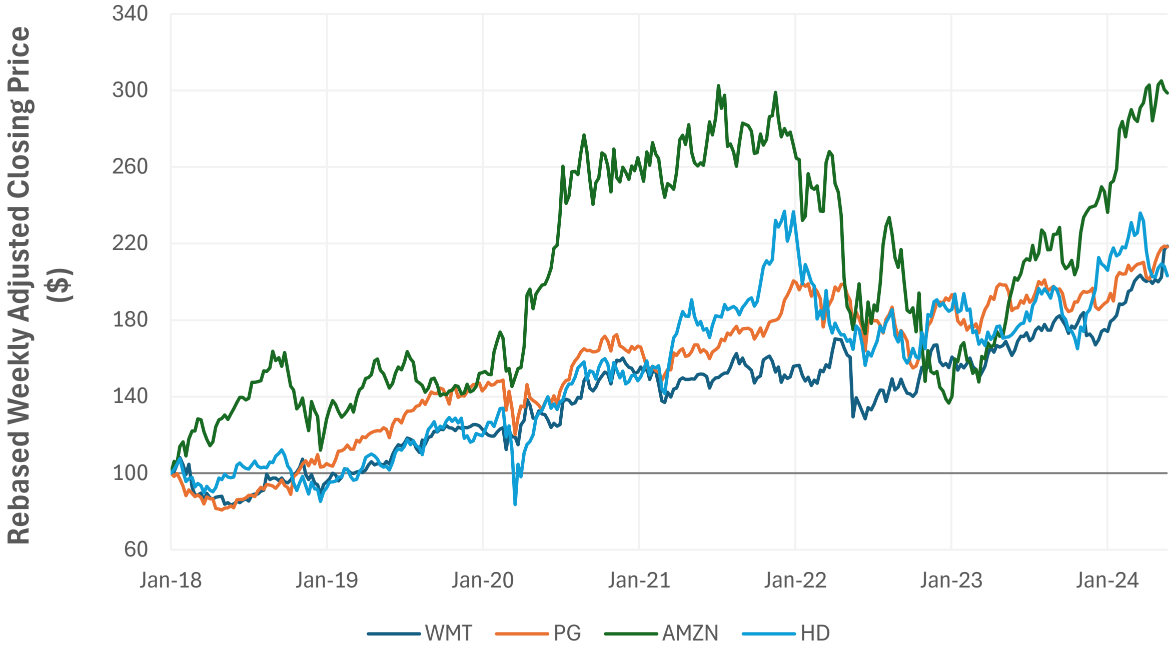


Figure 3: Time series of weekly price data for Consumer Defensive and Consumer Cyclical stocks.

Consumer defensive (WMT and PG) stocks exhibit relatively steady growth throughout the sample period compared to the more fluctuating growth in consumer cyclical stocks (AMZN and HD). Walmart (WMT) and Procter & Gamble (PG) display moderate, steady growth with limited volatility, aligning with the defensive nature of these stocks. In contrast, Amazon (AMZN) and Home Depot (HD) have seen greater volatility, reflecting their exposure to consumer spending trends. AMZN, in particular, experienced substantial growth pre-COVID, followed by a highly volatile period as the pandemic impacted consumer spending habits and competitive pressures grew within the e-commerce space. Meanwhile, HD's growth remained steadier, aligning with trends in home improvement demand, which has shown resilience despite macroeconomic pressures.

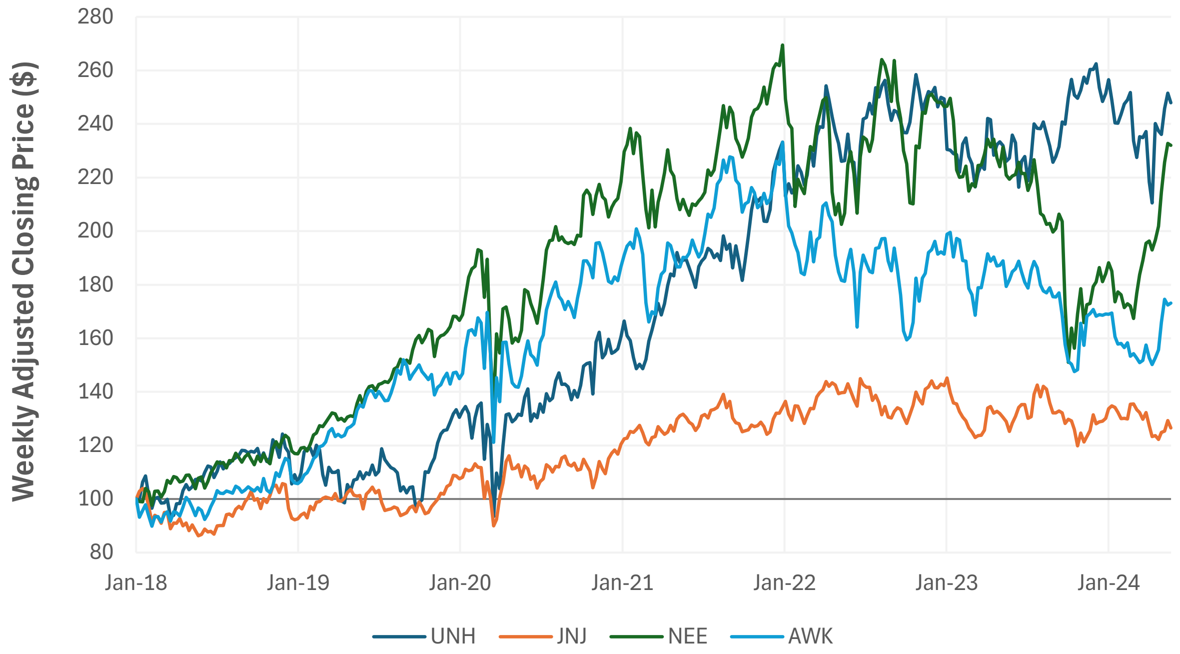


Figure 4: Time series of weekly price data for Healthcare and Utilities stocks.

The stocks in healthcare (UNH and JNJ) and utilities (NEE and AWK) illustrate different growth dynamics, reflecting their sectors' distinct responses to market changes.

From 2018 onward, UnitedHealth Group (UNH) shows the highest growth among the four stocks, suggesting strong demand and performance in the healthcare sector. UNH's trajectory is relatively stable until 2021, after which it shows a sharp upward trend, likely benefiting from increased attention on healthcare post-COVID. By the end of the period, UNH has grown significantly, indicating resilience and strong long-term growth.

NextEra Energy (NEE) also exhibits strong growth, but with more fluctuations. NEE's performance is characterised by moderate volatility, especially around 2019 and early 2020, possibly reflecting shifts in energy policy or sectoral investment patterns. After a dip in early 2020, likely linked to pandemic-induced market shocks, NEE recovers and continues its upward trend, suggesting renewed interest in clean energy investments.

American Water Works (AWK) and Johnson & Johnson (JNJ) in particular follow relatively more stable but lower growth trajectories, aligning with their defensive sector profiles. JNJ, a large healthcare company, displays consistent growth with minimal volatility, especially in comparison to UNH. This pattern is typical for large-cap healthcare firms with diversified product lines. AWK, as a water utility, experiences steady, moderate growth, maintaining a low volatility profile that is characteristic of essential services providers.

Figure 5: Time series of weekly price data for Industrials and Energy stocks.

The industrial (CAT and RTX) and energy (XOM and SLB) stocks display distinct recovery trajectories after the significant decline caused by the COVID-19 pandemic in early 2020. Caterpillar (CAT), a bellwether for the global construction and mining equipment industries, exhibits steady and robust growth post-2020. This is likely tied to increased infrastructure spending worldwide and the post-pandemic rebound in construction activity. CAT's resilience and long-term outperformance could also be attributed to rising demand for heavy machinery, driven by government-funded projects and private-sector investments.

Raytheon Technologies (RTX), by contrast, has a more subdued recovery. Its reliance on the aerospace sector, one of the hardest-hit industries during the pandemic due to grounded flights and diminished air travel, likely contributed to its weaker performance. However, as aviation demand has recovered and geopolitical tensions have spurred demand for defence equipment, RTX has seen steady improvement since 2022. This dual exposure to civil aviation and defence markets offers RTX a diversified revenue base, though it has taken longer for both arms of its business to regain full momentum.

In the energy sector, Exxon Mobil (XOM) and Schlumberger (SLB) both benefit from the rebound in oil and gas prices following the pandemic's initial economic disruptions. Exxon Mobil's strong performance from 2021 onwards reflects its integration across the energy supply chain, with upstream exploration and downstream refining activities benefiting from rising crude oil prices. Its growth plateaus somewhat in 2023, potentially due to fluctuating energy prices and broader concerns about slowing global demand. Schlumberger (SLB), as an oilfield services company, exhibits a similar recovery pattern but with greater volatility. This may be indicative of its reliance on exploration and production activities, which are more sensitive to cyclical and geopolitical factors. SLB's underperformance relative to XOM may also suggest that it remains more exposed to capital spending constraints and regional instability in energy-producing regions.

Figure 6: Time series of weekly price data for Basic Materials and Real Estate stocks.

Lastly, the basic materials and real estate sectors' stocks display a varied set of behaviours influenced by sector-specific drivers. In the materials group, Sherwin-Williams (SHW) experiences consistent, steady growth, reflecting its market position as a provider of essential construction and consumer products such as paints and coatings, even amid economic slowdowns. Southern Copper (SCCO), however, stands out with its exceptional growth, driven by rising copper prices. Copper has seen structural demand growth due to its critical role in renewable energy infrastructure, electric vehicles, and electrification projects globally. SCCO's heightened volatility underscores its sensitivity to commodity price fluctuations, geopolitical risks (e.g., mining policy changes), and macroeconomic uncertainty.

In the real estate sector, Prologis (PLD) and Simon Property Group (SPG) exhibit contrasting performance. Prologis, as a logistics-focused real estate investment trust (REIT), benefits from the e-commerce boom that followed the pandemic, driving demand for warehouses and distribution centres. Its performance post-2020 reflects this favourable structural trend, although rising interest rates in 2022 begin to moderate its growth, highlighting the sensitivity of REITs to financing costs. On the other hand, SPG, which focuses on retail properties such as malls, struggles more significantly. Its slower recovery reflects the ongoing challenges in the retail property market, particularly as consumers increasingly shift toward online shopping. While SPG stabilises in recent years, its growth lags behind PLD, reflecting the structural challenges facing traditional retail real estate.

Table 4. Descriptive statistics of weekly stock-level returns. 01-Jan-18 – 20-May-24.

Stock	Mean (%)	Max. (%)	Min. (%)	Std. Dev. (%)	Skewness	Kurtosis	Jarque-Bera
JPM	0.32	23.53	−19.64	4.05	0.14	5.07	358.18 **
GS	0.33	25.41	−21.88	4.51	0.32	4.32	264.94 **
BAC	0.23	24.11	−18.58	4.63	0.19	3.23	146.28 **
WFC	0.17	26.57	−16.72	4.95	0.26	3.66	190.00 **
AAPL	0.54	14.73	−17.53	4.02	−0.16	1.86	49.29 **
MSFT	0.56	12.41	−13.52	3.42	0.04	1.26	22.11 **
NVDA	1.07	24.57	−20.05	6.43	0.13	0.82	10.13 **
GOOGL	0.43	14.31	−12.03	4.06	0.18	0.72	8.95 *
NFLX	0.52	25.90	−36.82	6.06	−0.30	6.05	512.32 **
WMT	0.28	9.62	−19.49	2.95	−0.84	6.98	715.45 **
PG	0.27	10.42	−10.63	2.69	−0.22	2.42	84.13 **
AMZN	0.43	16.26	−13.90	4.41	−0.02	1.21	20.40 **
HD	0.29	25.24	−25.53	4.03	−0.15	8.84	1085.71 **
UNH	0.36	17.36	−23.73	4.06	−0.30	5.61	441.05 **
JNJ	0.10	8.94	−10.72	2.62	−0.23	1.66	41.31 **
NEE	0.33	20.96	−18.49	3.81	−0.36	5.53	431.87 **
AWK	0.24	19.94	−21.51	3.76	0.14	6.63	610.44 **
CAT	0.39	16.08	−17.93	4.47	−0.11	0.93	12.82 **
RTX	0.22	29.60	−21.70	4.47	0.23	9.01	1130.23 **
XOM	0.29	16.74	−20.07	4.42	−0.10	2.77	107.34 **
SLB	0.13	19.57	−32.24	6.33	−0.12	2.79	108.59 **
SHW	0.33	17.17	−17.35	3.87	0.07	2.63	95.88 **
SCCO	0.51	19.38	−16.93	5.17	0.10	0.86	10.89 **
PLD	0.31	26.03	−18.60	4.28	0.62	8.10	932.15 **
SPG	0.29	53.55	−46.37	6.82	1.35	23.59	7823.79 **

Note: ** denotes significance at the 1% level. * denotes significance at the 5% level
Output using R.

Table 4 shows descriptive statistics of the weekly stock-level returns over the full sample period, with a few general trends being observed across the 25 stocks. The mean weekly returns show variation across the sample, ranging from as low as 0.10% for JNJ to as high as 1.07% for NVDA. This highlights differences in long-term performance, likely driven by sector-specific dynamics, firm strategies, and broader market trends. Despite these differences, most stocks exhibit positive mean returns, reflecting the general upward trajectory of equity markets during the sample period, which included notable growth phases despite intermittent periods of volatility.

The maximum and minimum weekly returns reveal substantial heterogeneity in the magnitude of price swings. Maximum returns range from the low single digits, such as JNJ's (8.94%), to extreme surges, such as SPG's (53.55%). Similarly, minimum returns range from moderate declines, like PG's (-10.63%), to significant crashes, such as NFLX's (-36.82%). These variations suggest that some stocks, particularly those in cyclical or high-growth sectors, are far more prone to dramatic weekly price movements, likely reflecting heightened investor expectations or sensitivity to macroeconomic shocks.

Weekly volatility, as measured by standard deviation, further supports these observations. Stocks like NVDA (6.43%) and SPG (6.82%) display the highest levels of volatility, consistent with their higher susceptibility to speculative trading or sector-specific risks. Conversely, stocks such as JNJ (2.62%) and PG (2.69%) exhibit lower volatility, reinforcing their status as relatively stable, defensive assets that provide consistency during turbulent market conditions.

Most stocks show skewness values near zero, indicating a relative balance between upward and downward extremes. However, notable deviations occur. Defensive stocks such as WMT (-0.84) and UNH (-0.30) exhibit negative skewness, suggesting frequent small gains relative to downside risks, whereas SPG (1.35) displays positive skewness, indicating a greater frequency of large positive returns. The kurtosis values are overwhelmingly leptokurtic, highlighting the frequent occurrence of extreme returns compared to a normal distribution. Stocks like SPG (23.59) and RTX (9.01) have notably high kurtosis, suggesting their susceptibility to significant outliers.

Lastly, the Jarque-Bera (JB) statistic complement the observations on skewness and kurtosis, indicating that none of the stocks in the sample follow a normal return distribution. Nearly all stocks exhibit JB statistics significant at the 1% level, except for GOOGL, which is significant at the 5% level. This reinforces the general observation that extreme returns and asymmetric distributions are pervasive, with deviations from normality largely driven by the high kurtosis and occasional skewness observed across the sample.

The correlation matrices for weekly stock returns show notable shifts in inter-stock relationships between the pre-pandemic and pandemic periods, outlined in Tables 5 and 6.

Table 5. Sample weekly stock-level return correlations over the pre-pandemic period 01-Jan-8 – 30-Dec-19 (104 weeks).

Sample ρ	*JPM*	*GS*	*BAC*	*WFC*	*AAPL*	*MSFT*	*NVDA*	*GOOGL*	*NFLX*	*WMT*	*PG*	*AMZN*	*HD*
JPM	**1**												
GS	0.82	**1**											
BAC	0.91	0.85	**1**										
WFC	0.69	0.67	0.74	**1**									
AAPL	0.38	0.39	0.33	0.34	**1**								
MSFT	0.58	0.53	0.57	0.52	0.52	**1**							
NVDA	0.49	0.58	0.53	0.36	0.55	0.57	**1**						
GOOGL	0.57	0.50	0.53	0.55	0.43	0.71	0.46	**1**					
NFLX	0.34	0.42	0.34	0.37	0.28	0.56	0.57	0.51	**1**				
WMT	0.37	0.45	0.38	0.41	0.35	0.37	0.31	0.24	0.15	**1**			
PG	0.38	0.35	0.26	0.34	0.37	0.39	0.13	0.24	0.06	0.37	**1**		
AMZN	0.52	0.51	0.50	0.41	0.44	0.71	0.65	0.61	0.63	0.35	0.14	**1**	
HD	0.47	0.41	0.45	0.40	0.33	0.50	0.52	0.46	0.37	0.40	0.24	0.54	**1**
UNH	0.40	0.36	0.42	0.38	0.28	0.51	0.33	0.42	0.33	0.34	0.26	0.48	0.37
JNJ	0.50	0.44	0.42	0.49	0.34	0.43	0.21	0.33	0.27	0.30	0.51	0.27	0.29
NEE	0.07	0.05	0.03	0.18	0.16	0.27	0.02	0.11	0.17	0.24	0.35	0.19	0.18
AWK	0.05	0.08	0.01	0.24	0.15	0.35	−0.01	0.21	0.19	0.22	0.43	0.19	0.16
CAT	0.75	0.63	0.71	0.55	0.39	0.55	0.58	0.51	0.37	0.31	0.26	0.50	0.56
RTX	0.68	0.62	0.61	0.54	0.37	0.56	0.40	0.44	0.36	0.22	0.40	0.42	0.47
XOM	0.58	0.59	0.49	0.56	0.34	0.53	0.45	0.51	0.43	0.30	0.39	0.51	0.52
SLB	0.61	0.61	0.62	0.47	0.33	0.45	0.51	0.48	0.42	0.24	0.14	0.48	0.33
SHW	0.58	0.47	0.52	0.50	0.39	0.48	0.38	0.46	0.34	0.24	0.41	0.33	0.45
SCCO	0.49	0.48	0.45	0.39	0.41	0.40	0.43	0.35	0.32	0.17	0.24	0.33	0.37
PLD	0.37	0.35	0.33	0.40	0.37	0.50	0.24	0.37	0.29	0.39	0.43	0.41	0.34
SPG	0.25	0.23	0.25	0.24	0.34	0.33	0.20	0.21	0.18	0.21	0.40	0.27	0.41

Sample ρ	*UNH*	*JNJ*	*NEE*	*AWK*	*CAT*	*RTX*	*XOM*	*SLB*	*SHW*	*SCCO*	*PLD*	*SPG*
UNH	**1**											
JNJ	0.36	**1**										
NEE	0.23	0.34	**1**									
AWK	0.31	0.37	0.72	**1**								
CAT	0.37	0.38	−0.02	−0.01	**1**							
RTX	0.34	0.45	0.16	0.15	0.61	**1**						
XOM	0.40	0.55	0.18	0.19	0.60	0.61	**1**					
SLB	0.32	0.36	−0.03	−0.10	0.58	0.58	0.66	**1**				
SHW	0.33	0.47	0.14	0.22	0.54	0.54	0.44	0.33	**1**			
SCCO	0.18	0.30	−0.04	−0.02	0.62	0.43	0.51	0.50	0.34	**1**		
PLD	0.35	0.47	0.46	0.49	0.25	0.52	0.39	0.26	0.38	0.24	**1**	
SPG	0.28	0.27	0.31	0.27	0.13	0.37	0.24	0.18	0.17	0.12	0.50	**1**

Output using R.

Table 6. Sample weekly stock-level return correlations over the pandemic period 06-Jan-20 – 27-Dec-21 (104 weeks).

Sample ρ	*JPM*	*GS*	*BAC*	*WFC*	*AAPL*	*MSFT*	*NVDA*	*GOOGL*	*NFLX*	*WMT*	*PG*	*AMZN*	*HD*
JPM	**1**												
GS	0.91	**1**											
BAC	0.95	0.87	**1**										
WFC	0.85	0.79	0.86	**1**									
AAPL	0.37	0.49	0.36	0.23	**1**								
MSFT	0.35	0.45	0.37	0.26	0.74	**1**							
NVDA	0.24	0.39	0.26	0.23	0.59	0.67	**1**						
GOOGL	0.50	0.58	0.52	0.42	0.65	0.72	0.60	**1**					
NFLX	0.04	0.14	0.09	0.02	0.42	0.56	0.51	0.43	**1**				
WMT	0.15	0.20	0.13	0.07	0.33	0.42	0.26	0.32	0.43	**1**			
PG	0.32	0.40	0.32	0.22	0.56	0.57	0.38	0.43	0.30	0.52	**1**		
AMZN	0.05	0.17	0.09	0.02	0.54	0.61	0.56	0.51	0.62	0.42	0.35	**1**	
HD	0.54	0.62	0.52	0.48	0.64	0.63	0.62	0.54	0.29	0.20	0.58	0.29	**1**
UNH	0.57	0.66	0.54	0.46	0.59	0.59	0.46	0.49	0.24	0.36	0.68	0.20	0.73
JNJ	0.34	0.44	0.30	0.22	0.49	0.51	0.30	0.40	0.26	0.58	0.69	0.25	0.51
NEE	0.45	0.57	0.39	0.43	0.54	0.49	0.49	0.44	0.31	0.33	0.61	0.31	0.72
AWK	0.50	0.60	0.43	0.42	0.59	0.50	0.47	0.44	0.23	0.30	0.63	0.25	0.75
CAT	0.67	0.68	0.66	0.67	0.22	0.22	0.27	0.39	0.08	0.24	0.40	0.05	0.42
RTX	0.83	0.82	0.79	0.78	0.40	0.41	0.34	0.54	0.07	0.12	0.37	0.16	0.62
XOM	0.69	0.71	0.73	0.69	0.27	0.28	0.24	0.46	0.07	0.09	0.33	0.04	0.43
SLB	0.65	0.64	0.71	0.66	0.16	0.15	0.16	0.39	0.03	0.16	0.21	0.01	0.26
SHW	0.54	0.63	0.53	0.47	0.61	0.53	0.51	0.55	0.24	0.26	0.60	0.22	0.81
SCCO	0.61	0.68	0.61	0.61	0.44	0.36	0.35	0.53	0.15	0.26	0.38	0.21	0.52
PLD	0.61	0.67	0.60	0.54	0.59	0.54	0.51	0.50	0.23	0.16	0.53	0.28	0.78
SPG	0.79	0.73	0.77	0.75	0.37	0.29	0.28	0.40	0.00	−0.03	0.25	−0.01	0.52

Sample ρ	*UNH*	*JNJ*	*NEE*	*AWK*	*CAT*	*RTX*	*XOM*	*SLB*	*SHW*	*SCCO*	*PLD*	*SPG*
UNH	**1**											
JNJ	0.67	**1**										
NEE	0.61	0.54	**1**									
AWK	0.71	0.54	0.87	**1**								
CAT	0.44	0.44	0.45	0.35	**1**							
RTX	0.64	0.39	0.56	0.59	0.60	**1**						
XOM	0.42	0.36	0.39	0.29	0.73	0.67	**1**					
SLB	0.31	0.33	0.24	0.14	0.72	0.61	0.84	**1**				
SHW	0.74	0.52	0.69	0.75	0.42	0.61	0.37	0.30	**1**			
SCCO	0.44	0.36	0.44	0.41	0.66	0.53	0.60	0.62	0.55	**1**		
PLD	0.70	0.46	0.77	0.84	0.44	0.69	0.44	0.31	0.76	0.48	**1**	
SPG	0.50	0.25	0.41	0.49	0.55	0.76	0.61	0.57	0.48	0.47	0.62	**1**

Output using R.

Over the pre-pandemic period 06-Jan-20 – 27-Dec-21, correlations between financial sector stocks (JPM, GS, BAC, WFC) were notably high, with values ranging between +0.67 (WFC-GS) and +0.91 (JPM-BAC). This suggests strong interconnectedness within the financial sector under stable market conditions. Technology and communication services stocks such as AAPL, MSFT, NVDA, and GOOGL exhibited moderate to high correlations, particularly between MSFT and GOOGL (+0.71), underscoring a degree of homogeneity in their responses to broader market trends. Correlations within consumer staples and discretionary stocks, such as WMT, PG, and AMZN, were weaker, suggesting sector-specific demand factors driving their returns.

Utilities stocks (NEE, AWK) exhibited the lowest correlations with all other stocks across the matrix, with negligible or even slightly negative relationships with industrial (CAT) and energy stocks (SLB). This aligns with the defensive nature of utilities during stable periods, as their performance is largely decoupled from broader cyclical trends. Meanwhile, industrials and energy stocks showed moderate to high intra-sector correlations, such as SLB-XOM (+0.66) and CAT-SLB (+0.58), reflecting shared exposure to macroeconomic conditions such as commodity prices and global demand.

Over the pandemic period, a notable rise in general correlations was observed, particularly among stocks in cyclical and growth sectors, reflecting heightened systemic risk and increased market integration. Financial stocks demonstrated significantly stronger correlations, with JPM, GS, and BAC correlations increasing to +0.87 to +0.95, reflecting the heightened co-movement of these stocks during volatile market conditions. Notably, WFC's correlations with other financial stocks, such as JPM and BAC, rose to +0.85 and +0.86 respectively, suggesting an alignment of risk exposures across the financial sector. The correlations between technology and communication services stocks also generally increased during the pandemic. For instance, AAPL and MSFT increased their correlation from +0.52 to +0.74. AAPL's correlations with other technology stocks, such as GOOGL (+0.65), and NVDA (+0.59), also grew substantially, indicative of the sector's alignment during the tech-driven market recovery and digitalisation trends spurred by the pandemic.

Consumer staples and discretionary stocks exhibited varying dynamics. The correlation between PG and WMT increased significantly, from +0.37 to +0.52, highlighting the growing importance of essential goods during the pandemic. In contrast, AMZN's correlations with financial and technology stocks generally weakened, especially AMZN-HD decreasing from

+0.54 to +0.29, indicating unique demand drivers for the e-commerce stock during lockdowns. Defensive sectors, mainly utilities, displayed increasing correlations with other sectors during the pandemic, a departure from their historically low pre-pandemic levels. For example, NEE's correlations with CAT rose from -0.03 to +0.45, while AWK exhibited heightened correlations with SHW (+0.22 to +0.75) and PLD (+0.49 to +0.84). This may reflect the reduced segmentation of market behaviour during periods of systemic risk, as well as the influence of fiscal and monetary policy measures in stabilising market conditions.

Stocks in the energy and materials sectors also exhibited pronounced increases in correlations during the pandemic, reflecting heightened commodity price volatility and global economic uncertainty. XOM, SLB, and SCCO typically saw their correlations with financials and industrials rise. Correlations between industrial and energy stocks strengthened, with CAT-SLB increasing from +0.58 to +0.72, signalling a convergence in their performance due to pandemic-related global economic contractions and supply chain challenges.

The general rise in sector correlations during the pandemic highlights the challenges in achieving diversification benefits when systemic risk dominates. For instance, while pre-pandemic correlations between utilities and industrials offered diversification opportunities, the pandemic-induced convergence diminished this benefit. Technology stocks, despite their relative resilience, displayed tighter clustering during the pandemic, indicating reduced opportunities for intra-sector diversification. The behaviour of defensive stocks suggests that their diversification benefits are more effective under stable conditions than during crises. The stronger interlinkages among cyclical stocks highlight the risk of synchronised downturns occurring during periods of economic contraction (see Table 7).

Table 7. Sample weekly stock-level return correlations over the period 3-Jan-22 – 25-Dec-23 (104 weeks) subject to successive federal funds rate hikes.

Sample p	*JPM*	*GS*	*BAC*	*WFC*	*AAPL*	*MSFT*	*NVDA*	*GOOGL*	*NFLX*	*WMT*	*PG*	*AMZN*	*HD*
JPM	**1**												
GS	0.79	**1**											
BAC	0.87	0.85	**1**										
WFC	0.81	0.78	0.89	**1**									
AAPL	0.49	0.44	0.45	0.41	**1**								
MSFT	0.35	0.40	0.30	0.24	0.71	**1**							
NVDA	0.46	0.47	0.37	0.37	0.64	0.70	**1**						
GOOGL	0.28	0.26	0.24	0.22	0.63	0.68	0.54	**1**					
NFLX	0.40	0.44	0.35	0.29	0.49	0.47	0.51	0.49	**1**				
WMT	0.29	0.30	0.29	0.25	0.40	0.27	0.18	0.22	0.27	**1**			
PG	0.39	0.31	0.29	0.21	0.49	0.38	0.24	0.28	0.22	0.57	**1**		
AMZN	0.47	0.50	0.43	0.43	0.58	0.68	0.62	0.68	0.61	0.27	0.24	**1**	
HD	0.48	0.57	0.49	0.44	0.52	0.47	0.44	0.36	0.40	0.37	0.45	0.44	**1**
UNH	0.16	0.12	0.05	0.06	0.34	0.31	0.11	0.20	0.17	0.29	0.33	0.13	0.16
JNJ	0.30	0.28	0.31	0.23	0.26	0.16	0.01	0.18	0.09	0.34	0.46	0.05	0.28
NEE	0.24	0.36	0.26	0.21	0.28	0.30	0.25	0.17	0.25	0.27	0.42	0.28	0.37
AWK	0.34	0.44	0.33	0.22	0.48	0.48	0.36	0.36	0.36	0.39	0.62	0.39	0.53
CAT	0.56	0.53	0.58	0.54	0.29	0.16	0.32	0.14	0.26	0.30	0.12	0.27	0.41
RTX	0.36	0.35	0.38	0.36	0.29	0.21	0.20	0.02	0.29	0.36	0.30	0.22	0.17
XOM	0.28	0.36	0.41	0.40	0.09	0.02	−0.02	0.07	0.04	0.15	0.02	0.07	0.08
SLB	0.46	0.40	0.51	0.51	0.12	0.05	0.07	0.05	0.14	0.16	0.00	0.18	0.12
SHW	0.46	0.47	0.41	0.35	0.55	0.51	0.42	0.34	0.42	0.37	0.54	0.32	0.73
SCCO	0.42	0.43	0.41	0.37	0.20	0.19	0.32	0.17	0.23	0.06	0.08	0.30	0.25
PLD	0.46	0.55	0.47	0.43	0.61	0.57	0.51	0.39	0.32	0.42	0.50	0.51	0.63
SPG	0.68	0.77	0.75	0.68	0.46	0.38	0.42	0.24	0.39	0.40	0.37	0.41	0.56

Sample p	*UNH*	*JNJ*	*NEE*	*AWK*	*CAT*	*RTX*	*XOM*	*SLB*	*SHW*	*SCCO*	*PLD*	*SPG*
UNH	**1**											
JNJ	0.45	**1**										
NEE	0.33	0.24	**1**									
AWK	0.43	0.51	0.68	**1**								
CAT	0.03	0.13	0.27	0.26	**1**							
RTX	0.25	0.22	0.35	0.35	0.41	**1**						
XOM	0.16	0.13	0.16	0.11	0.55	0.42	**1**					
SLB	0.02	0.06	0.08	0.02	0.64	0.45	0.76	**1**				
SHW	0.22	0.23	0.37	0.48	0.31	0.15	0.06	0.10	**1**			
SCCO	−0.08	0.01	0.26	0.22	0.59	0.25	0.39	0.51	0.27	**1**		
PLD	0.27	0.32	0.46	0.70	0.45	0.26	0.12	0.13	0.61	0.35	**1**	
SPG	0.06	0.25	0.38	0.49	0.62	0.35	0.33	0.42	0.53	0.47	0.64	**1**

Output using R.

In March 2022, following the initiation of the Federal Reserve's rate hikes, stock return correlations shifted significantly. This indicated broader changes in market dynamics as the monetary tightening aimed at combating inflation took effect. These changes are illustrated in Table 8, which presents the weekly correlations of stock returns from Jan-22 through Dec-23, following the pandemic period of 2020-2021.

The financial sector, which had previously exhibited strong inter-stock correlations during the pandemic period, saw a general decline in correlation strength across its constituent stocks in the rate hike period. For example, the correlations between JPM and GS, BAC, and WFC decreased from the pandemic highs to pre-pandemic levels: JPM-GS dropped from +0.91 to +0.79, JPM-BAC from +0.95 to +0.87, and JPM-WFC from +0.85 to +0.81. These declines may suggest a less synchronised response to the macroeconomic conditions imposed by the Federal Reserve's tightening, likely reflecting adjustments to rising interest rates, changing credit conditions, and sector-specific risks. The weaker correlations may also indicate a more varied market perception of financial institutions' ability to manage the impacts of inflation and rising borrowing costs.

Technology and communication services stocks continued to display elevated correlations for most pairs during the rate hike period. Despite market fluctuations as tech stocks adjusted to higher interest rates, which typically depress the valuation multiples of growth stocks, correlations between these stocks remained strong and stable. This may indicate continued alignment within the sector as it adapted to changing monetary conditions, such as higher borrowing costs and a focus on profitability over expansion.

Consumer staples and discretionary stocks showed somewhat stable correlations between the pandemic and rate hike periods. For example, the correlation between WMT and PG increased slightly from +0.52 to +0.57, reflecting an enduring alignment between companies in the consumer defensive sector as consumers continued to prioritise essential goods. In contrast, the correlation between AMZN and HD in the consumer discretionary sector increased from +0.29 to +0.44. However, correlations between consumer staples and discretionary stocks weakened. For instance, AMZN's correlations with WMT and PG dropped during the rate hike period (e.g., AMZN-WMT decreased from +0.42 to +0.27). This change may reflect the divergent impacts of rising interest rates on discretionary spending and e-commerce growth, with some consumers becoming more price-sensitive and others shifting spending patterns.

Utilities showed a notable decline in correlations compared to their heightened pandemic-period values, suggesting a partial return to their traditional defensive characteristics. For instance, NEE's correlation with CAT dropped significantly, from +0.45 during the pandemic to +0.27. Similarly, AWK, which exhibited strong correlations with industrials and materials stocks during the pandemic (e.g., +0.75 with SHW), saw these connections weaken (to 0.48). However, some correlations, such as AWK-PLD, remained relatively elevated at +0.70 (down from +0.84), reflecting continued exposure to macroeconomic factors such as interest rate sensitivity. These shifts imply that utilities began to regain their diversification potential, though higher rates likely sustained some residual interconnectedness.

Energy stocks displayed a pronounced decoupling from other sectors, reflecting reduced systemic shocks but persistent commodity price volatility under a tightening monetary environment. For example, XOM’s correlation with industrials, such as CAT, declined from +0.73 during the pandemic to +0.55. Similarly, SLB's correlation with RTX weakened from +0.61 to +0.45, indicating a reversal of the pandemic-induced interlinkages driven by synchronised global supply chain disruptions. In materials, SHW’s correlation with SCCO saw a marked decline, dropping from +0.55 to +0.27, suggesting a divergence in sector performance driven by differentiated impacts of rising interest rates and demand-side pressures.

The period of Fed rate hikes from Mar-22 – Jul-23 indicates a notable shift in sectoral correlations following the pandemic period. The financial sector saw a broad decline in intra-sector correlations, suggesting a more diverse market response to macroeconomic conditions such as rising interest rates and inflation. Technology and communication services maintained elevated correlations, indicating sustained sector alignment despite valuation pressures from tightening monetary policy. Consumer staples and discretionary stocks exhibited stable intra-sector correlations, while inter-sector correlations weakened, highlighting divergent impacts of inflation on essential versus discretionary spending. Utilities experienced a partial reversion to their traditional defensive role, with declining correlations reflecting improved diversification potential. Meanwhile, energy and materials sectors demonstrated significant decoupling from other sectors, driven by persistent commodity price volatility and reduced systemic shocks. Overall, the observed trends highlight the impact of changes in monetary policy and sector-specific risk factors during periods of economic adjustment.

3.4 Portfolio and shrinkage estimator overview

Table 8 provides an overview of the portfolios, and VCV matrix shrinkage estimators used in the subsequent analyses.

Table 8: Description of portfolios and VCV matrix estimators. All portfolios are unconstrained in terms of short-selling and asset weights.

Panel A: Efficient Frontier & Benchmark Portfolios	
Efficient Frontier Portfolio	Description
GMV	Global minimum variance portfolio
TG	Tangency portfolio
Benchmark Portfolio	Description
EW	Equally weighted portfolio
IV	Inverse volatility portfolio
Panel B: VCV Matrix Estimators	
Estimator	Description
Sample	Sample estimator (unshrunk)
CC-BI	Linear shrinkage towards a constant correlation target; bias ignored
CC-BC	Linear shrinkage towards a constant correlation target; bias corrected
ZC-BI	Linear shrinkage towards a zero-correlation target; bias ignored
ZC-BC	Linear shrinkage towards a zero-correlation target; bias corrected
AV	Average of sample correlations as a predictor of individual asset correlations

4 Empirical results

4.1 Overview of portfolio performance statistics (full out-of-sample period)

Table 9 shows the full out-of-sample performance statistics of the benchmark portfolios, efficient frontier portfolios, and VCV matrix shrinkage estimators detailed in Table 8. These include annualised volatility, defined by the standard deviation of the annualised cumulative (excess) return (%) (Vol.), annualised cumulative (excess) return (%) (Ret.), and annualised Sharpe ratio (SR). The first two rows present the performance statistics for the two benchmark portfolios. For each efficient frontier portfolio and shrinkage estimator combination, we use a two-year VCV estimator as input to the respective portfolio method. Therefore, the results presented are for the defined S&P 500 index asset universe using a rolling two-year (104 week) estimation period over the full out-of-sample period from 30-Dec-19 – 20-May-24. For example, the first annualised cumulative return, estimated on 30-Dec-19, is based on the preceding two-year period of weekly portfolio excess returns from 01-Jan-18 – 30-Dec-19. The next value, estimated on the following weekly date (06-Jan-20), is then based on the 104-week period from 08-Jan-18 – 06-Jan-20, and so forth.

Table 9: Performance overview of the benchmark and efficient frontier portfolios. Reported values indicate the mean of the respective performance statistic measured over the full out-of-sample period from 30-Dec-19 – 20-May-24.

Portfolio	Estimator	Mean Vol. (%)	Mean Ret. (%)	Mean SR
EW	-	22.14	15.52	0.70
IV	-	20.26	15.82	0.78
GMV	Sample	13.19	13.17	1.00
	CC-BI	14.09	13.89	0.99
	CC-BC	14.08	13.89	0.99
	ZC-BI	13.68	13.48	0.99
	ZC-BC	13.67	13.47	0.99
	AV	13.00	17.19	1.32
TG	Sample	181.25	408.08	2.25
	CC-BI	507.28	-129.98	-0.26
	CC-BC	698.31	-753.34	-1.08
	ZC-BI	252.82	-111.51	-0.44
	ZC-BC	273.71	-169.64	-0.62
	AV	137.71	101.38	0.74

Vol., annualised volatility; Ret., annualised cumulative return; SR, annualised Sharpe ratio.
Output using Excel.

4.1.1 Global minimum variance portfolio

Firstly, the benchmark EW and IV portfolios exhibit mean annualised volatilities of 22.14% and 20.26%, respectively, with mean Sharpe ratios of 0.70 and 0.78. In comparison, the GMV portfolio achieves mean annualised volatility levels significantly lower than those of the benchmark portfolios for all estimators, ranging from 13.00% to 14.09%.

The average correlation (AV) estimator stands out as an exception, delivering the lowest mean annualised volatility (13.00%), the highest mean annualised cumulative return (17.19%), and the best mean annualised Sharpe ratio (+1.32) among all GMV portfolios. Notably, the AV estimator is not a linear shrinkage method in the strict sense; instead, it homogenises individual asset correlations by setting them all equal to their sample average. This simplification, while potentially neglecting unique inter-asset relationships, appears to mitigate estimation noise effectively on average, thereby improving portfolio stability and performance.

Among the linear shrinkage estimators of Ledoit and Wolf (2003, 2004a, 2004b) and Kwan (2017), the constant correlation (CC-BI and CC-BC) and zero correlation (ZC-BI and ZC-BC) target estimators yield similar annualised Sharpe ratios (approximately 0.99). Correcting for statistical bias with the CC-BC and ZC-BC estimators does not result in significantly different performance statistics from their bias ignored counterparts CC-BI and ZC-BI, respectively; this observation holds for both GMV and TG portfolio applications. The zero correlation target estimators exhibit slightly lower volatility and cumulative return values compared to the constant correlation counterparts, reflecting the subtle differences in their approach to correlation structure. However, these methods fail to meaningfully improve upon the sample covariance matrix estimator, suggesting that linear shrinkage may not sufficiently address estimation errors for GMV portfolio construction over this sample period on average.

4.1.2 Tangency portfolio

The performance statistics of the TG portfolio highlight the challenges of mean-variance optimisation under unstable covariance matrix estimation. The sample covariance matrix produces extreme results, with a mean annualised cumulative return of 408.08%, a Sharpe ratio of +2.25, and annualised volatility of 181.25%. These implausible figures underscore the sensitivity of the TG portfolio to estimation errors and input instability over turbulent market periods.

The application of linear shrinkage estimators seemingly exacerbates this instability. For example, the CC-BC estimator produces a strikingly negative mean annualised cumulative return of -753.34% and a Sharpe ratio of -1.08, while the ZC-BI estimator yields -111.51% and a Sharpe ratio of -0.44. These results initially suggest that linear shrinkage techniques may amplify estimation errors when applied to the TG portfolio, particularly in volatile market conditions.

The AV estimator offers relatively better performance within this context, achieving a mean annualised cumulative return of 101.38% and a Sharpe ratio of +0.74. Despite these improvements, the AV estimator's resulting annualised volatility of 137.71% still indicates impractical performance. The inability of any estimator to produce plausible and stable TG portfolio returns on average may suggest inherent weaknesses in the mean-variance optimisation process during periods of market turbulence.

4.2 Efficient frontier instability during market turbulence

The efficient frontier represents the set of investment portfolios offering the highest expected return for a given level of risk, or the lowest risk for a defined level of expected return. This includes the GMV and TG portfolios. Figures 7-14 illustrate the behaviour of efficient frontiers at two distinct points in the out-of-sample period: the first estimation date, 30-Dec-19, and 03-Oct-22. These dates represent the contrasting periods of relative market stability and significant turbulence induced by the Federal Reserve's interest rate hikes, respectively.

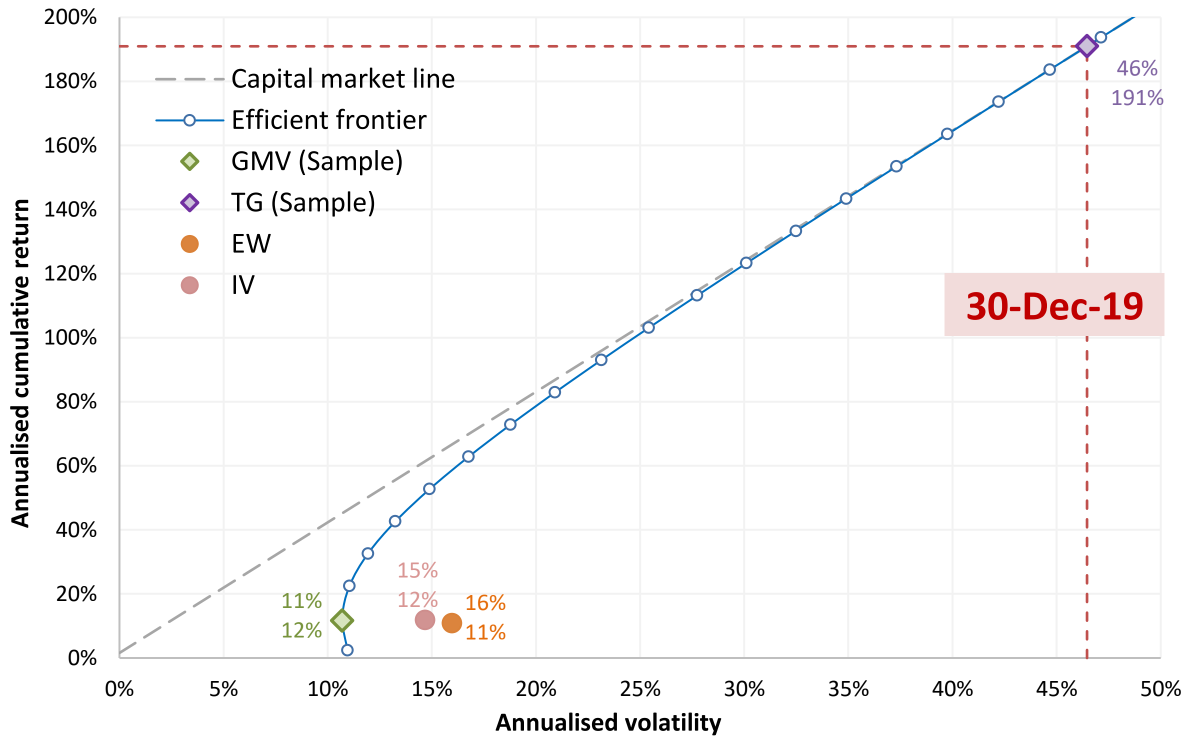


Figure 7. Efficient frontier plot as at 30-Dec-19. GMV and TG portfolios are constructed using the Sample estimator.

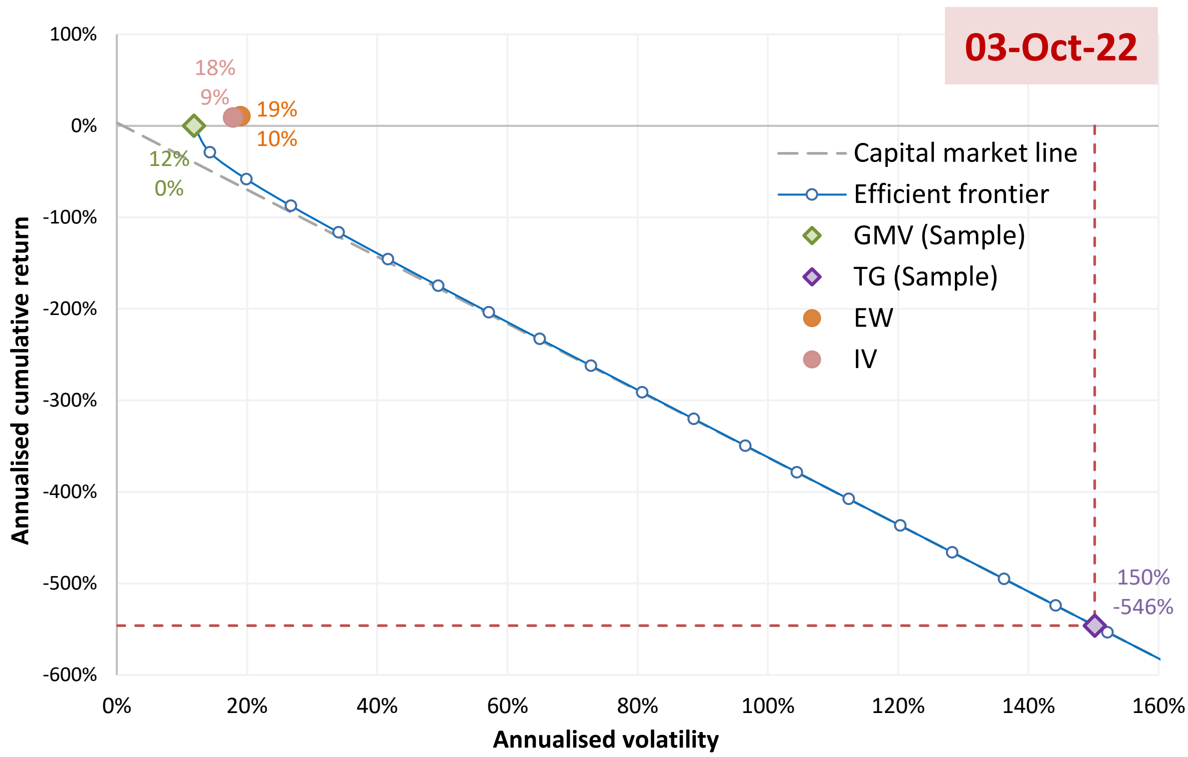


Figure 8. Efficient frontier plot as at 03-Oct-22. GMV and TG portfolios are constructed using the Sample estimator.

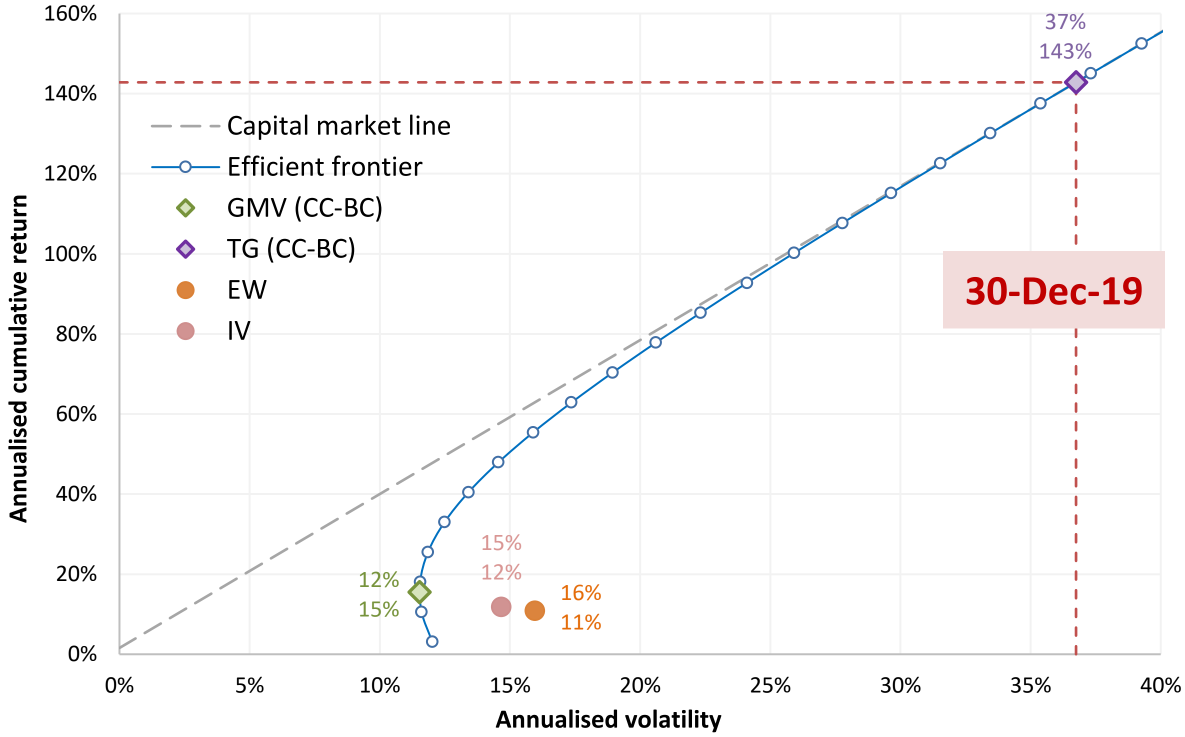


Figure 9. Efficient frontier plot as at 30-Dec-19. GMV and TG portfolios are constructed using the CC-BC estimator.

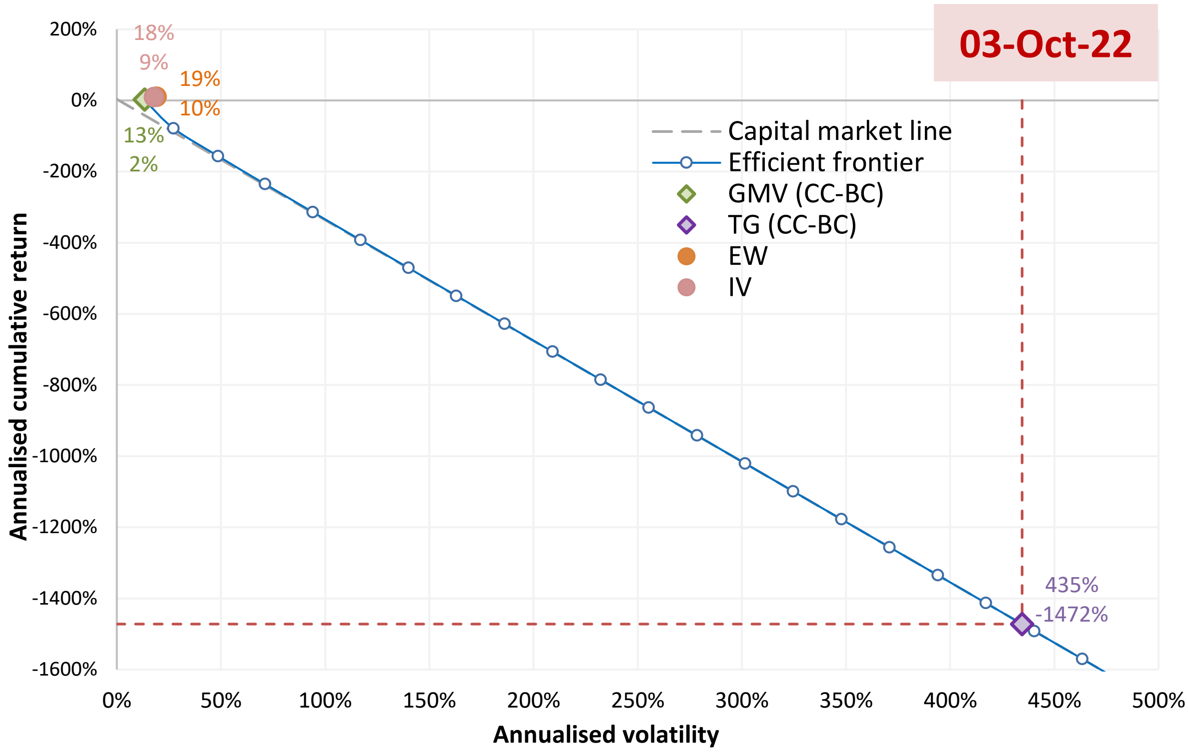


Figure 10. Efficient frontier plot as at 03-Oct-22. GMV and TG portfolios are constructed using the CC-BC estimator.

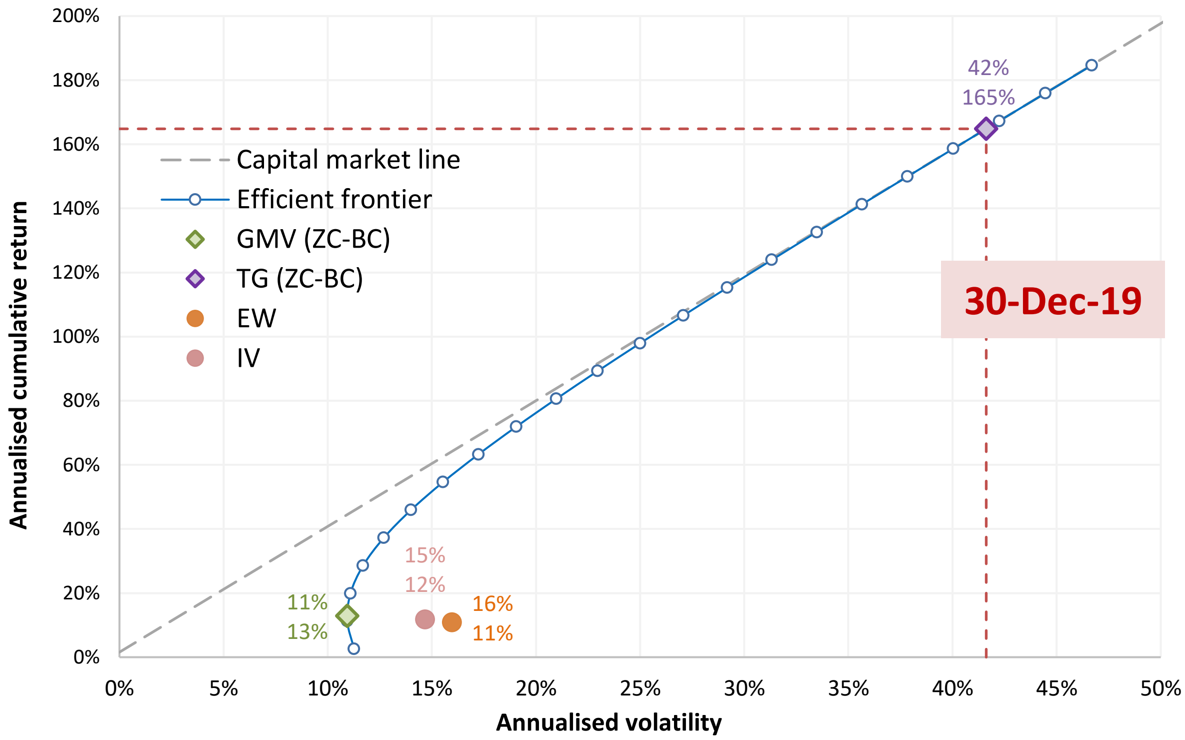


Figure 11. Efficient frontier plot as at 30-Dec-19. GMV and TG portfolios are constructed using the ZC-BC estimator.

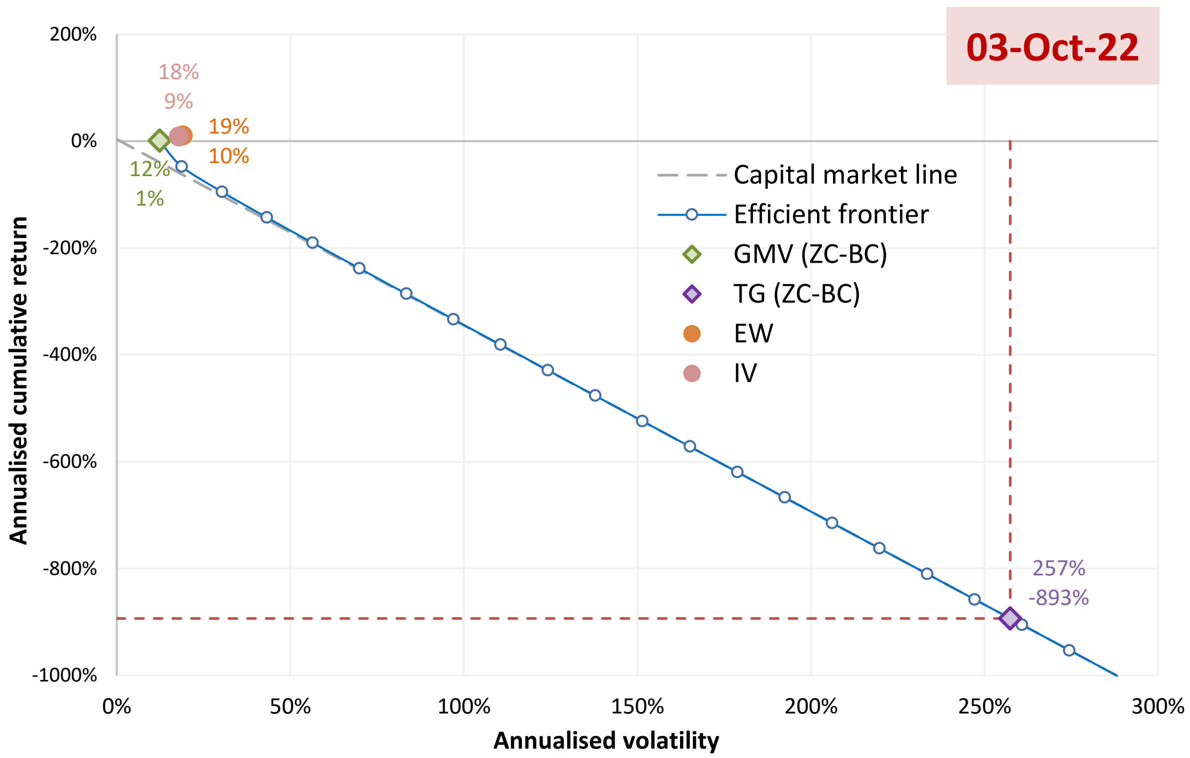


Figure 12. Efficient frontier plot as at 03-Oct-22. GMV and TG portfolios are constructed using the ZC-BC estimator.

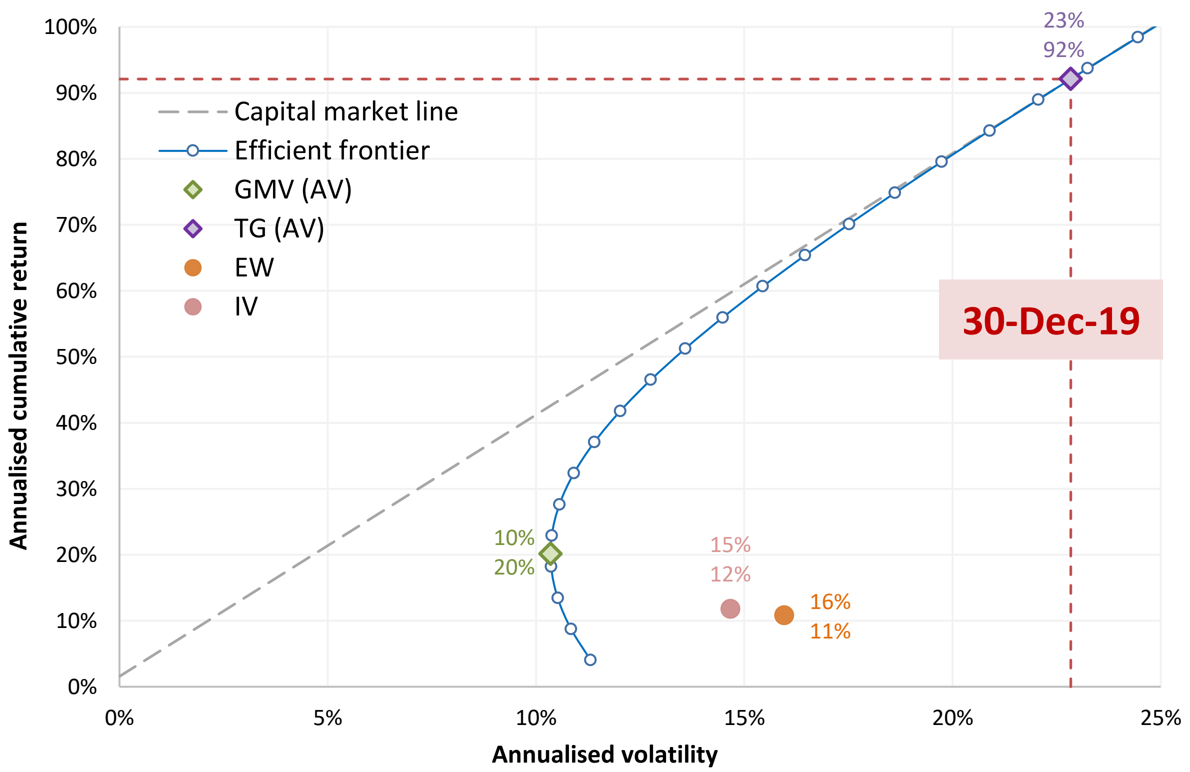


Figure 13. Efficient frontier plot as at 30-Dec-19. GMV and TG portfolios are constructed using the AV estimator.

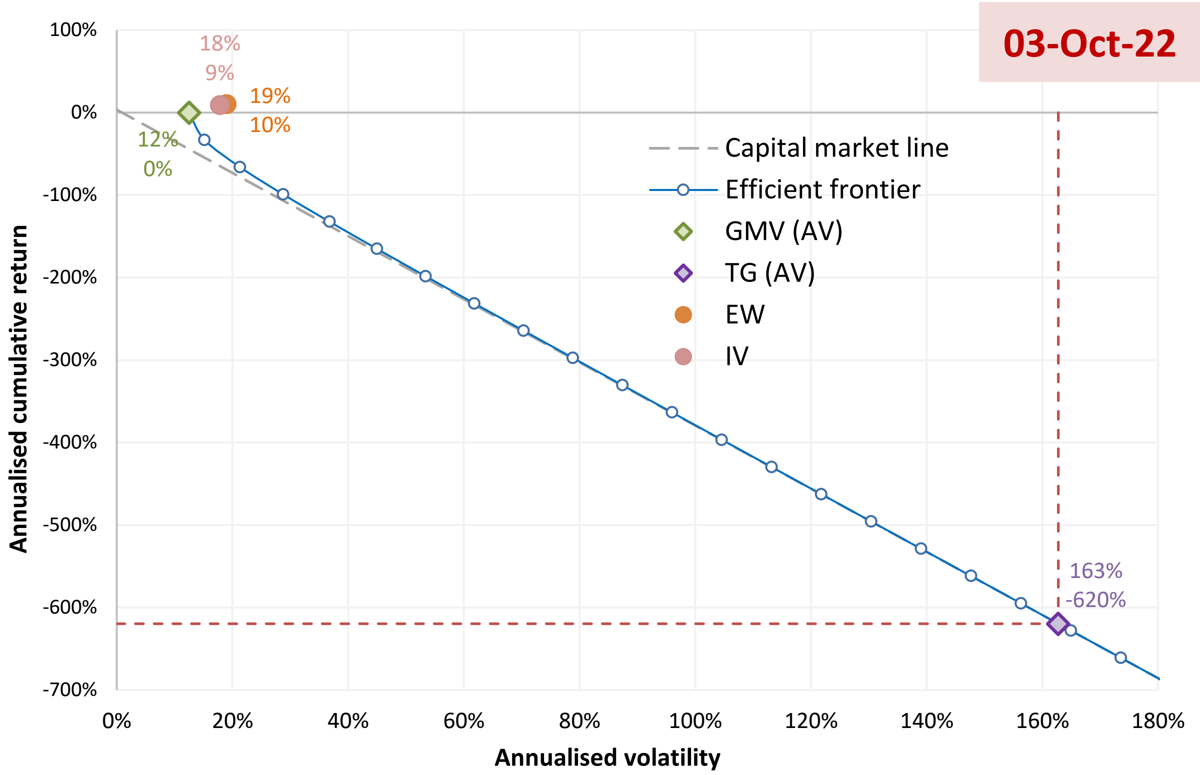


Figure 14. Efficient frontier plot as at 03-Oct-22. GMV and TG portfolios are constructed using the AV estimator.

On 30-Dec-19, the efficient frontiers derived from all VCV matrix estimators exhibit the expected upward-sloping curve. For instance, the efficient frontier constructed using the sample covariance matrix is consistent with theoretical expectations, producing practical portfolios with increasing returns as risk (volatility) increases. All other estimators, such as CC-BC, ZC-BC, and AV, similarly produce stable efficient frontiers on this initial date. In contrast, the efficient frontiers on 03-Oct-22, demonstrate significant instabilities and inversion. This period corresponds to the Federal Reserve's aggressive monetary tightening, occurring from Mar-22 to Jul-23, leading to elevated interest rates and market stress. Frequent and consecutive rising of risk-free rates substantially impact the mean-variance optimisation process, with the inverted capital market line reflecting a negative risk-return trade-off. The inversion of the efficient frontier is observed across all test portfolio and estimator combinations on several later dates within the out-of-sample period.

The sample covariance matrix produces a severely inverted efficient frontier on 03-Oct-22, where the TG portfolio's annualised cumulative return drops to -546% with an exceedingly high annualised volatility of 150%. Such extreme values render the TG portfolio impractical. Similarly, while the GMV portfolio shows modest volatility at 11.9%, it fails to deliver a positive, non-zero return. The linear shrinkage estimators, such as CC-BC and ZC-BC, fail to mitigate these observed instabilities. Their respective efficient frontiers invert similarly, with negative TG portfolio annualised cumulative returns and Sharpe ratios. Even the AV estimator, which outperforms on average with the GMV portfolio, fails to stabilise the frontier. Although it demonstrates some improvement over the other linear shrinkage estimators, the AV-based TG portfolio still suffers from impractically negative and volatile returns during this period.

The inversion of efficient frontiers during the observed period of heightened market stress potentially indicates limitations of both the sample and the tested linear shrinkage estimators in producing reliable mean-variance optimisation outputs. Elevated risk-free rates and increased market volatility, observed during the Federal Reserve's monetary tightening period, seem to disrupt the optimisation process. This results in theoretically inconsistent risk-return trade-offs across all estimators, particularly when optimising for risk-adjusted returns with the TG portfolio.

4.3 Overview of portfolio performance statistics (truncated out-of-sample period)

To further investigate the nature of shrinkage-based portfolio performance, we repeat the analysis from Section 4.1 but restrict the out-of-sample period to 30-Dec-19, through 30-May-22. This truncated period primarily captures the pre-Federal Reserve rate hike environment and the immediate aftermath of the COVID-19 market recovery, excluding the prolonged turbulence that followed. This allows for a comparison of the mean performance statistics over this more "stable" market period relative to those measured over the full out-of-sample period. We note the degree to which the TG portfolio’s performance was disproportionately affected by the rate hike period. Table 10 reports the mean annualised volatility, annualised cumulative return, and annualised Sharpe ratio for each portfolio and estimator combination over the truncated out-of-sample period (30-Dec-19 – 30-May-22).

Table 10: Performance overview of the benchmark and efficient frontier portfolios. Reported values indicate the mean of the respective performance statistic measured over the truncated out-of-sample period from 30-Dec-19 – 30-May-22.

Portfolio	Estimator	Mean Vol. (%)	Mean Ret. (%)	Mean SR
EW	-	24.57	20.17	0.82
IV	-	22.27	21.62	0.97
GMV	Sample	14.24	17.89	1.26
	CC-BI	15.08	19.89	1.32
	CC-BC	15.08	19.89	1.32
	ZC-BI	14.77	18.75	1.27
	ZC-BC	14.76	18.73	1.27
	AV	13.46	29.62	2.20
TG	Sample	49.86	201.01	4.03
	CC-BI	46.83	177.74	3.79
	CC-BC	46.85	177.80	3.80
	ZC-BI	48.59	188.09	3.87
	ZC-BC	48.62	188.32	3.87
	AV	30.64	141.42	4.61

Vol., annualised volatility; Ret., annualised cumulative return; SR, annualised Sharpe ratio.
Output using Excel.

4.3.1 Global minimum variance portfolio

As observed in Table 10, the benchmark portfolios (EW and IV) exhibit marginally higher mean annualised volatility in the truncated period relative to the full sample results, but also a greater increase in annualised cumulative return. The EW and IV portfolios' respective annualised Sharpe ratios are +0.82 and +0.97, reflecting a noticeable improvement in risk-adjusted return compared to the full sample period.

The GMV portfolio follows a similar trend, showing increased risk-adjusted returns across the board when the most volatile market period of the sample is excluded. The sample covariance estimator results in a mean annualised volatility of 14.24%, while the linear shrinkage estimators (CC-BI, CC-BC, ZC-BI, ZC-BC) yield slightly higher volatility values in the range of

14.76% to 15.08%. Notably, the AV estimator achieves the lowest mean volatility at 13.46%. Whilst the linear shrinkage estimators still do not seem to reduce risk compared to the sample estimator, they produce higher annualised cumulative returns on average. This in turn results in greater annualised Sharpe ratios for the linear and AV estimators relative to the sample estimator; the AV estimator achieves the highest mean annualised Sharpe ratio at +2.20, followed by CC-BI and CC-BC at +1.32, and sample at +1.26. Despite the marginal increase in annualised volatility for all estimators during this truncated period, the improved risk-adjusted performance of the GMV portfolio when using linear shrinkage estimators suggests that significant market turbulence may hinder their efficacy.

4.3.2 Tangency Portfolio

The impact of excluding the post-2022 rate hike period is most pronounced in the TG portfolio, which optimises for risk-adjusted returns and is thus highly sensitive to changes in the underlying risk environment. Over the truncated period, the sample estimator produces an annualised cumulative return of 201.01%, with a corresponding annualised volatility of 49.86%. The CC-BI and CC-BC estimators generate slightly lower returns of approximately 177.74% and 177.80%, respectively, with corresponding volatilities of around 46.83% and 46.85%. The ZC-BI and ZC-BC estimators yield marginally higher returns at approximately 188.09% and 188.32%, with volatilities of 48.59% and 48.62%, respectively. The AV estimator, while reducing annualised volatility significantly to 30.64%, still achieves a relatively high annualised cumulative return of 141.42%.

Although these TG risk and return values are still somewhat impractical due to our assumption of fully unconstrained portfolios, they are much more stable prior to the period of Fed rate hikes. The sample estimator yields a mean annualised Sharpe ratio of +4.03, almost double than what was observed in the full out-of-sample period. In contrast to their previously negative values, the four linear shrinkage estimators (excluding AV) now produce positive and strong annualised Sharpe ratios in the range of +3.79 to +3.87. The AV estimator now stands out with the highest Sharpe ratio of +4.61, further suggesting its effectiveness in more stable market environments.

Compared to the full out-of-sample period, the truncated period results indicate that the performance of the TG portfolio was disproportionately affected by the Federal Reserve rate hikes and subsequent market dislocations. The much higher mean Sharpe ratios across all estimators in the truncated period suggest that the extreme fluctuations in market risk premia during the latter part of the sample period introduced significant instability, distorting the portfolio's expected risk-adjusted return profile. This further supports the notion that mean-variance optimization strategies are highly sensitive to shifts in the market regime. However, the four linear shrinkage estimators still fail, on average, to improve upon the risk-adjusted return of the TG portfolio relative to the sample estimator.

4.4 Descriptive time series analysis of portfolio performance statistics

To further investigate the temporal dynamics of the portfolio performance statistics, we conduct a descriptive time series analysis of these measures over the full out-of-sample period. Figures 15-20 show plots of annualised volatility, annualised cumulative return, and annualised Sharpe ratio as a function of time for both GMV and TG portfolios from 30-Dec-19 – 20-May-24.

4.4.1 Global minimum variance portfolio

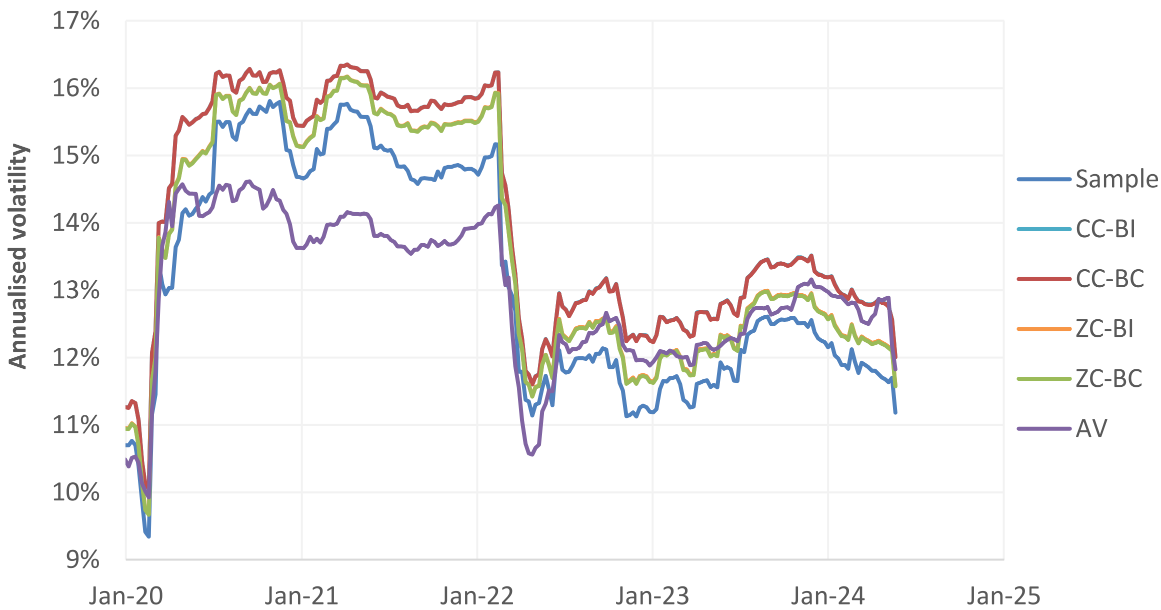


Figure 15. Time series plots of GMV portfolio annualised volatility (Vol.) using sample, linear shrinkage, and average estimators. The full out-of-sample period is 30-Dec-19 – 20-May-24.

Figures 15, 16, and 17 depict the time series of annualised volatility, annualised cumulative return, and annualised Sharpe ratio, respectively, for the GMV portfolio. A key feature observed in the data is the sharp increase in portfolio risk during the first quarter of 2020, aligning with the onset of the COVID-19 pandemic and the accompanying financial market turmoil. Across all estimators, annualised volatility spikes from an initial level of approximately 10–11% to a peak of around 16%. However, while this increase is a common feature among all methods, variations in magnitude highlight differences in the ability of shrinkage estimators to mitigate extreme market-induced fluctuations.

Following the initial COVID-19-induced volatility shock, portfolio risk remains persistently elevated throughout 2020 and 2021. During this period, the linear shrinkage estimators all consistently exhibit greater volatility levels than the sample estimator, whereas the AV estimator produces relatively lower risk estimates. While all estimators exhibit broadly similar trajectories of annualised volatility, the sample estimator exhibits somewhat greater fluctuations. The linear shrinkage and AV estimators seem to produce slightly smoother volatility paths, suggesting that imposing structure on correlation estimates may lead to greater stability in risk estimates over time.

A decline in annualised volatility across all estimators occurs in March 2022, potentially due to broader shifts in market conditions as the series of Federal Reserve rate hikes begins on March 16, 2022. This results in a stabilisation of volatility at lower levels (11–13%) from June 2022 onwards, with the sample estimator then consistently achieving the lowest annualised volatility levels for the remainder of the out-of-sample period. In contrast, the linear shrinkage estimators still fail to significantly outperform the sample covariance matrix in volatility reduction, contradicting their theoretical motivation. Despite the findings of Ledoit and Wolf (2003, 2004a, 2004b) which suggest that linear shrinkage towards structured correlation matrices can mitigate estimation error, these methods still yield high levels of annualised risk relative to the sample covariance matrix over this sample period from 30-Dec-19 – 20-May-24.

The time series of annualised cumulative return and annualised Sharpe ratio for the GMV portfolio show additional long-term performance trends of the tested estimators. During early 2020, all portfolio estimators experience a sharp increase in annualised cumulative return, coinciding with the heightened market volatility. The AV estimator achieves the highest annualised cumulative return throughout 2020 and 2021, peaking above 40%, whereas the sample covariance and linear shrinkage estimators cluster together with no significant differences, exhibiting returns in the range of 20–30%. Notably, as shown in Table 10, linear shrinkage estimators achieve marginally higher annualised cumulative returns and annualised Sharpe ratios, on average, compared to the sample estimator during this truncated period.

Figure 16. Time series plots of GMV portfolio annualised cumulative return (Ret.) using sample, linear shrinkage, and average estimators. The full out-of-sample period is from 30-Dec-19 – 20-May-24.

Figure 17. Time series plots of GMV portfolio annualised Sharpe ratio (SR) using sample, linear shrinkage, and average estimators. The full out-of-sample period is from 30-Dec-19 – 20-May-24.

However, from March 2022 onward, coinciding with the Federal Reserve's rate hike cycle, the linear shrinkage estimators begin to slightly underperform the sample estimator in both cumulative return and Sharpe ratio. This pattern explains the equalisation of annualised Sharpe ratios, on average, across the sample and linear shrinkage estimators for the full out-of-sample period in Table 9. The AV estimator, while initially outperforming all other approaches, deviates significantly with negative risk-adjusted returns from mid-2023 onwards.

4.4.1.1 Do shrinkage estimators result in statistically significant reductions in GMV portfolio risk?

The GMV portfolio is explicitly designed to achieve the lowest possible level of risk for a given set of assets. Therefore, any statistically significant increase in annualised volatility (Vol.) when applying shrinkage estimators, relative to the sample VCV matrix, suggests a failure to meet the portfolio's primary objective. To test this, paired samples t-tests are conducted for each sample and shrinkage estimator combination. The null hypothesis states that there is no statistically significant difference between the mean Vol. estimated using the sample VCV matrix and that estimated using shrinkage estimators. The analysis is performed over two sub-periods within the full out-of-sample period, namely the "pre-rate hike period" and "rate hike period" as discussed previously. This aims to further analyse the impact of differing market conditions on shrinkage estimator efficacy. The results are displayed in Table 11 below.

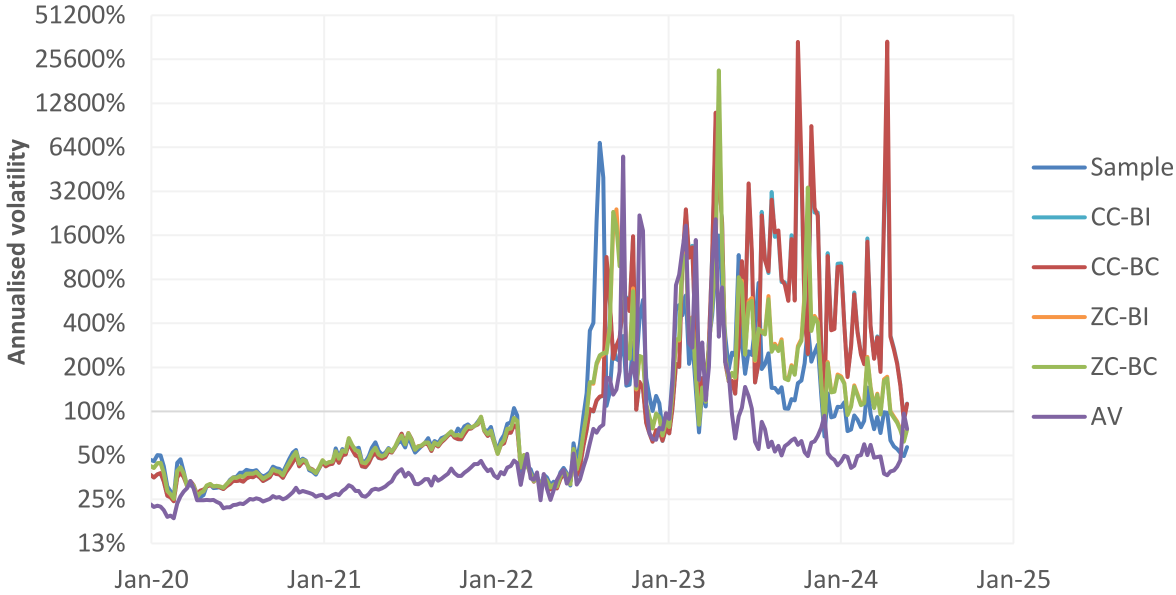
Table 11. Paired samples t-test of GMV portfolio annualised volatility (Vol.). The null hypothesis states that there is no statistically significant difference between the sample VCV matrix estimated mean Vol. and shrinkage estimator mean Vol.

Out-of-sample period	Estimator	Mean Vol. (%)	t-statistic	p-value (two-tail)
Pre-rate hike period (Dec 30, 2019 - May 30, 2022)	Sample	14.24	-	-
	CC-BI	15.08	-33.53	0.0000**
	CC-BC	15.08	-33.48	0.0000**
	ZC-BI	14.77	-31.36	0.0000**
	ZC-BC	14.76	-31.10	0.0000**
	AV	13.46	13.55	0.0000**
Rate hike period (May 30, 2022 - May 20, 2024)	Sample	11.89	-	-
	CC-BI	12.85	-86.35	0.0000**
	CC-BC	12.84	-86.27	0.0000**
	ZC-BI	12.33	-90.17	0.0000**
	ZC-BC	12.32	-89.60	0.0000**
	AV	12.43	-19.36	0.0000**

Note: ** denotes significance at the 1% level. * denotes significance at the 5% level.
Output using R.

During the pre-rate hike period, all linear shrinkage estimators result in statistically significant increases in GMV portfolio annualised volatility, on average, relative to the sample estimator. This indicates that they fail to optimise the GMV portfolio over the sample VCV matrix. The AV estimator is the only exception, where it achieves a statistically significant reduction in risk relative to the sample estimator. In the subsequent Fed rate hike period, all shrinkage estimators still produce statistically significant differences relative to the sample estimator. Unlike in the more stable pre-rate hike period, however, none of the shrinkage estimators — including AV — lead to a reduction in risk. Instead, all shrinkage estimators result in statistically significant increases in annualised volatility, suggesting that shrinkage does not significantly improve risk minimisation in this particular monetary policy environment.

4.4.2 Tangency portfolio



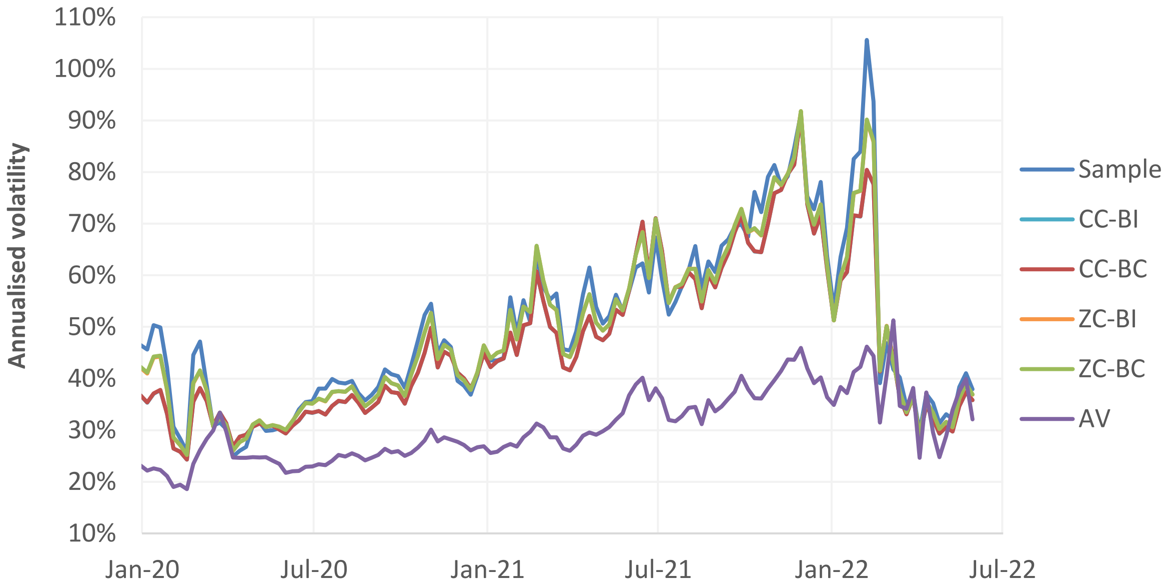


Figure 18. Time series plots of TG portfolio annualised volatility (Vol.) using sample, linear shrinkage, and average estimators. The top panel shows the full out-of-sample period from 30-Dec-19 – 20-May-24 (logarithmic scale), while the bottom panel shows the truncated period from 30-Dec-19 – 30-May-22.

The instability of the TG portfolio's risk and return estimates following the period of Fed rate hikes is evident in Figures 18-20. Annualised volatility of the TG portfolio exhibits pronounced fluctuations over the full out-of-sample period, particularly sharp increases in risk estimates following the onset of the Federal Reserve's rate hikes in March 2022. Compared to the GMV portfolio, which prioritises risk minimisation, the TG portfolio aims to achieve the highest risk-adjusted return. The mean-variance optimisation process is evidently highly sensitive to estimation errors in this regard, given the extreme and unstable performance statistics of the TG portfolio from mid-2022 onwards for all estimators.

Prior to the Fed rate hikes, annualised risk estimates for the TG portfolio remain relatively stable, with linear shrinkage estimators marginally reducing volatility relative to the sample covariance matrix. Notably, the AV estimator achieves significantly lower annualised volatility relative to all other estimators from 2020 to early 2022. However, once market conditions become more turbulent, particularly after the initiation of rate hikes in March 2022, the instability in risk estimates are pronounced. All estimators respond with extremely high and implausible risk estimates, with the linear shrinkage estimators seemingly exacerbating the observed instability towards the end of the period. Again, while the AV estimator does not fully mitigate the extreme estimation error, it provides significantly lower and more stable annualised volatility estimates from mid-2023 onwards.

The performance of the TG portfolio in terms of both annualised cumulative return and risk-adjusted return (SR) exhibits notable characteristics across the pre-rate hike period. The annualised cumulative return estimates show a general upward trend from 2020 to late 2021, with all estimators following a broadly similar trajectory. The sharp initial decline, followed by a steady rebound, reflects the high levels of market uncertainty induced by the COVID-19 pandemic and subsequent market recovery. Although the sample estimator seems to yield the highest annualised cumulative return estimates, the linear shrinkage estimators show relatively similar levels of return. The AV estimator exhibits the lowest returns for most of the period, which is consistent with its more conservative risk estimates observed in Figure 18.

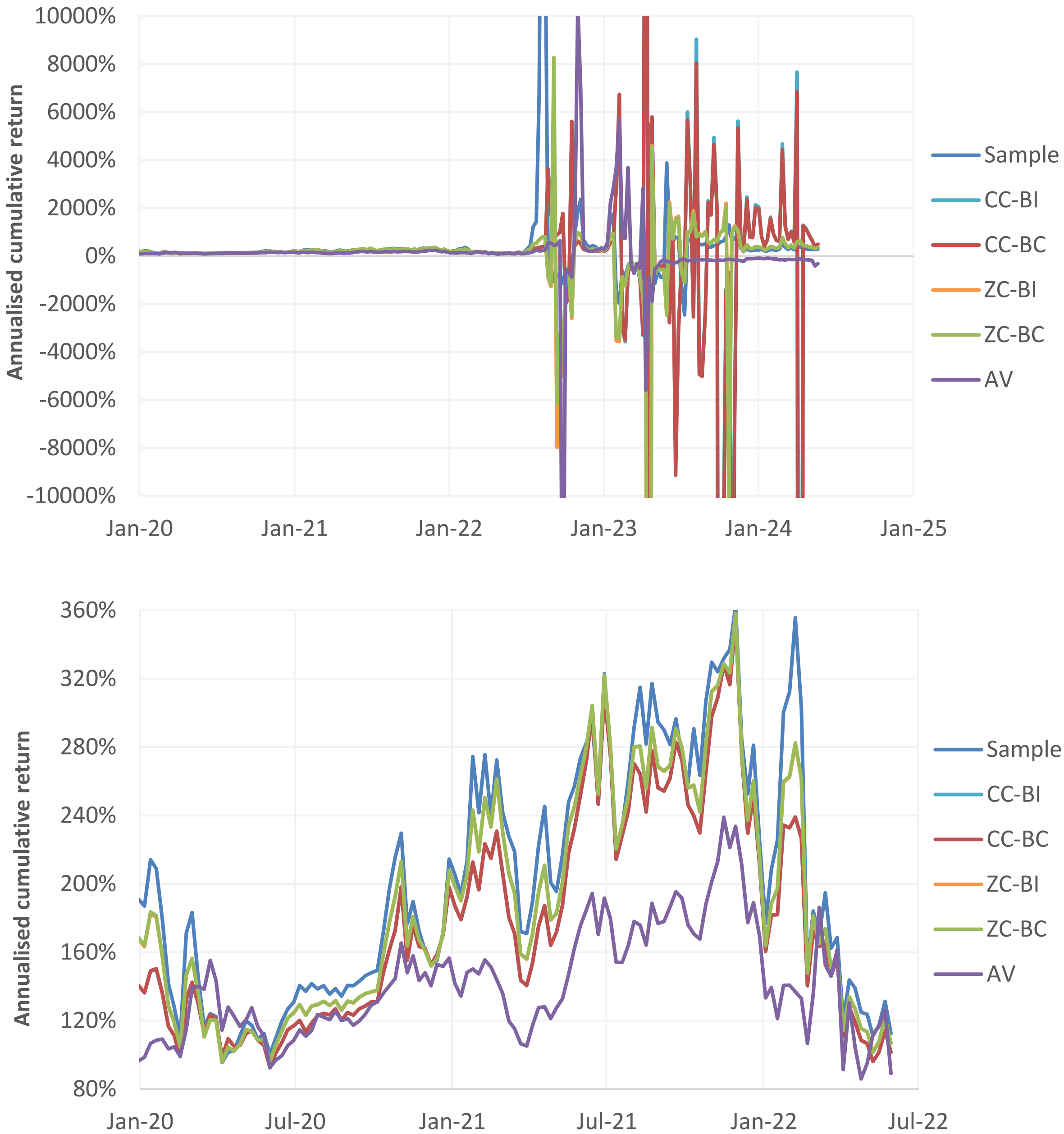


Figure 19. Time series plots of TG portfolio annualised cumulative return (Ret.) using sample, linear shrinkage, and average estimators. The top panel shows the full out-of-sample period from 30-Dec-19 – 20-May-24, while the bottom panel shows the truncated period from 30-Dec-19 – 30-May-22.

Figure 20. Time series plots of TG portfolio annualised Sharpe ratio (SR) using sample, linear shrinkage, and average estimators. The top panel shows the full out-of-sample period from 30-Dec-19 – 20-May-24, while the bottom panel shows the truncated period from 30-Dec-19 – 30-May-22.

The annualised Sharpe ratio initially exhibits a somewhat level trend, with relatively stable risk-adjusted performance in the pre-rate hike period. The AV estimator stands out as achieving consistently higher Sharpe ratios relative to the sample estimator throughout 2020. In contrast, the linear shrinkage estimators do not seem to offer notable improvement in risk-adjusted performance over this pre-rate hike period. Beginning in Mar-22, there is a brief downward trend in annualised Sharpe ratio estimates across all estimators, signalling a deterioration in risk-adjusted performance leading into the Fed rate hike period. This decline is consistent with increasing macroeconomic uncertainty, inflationary pressures, and a shift in market sentiment away from the conditions that had previously supported low interest rates.

Like Figure 18, Figures 19 and 20 depict extreme levels of instability in annualised cumulative return and Sharpe ratio estimates for the TG portfolio thereafter. This may further suggest that the observation of an inverted and highly unstable efficient frontier is primarily a consequence of aggressive monetary tightening. During this latter period, the linear shrinkage estimators do not mitigate the observed instability in the sample VCV matrix; in fact, they seemingly exacerbate the issue, resulting in both lower and more volatile risk-adjusted return estimates for the TG portfolio. Interestingly, the AV estimator exhibits a consistently negative annualised Sharpe ratio estimate from early 2023 onwards.

4.4.2.1 Do shrinkage estimators result in statistically significant increases in TG portfolio Sharpe ratio?

Again, we perform the same set of paired sample t-tests now for the TG portfolio (Table 12) – the aim of which is to maximise risk-adjusted return (Sharpe ratio) for a given set of assets. The null hypothesis states that there is no statistically significant difference between the mean Sharpe ratio estimated using the sample VCV matrix and that estimated using shrinkage estimators.

During the pre-rate hike period, all linear shrinkage estimators result in statistically significant reductions in TG portfolio annualised Sharpe ratio, on average, relative to the sample estimator. This suggests that these estimators fail to optimise the TG portfolio over the sample VCV matrix. The AV estimator is the only exception, where it achieves a statistically significant increase in Sharpe ratio relative to the sample estimator.

In the subsequent Fed rate hike period, only the AV estimator exhibits a statistically significant difference relative to the sample estimator, leading to a substantial reduction in Sharpe ratio. In contrast, none of the linear shrinkage estimators result in statistically significant deviations from the sample estimator during this period. This indicates that while the linear shrinkage estimators do not provide notable improvements, their impact on SR is not significantly different from the sample estimator under these market conditions. However, the AV estimator's significantly negative SR suggests that its risk-adjusted performance deteriorates in this period of monetary tightening.

Table 12. Paired samples t-test of TG portfolio annualised Sharpe ratio (SR). The null hypothesis states that there is no statistically significant difference between the sample VCV matrix estimated mean SR and shrinkage estimator mean SR.

Out-of-sample period	Estimator	Mean SR	t-statistic	p-value (two-tail)
Pre-rate hike period (30-Dec-19 – 30-May-22)	Sample	4.02	-	-
	CC-BI	3.78	21.29	0.0000**
	CC-BC	3.78	21.31	0.0000**
	ZC-BI	3.85	25.44	0.0000**
	ZC-BC	3.85	25.38	0.0000**
	AV	4.66	-12.74	0.0000**
Rate hike period (30-May-22 – 20-May-24)	Sample	1.74	-	-
	CC-BI	1.21	1.67	0.0986
	CC-BC	1.22	1.66	0.0991
	ZC-BI	1.51	1.77	0.0790
	ZC-BC	1.51	1.75	0.0837
	AV	-0.65	5.61	0.0000**

Note: ** denotes significance at the 1% level. * denotes significance at the 5% level.
Output using R.

4.5 Shrinkage intensity and performance statistics dynamics

Figure 21 presents the time series of λs applied by the various VCV matrix shrinkage estimators over the full out-of-sample period (30-Dec-19 – 20-May-24). Additionally, the figure overlays the US 3-month Treasury bill rate, representing the risk-free rate. This allows for an examination of how shifts in monetary policy, particularly the Fed rate hike period, may influence the magnitude of shrinkage applied by the estimators.

A key observation is the inverse relationship between the risk-free rate and shrinkage intensities. Shrinkage intensities spike sharply in early 2020, coinciding with the onset of the COVID-19 pandemic and the subsequent Fed rate cuts to near-zero levels. Among the linear shrinkage estimators, the CC-BI and CC-BC estimators show higher sustained levels of λ. This indicates a greater weighting towards the highly structured estimator, characterised by a constant-correlation model. In contrast, the ZC-BI and ZC-BC estimators exhibit relatively lower sustained levels of λ, suggesting a more balanced weighting between the sample VCV matrix and structured estimator. The latter shrinkage estimators are characterised by a zero-correlation model, where an identity matrix serves as the structured estimator.

The initiation of Federal Reserve rate hikes in Mar-22 to combat inflation results in a sharp reversal, with linear shrinkage intensities declining to pre-pandemic levels. The decline suggests that, as risk-free rates increase, the shrinkage estimators place lower weighting on the structured estimator relative to the sample VCV matrix. Similar behaviour is observed for the AV estimator, although λ in this case simply represents the average of sample correlations (which serves as a predictor of the individual asset correlations). These dynamics suggest that shrinkage constants are sensitive to shifts in monetary policy, with loose monetary conditions (low rates) leading to higher shrinkage application, while tightening cycles (rising rates) reduce λ.

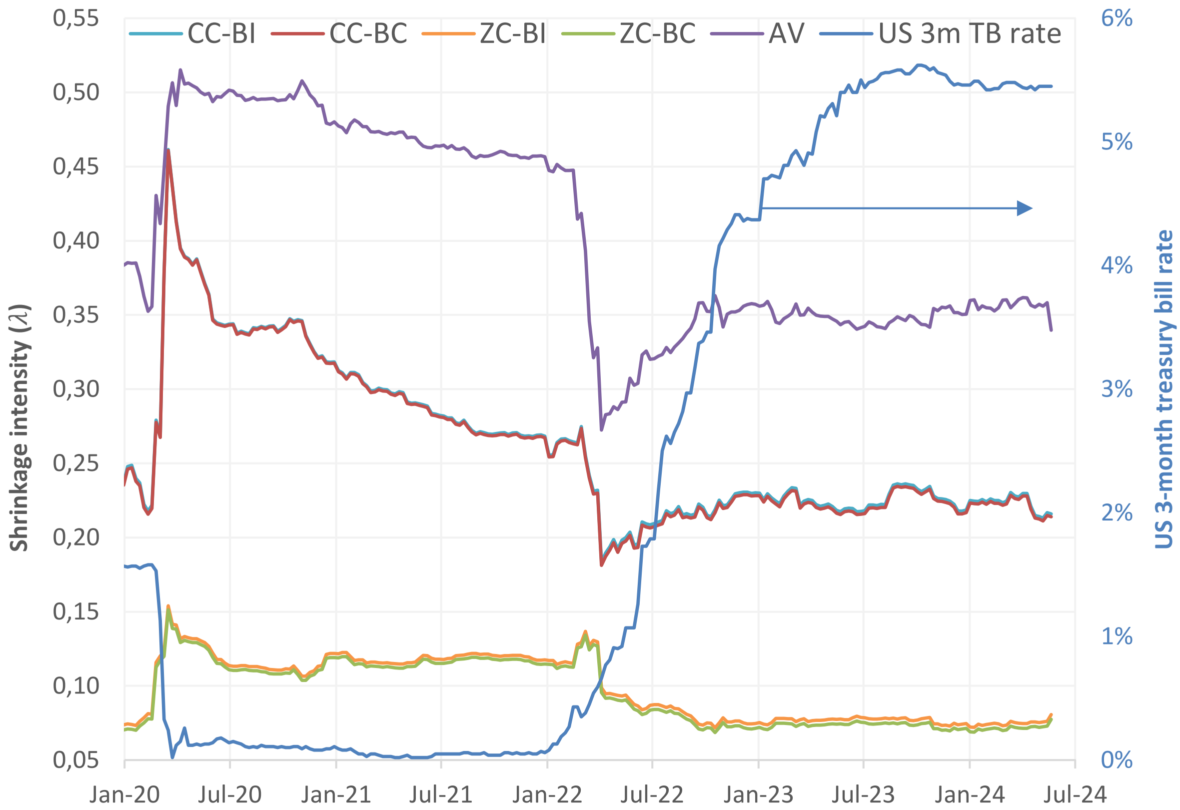


Figure 21. Time series plots of VCV matrix estimator shrinkage intensities and US 3-month treasury bill rate. The full out-of-sample period is from 30-Dec-19 – 20-May-24.

4.5.1 Global minimum variance portfolio

Figures 22-24 illustrate the relationship between GMV portfolio performance statistics and λ for all VCV matrix shrinkage estimators over the full out-of-sample period. The slope of each trend line illustrates the magnitude of the change in the performance statistic per unit increase in λ, while the R^2 value reflects the proportion of total variation in that performance measure that is explained by the λ.

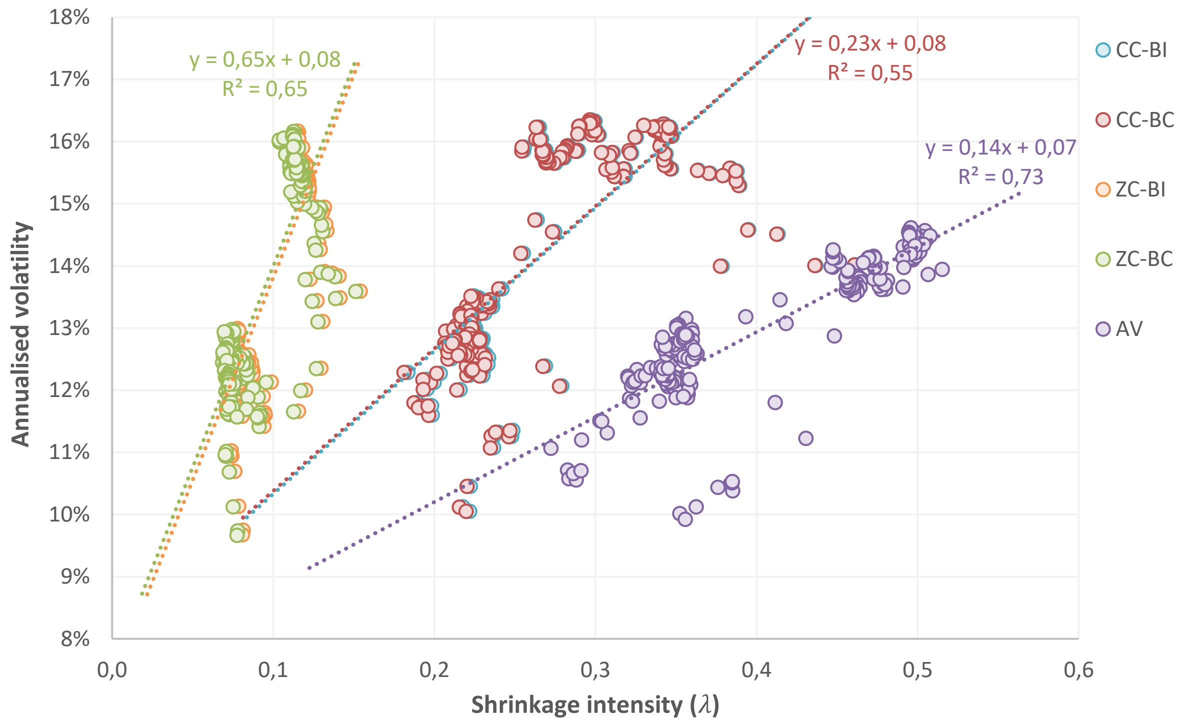


Figure 22. Scatter plot of GMV portfolio annualised volatility (Vol.) as a function of VCV matrix estimator shrinkage intensities. The full out-of-sample period is from 30-Dec-19 – 20-May-24.

Firstly, the scatterplot for annualised volatility shows a positive relationship between λ and GMV portfolio volatility across all estimators, though with varying degrees of strength. The AV estimator exhibits the most gradual increase in volatility with respect to λ, as indicated by its relatively low slope. However, its R^2 value of 0.73 shows a strong explanatory power, indicating that 73% of the variation in annualised volatility is explained by λ. In contrast, the ZC-based estimators (ZC-BI and ZC-BC) display the steepest trend line, signifying a sharp rise in volatility as λ increases. The R^2 value of 0.65 also indicates a strong explanatory power, albeit slightly weaker than that of the AV estimator. This suggests that the imposition of a zero-correlation structure leads to a more pronounced increase in risk among the shrinkage estimators. The CC-based estimators (CC-BI and CC-BC) lie somewhat between these extremes. They exhibit a moderate positive relationship between λ and volatility, with an R^2 of 0.55, suggesting that while λ moderately explains variability in risk, there is less explanatory power compared to the AV and ZC estimators.

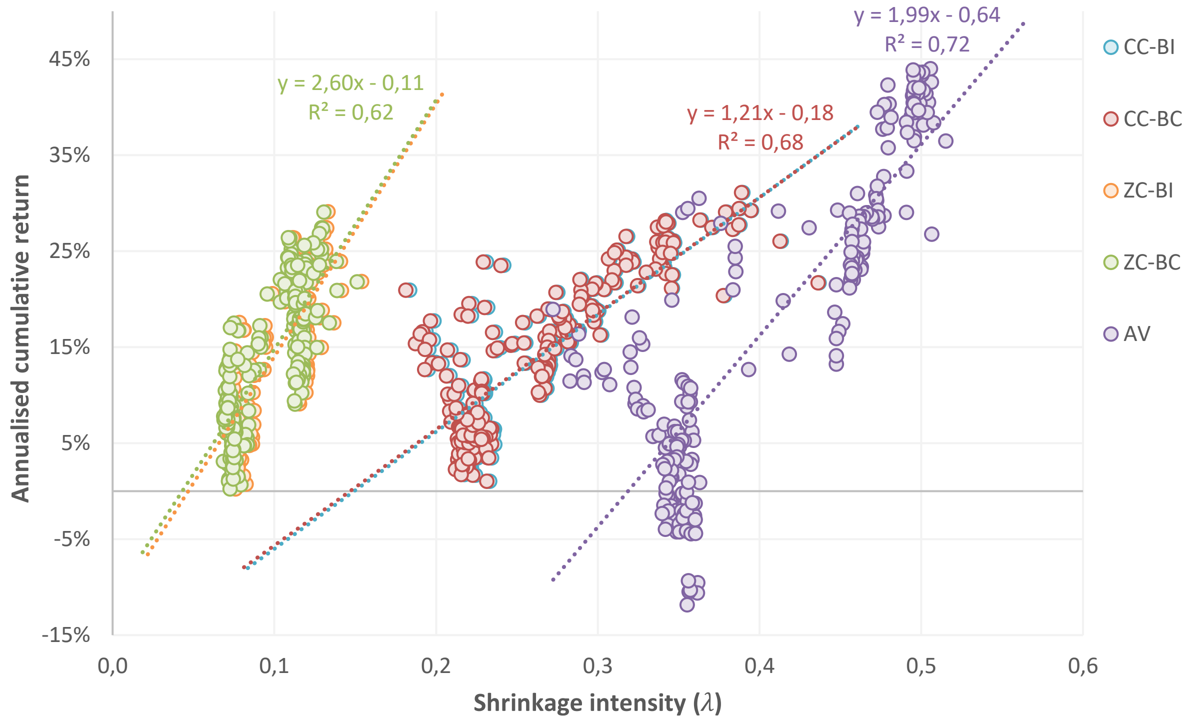


Figure 23. Scatter plot of GMV portfolio annualised cumulative return (Ret.) as a function of VCV matrix estimator shrinkage intensities. The full out-of-sample period is from 30-Dec-19 – 20-May-24.

When evaluating GMV portfolio annualised cumulative returns, the scatterplots show that higher shrinkage intensities generally correspond to increased return estimates across the board. The AV estimator again stands out; its steep slope and high R^2 of approximately 0.72 imply that a substantial portion of the return variability can be explained λ. This indicates that, as the estimator imposes a uniform correlation structure, the portfolio's expected return rises markedly – albeit with the accompanying cost of higher volatility. The CC-based estimators exhibit a more moderate positive relationship with cumulative returns, with R^2 values that are somewhat lower than those for the AV estimator. In contrast, the ZC-based estimators display the strongest positive correlation, although with the lowest R^2. This indicates that while ZC-based GMV returns are most sensitive to changes in λ, there remains more unexplained variation in this group.

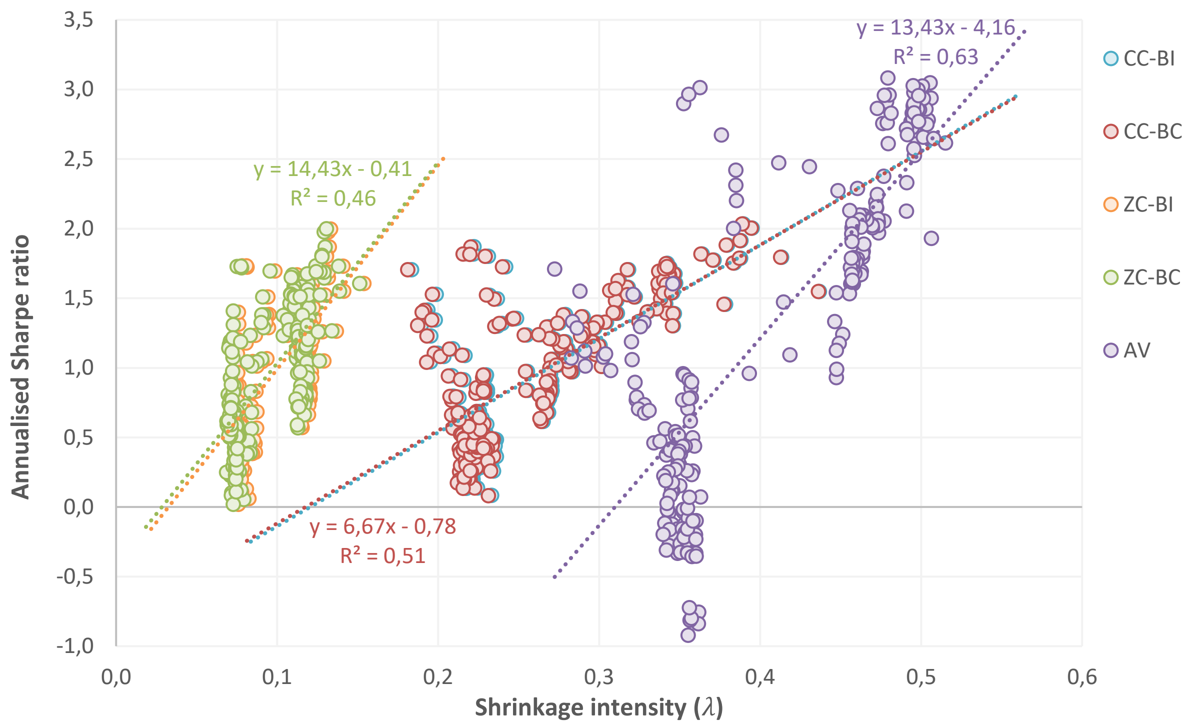


Figure 24. Scatter plot of GMV portfolio annualised Sharpe ratio (SR) as a function of VCV matrix estimator shrinkage intensities. The full out-of-sample period is from 30-Dec-19 – 20-May-24.

Lastly, the annualised Sharpe ratio shows a similar dependence on λ, though the relationships are somewhat more nuanced. For the AV estimator, the scatterplot displays a strongly positive slope with $R^2 = 0.63$. This robust association suggests that as the portfolio is characterised by a higher average of sample correlations, its risk-adjusted performance improves steadily. In contrast, the CC-based estimators have a relatively weaker correlation and explanatory power, with $R^2 = 0.51$. This indicates that while risk-adjusted performance is positively influenced by λ, the improvement in Sharpe ratio is more moderate. Again, the ZC-based estimators display the strongest positive correlation yet lowest R^2, suggesting that changes in λ have a relatively stronger but less predictable impact on risk-adjusted performance compared to the other shrinkage estimators.

Overall, a perhaps unintuitive finding is that a lower magnitude of shrinkage applied to the sample VCV matrix is generally associated with a decrease in GMV portfolio risk in this sample period. This observation may be explained by the impact of the risk-free rate on portfolio estimates, where both the sample and shrinkage estimators similarly exhibit decreased risk estimates at higher risk-free rate levels. However, as shown in Table 11, lower shrinkage intensities observed during the "rate hike period" interestingly result in marginally more statistically significant differences in sample and shrinkage estimator mean GMV portfolio risk estimates.

4.5.2 Tangency portfolio

Similarly, Figures 25-27 illustrate the relationship between TG portfolio performance statistics and λ for all VCV matrix shrinkage estimators over the truncated sample period. Due to the observed instability of the TG portfolio estimates during the period of Fed rate hikes, estimates from 30-May-22 onwards are excluded for the purpose of determining relatively stable relationships.

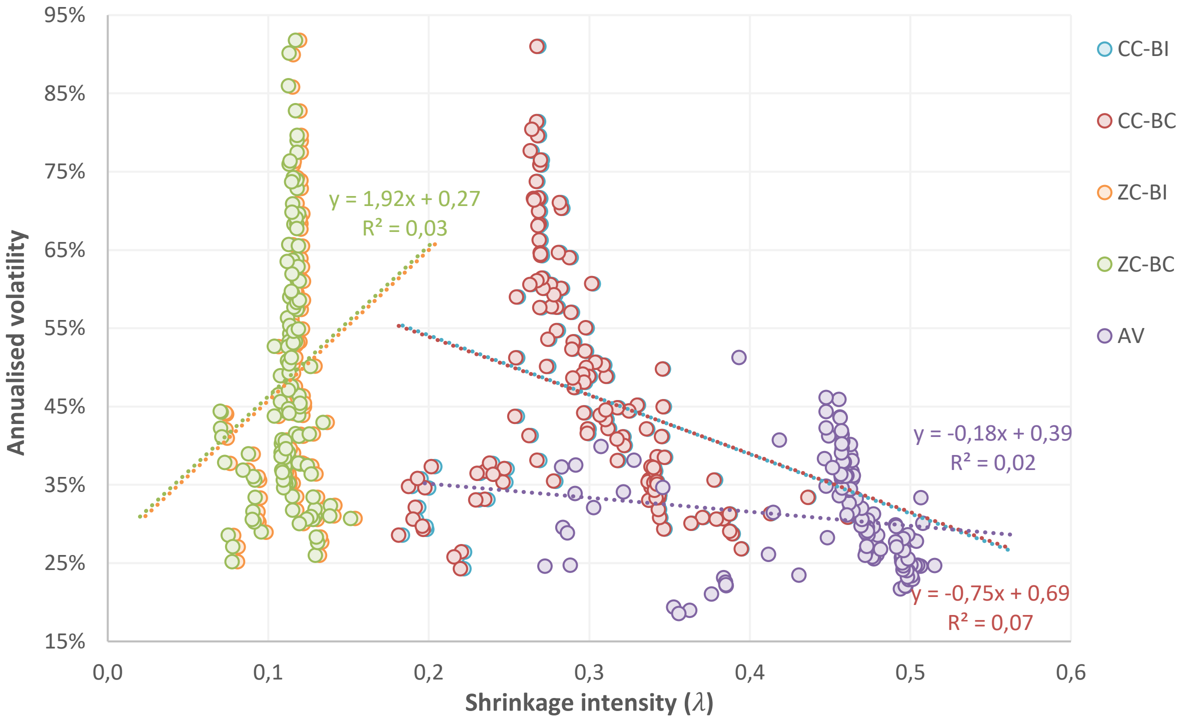


Figure 25. Scatter plot of TG portfolio annualised volatility (Vol.) as a function of VCV matrix estimator shrinkage intensities. The truncated out-of-sample period is from 30-Dec-19 – 30-May-22.

The scatterplot for annualised volatility presents mixed relationships across the different shrinkage estimators. The AV estimator exhibits a mild negative relationship between λ and annualised volatility, with an R^2 value of 0.02, suggesting a weak explanatory power. The CC-based estimators exhibit a stronger negative correlation and R^2 of 0.07, indicating slightly stronger explanatory power than the AV estimator. In contrast, the ZC-based estimators notably display a strong positive correlation. This suggests that greater shrinkage of the sample VCV matrix towards an identity matrix is associated with increased portfolio risk, whereas greater shrinkage towards a constant correlation target is associated with a declining volatility profile.

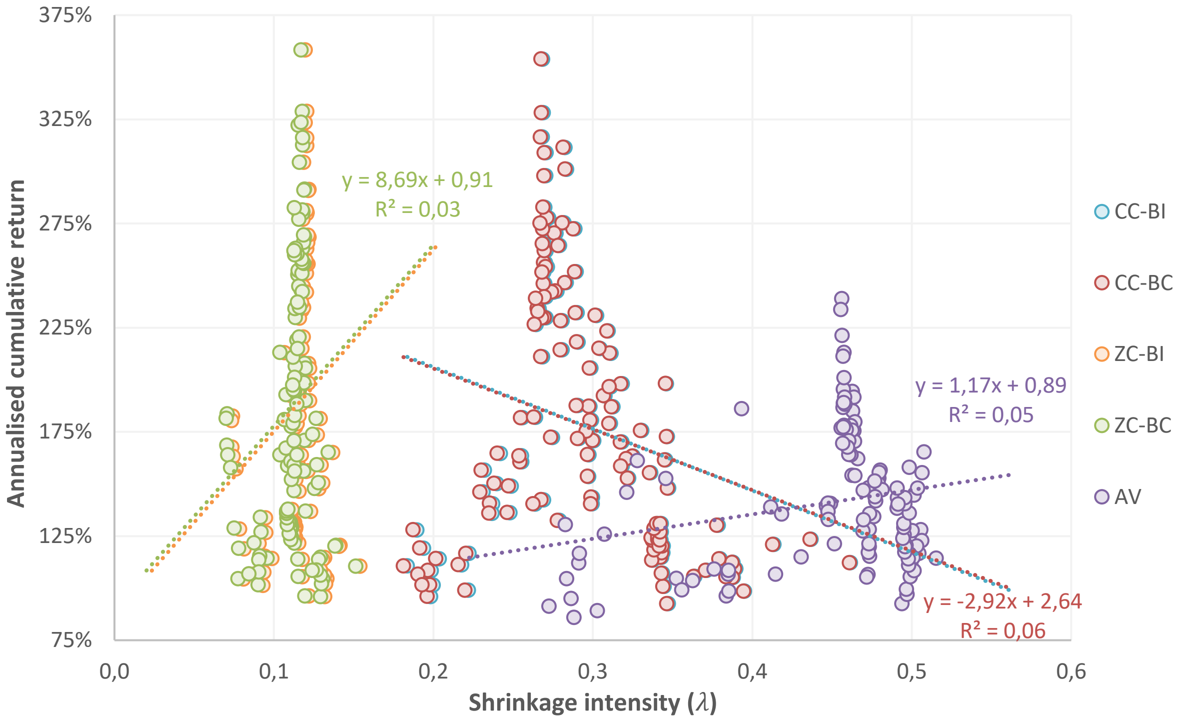


Figure 26. Scatter plot of TG portfolio annualised cumulative return (Ret.) as a function of VCV matrix estimator shrinkage intensities. The truncated out-of-sample period is from 30-Dec-19 – 30-May-22.

When considering TG portfolio annualised cumulative returns, the scatterplots again show differing patterns among the estimators. The AV estimator demonstrates a slight positive association with λ, while the CC-based estimators show the opposite. In contrast, the ZC-based estimators display the strongest positive relationship. Again, all estimators exhibit low explanatory power of shrinkage intensities for TG portfolio cumulative return. These risk and return relationships with λ translate to notable trends for the TG portfolio's key performance statistic; the AV estimator stands out with a strongly positive association between λ and annualised Sharpe ratio. This is accompanied by a significantly higher R^2 relative to the four linear shrinkage estimators. In contrast, the CC-based and ZC-based estimators exhibit very weak negative and positive relationships respectively.

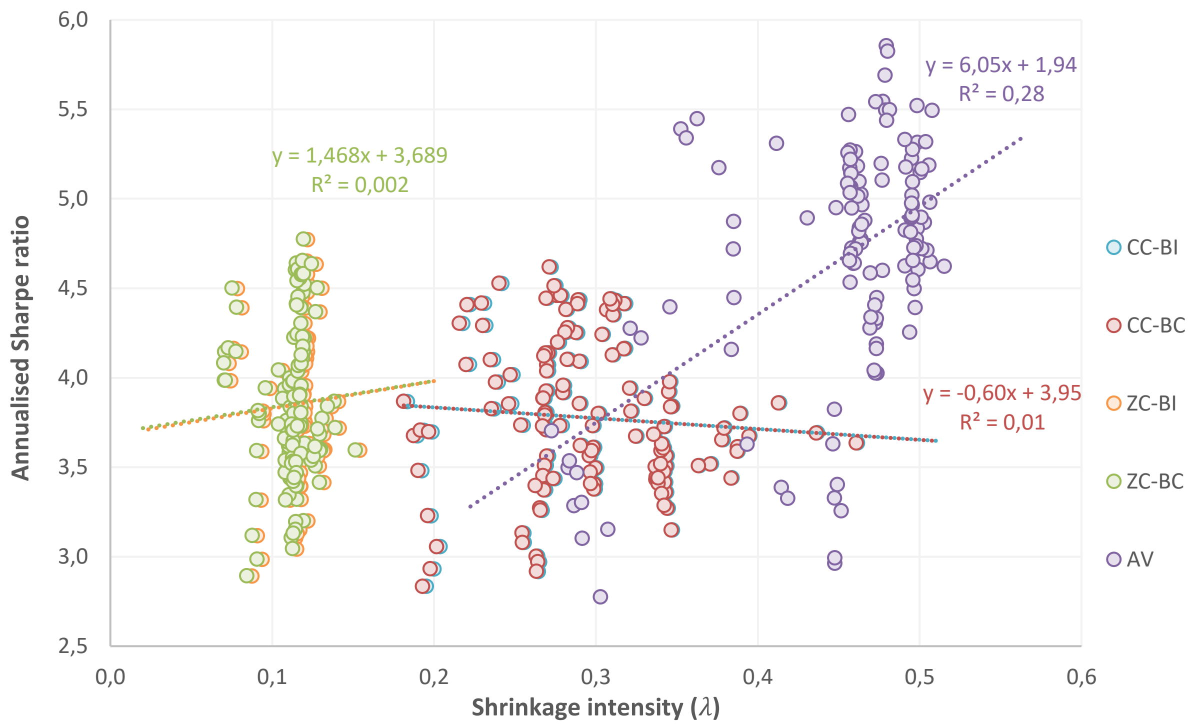


Figure 27. Scatter plot of TG portfolio annualised Sharpe ratio (SR) as a function of VCV matrix estimator shrinkage intensities. The truncated out-of-sample period is from 30-Dec-19 – 30-May-22.

Overall, in contrast to the GMV portfolio, where λ is generally associated with an increase in risk and return estimates, the TG portfolio displays more estimator-dependent patterns. In terms of optimising the TG portfolio, none of the four linear shrinkage estimators seem to exhibit strong or noticeably favourable relationships between λ and annualised Sharpe ratio.

4.6 Shrinkage estimator eigenvalues analysis

We now illustrate the application of shrinkage estimators by analysing their effect on the eigenvalues of the sample VCV matrix. Figure 27 displays the time series of the eigenvalue range – defined as the difference between the largest and smallest eigenvalues in the N = 25 asset sample set – for the full out-of-sample period (30-Dec-19 – 20-May-24).

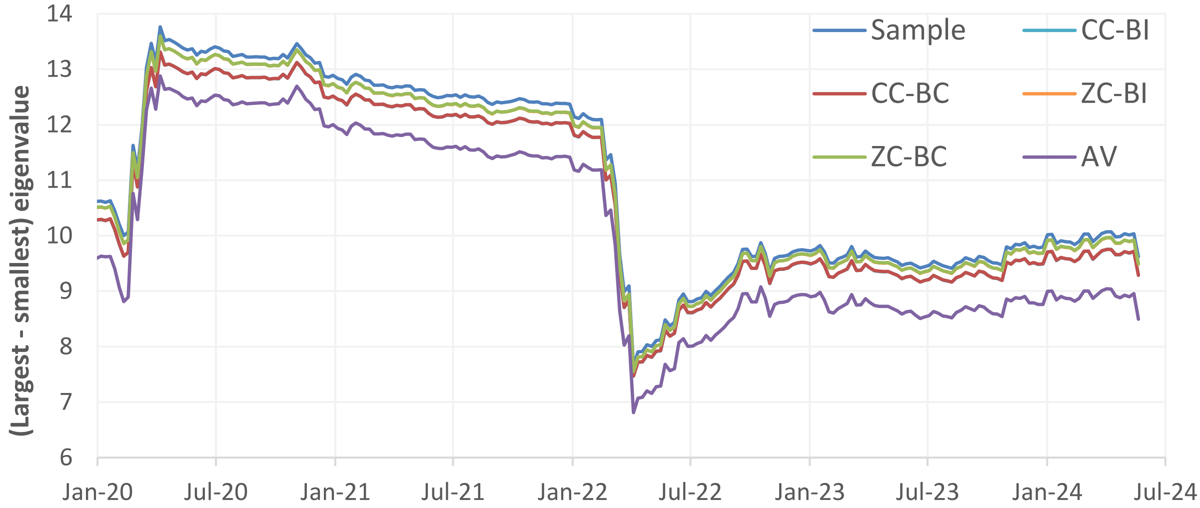


Figure 27. Time series plots of VCV matrix estimator eigenvalues; plotted values represent the weekly difference between the largest and smallest eigenvalue in the $N = 25$ asset sample set. The full out-of-sample period is from 30-Dec-19 – 20-May-24.

A smaller eigenvalue range indicates that shrinkage is effectively compressing the spectrum, by lowering the largest eigenvalues and raising the smallest ones. Throughout the full out-of-sample period, the linear shrinkage estimators consistently achieve a contraction of this range relative to the sample covariance matrix, indicating that they are operating as expected even under varying market conditions. Furthermore, the range of eigenvalues, both sample and shrunk, follow the same general trend as the shrinkage intensities of the estimators; a notable inverse relationship between the risk-free rate and range of eigenvalues is observed.

Figures 28 and 29 further illustrate the application by plotting the shrunk eigenvalues as a function of the sample eigenvalues for two key dates: 30-Dec-19 (pre-rate hike date) and 03-Oct-22 (during significant market instability following the Fed rate hikes). In both cases, the linear shrinkage estimator produces a trend that is less steep than the 45-degree sample estimator trendline, indicating that small sample eigenvalues are increased and large sample eigenvalues are decreased.

In line with the observed relationship between the risk-free rate and shrinkage intensities, the higher magnitude of shrinkage on 30-Dec-19 corresponds with a cross-over point at the 9th largest eigenvalue. The subsequent lower magnitude of shrinkage on 03-Oct-22 corresponds with a lower cross-over point at the 6th largest eigenvalue. This behaviour is consistent with the theoretical expectation that linear shrinkage can help to mitigate estimation errors by "averaging out" extreme eigenvalue coefficients.

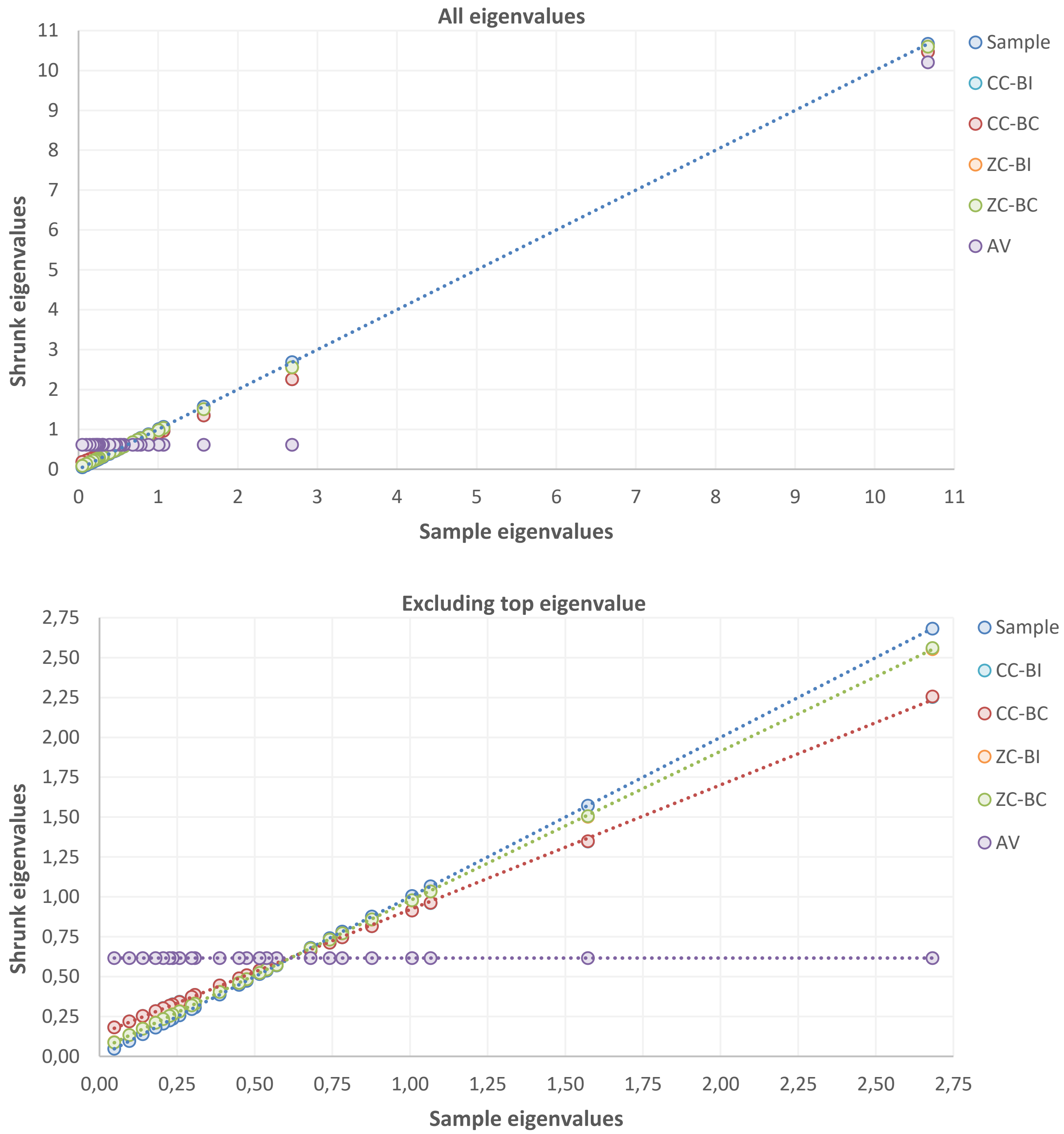


Figure 28. Shrunk eigenvalues as a function of sample eigenvalues for sample covariance matrix, linear shrinkage, and average estimators. Estimated on the out-of-sample date 30-Dec-19, for the $N = 25$ asset sample set.

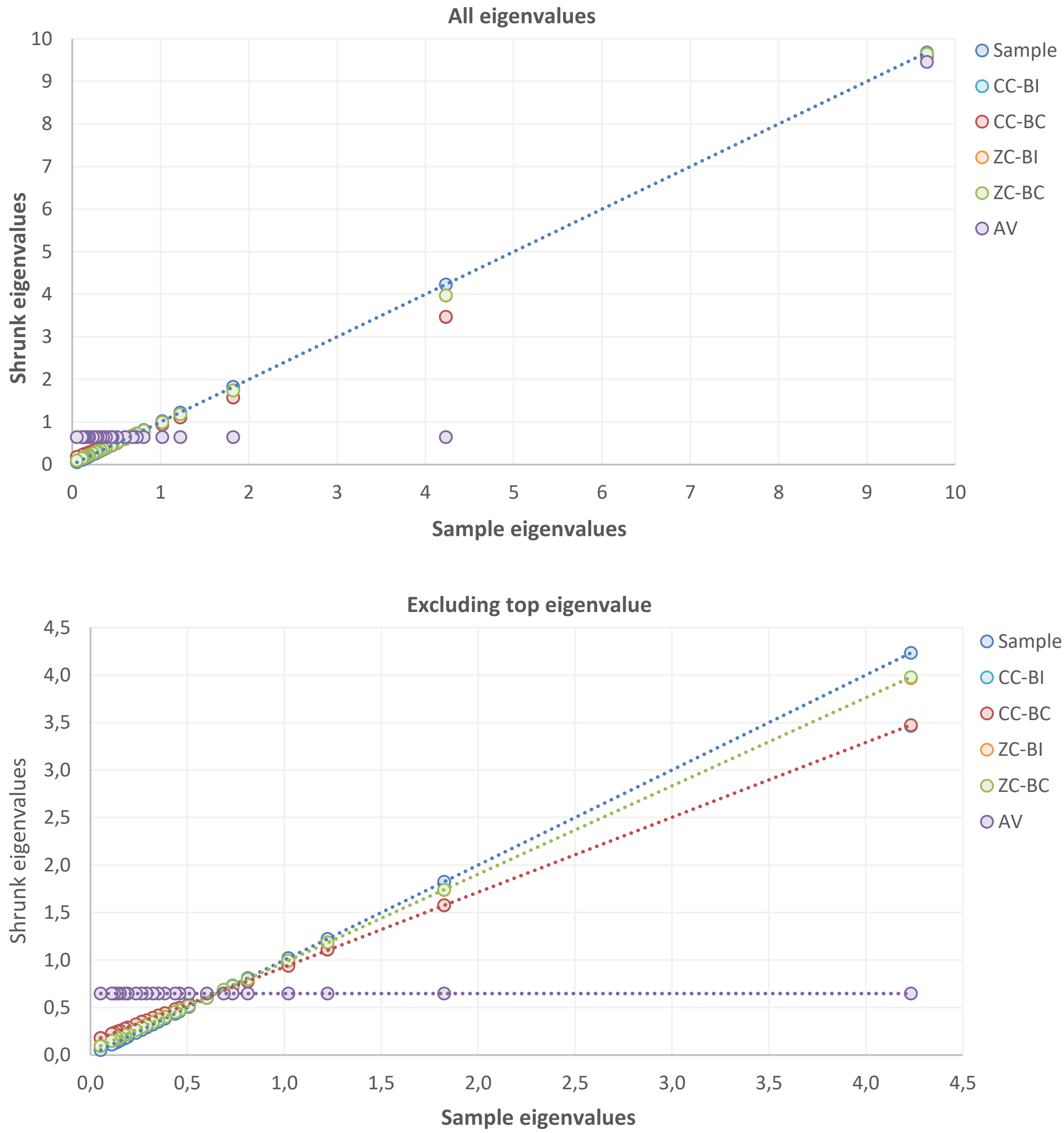


Figure 29. Shrunk eigenvalues as a function of sample eigenvalues for sample covariance matrix, linear shrinkage, and average estimators. Estimated on the out-of-sample date 03-Oct-22, for the $N = 25$ asset sample set.

Despite confirming the expected impact of VCV matrix shrinkage over the full out-of-sample period, it is observed that the four linear shrinkage estimators generally fail to improve upon and stabilise efficient frontier portfolio risk and return estimates relative to the sample VCV matrix. Barring the instability associated with consecutive increases in the risk-free rate, potential reasons for this underperformance are posited by Ledoit and Wolf (2017). They suggest that linear shrinkage may overshrink the market factor; this results in insufficient diversification of market risk and reduction of the portfolio beta. Furthermore, linear shrinkage can undershrink the few dimensions that appear safest in the sample, leading to an excessive concentration of portfolio weights in those components. Lastly, Ledoit and Wolf (2017) show that the quantity of shrinkage applied by linear estimators is optimal only on average across the entire eigenvalue spectrum. This means that while they can potentially improve overall VCV matrix conditioning, they can be suboptimal in certain segments of the spectrum, particularly when market conditions change significantly.

5 Conclusions, limitations and future work

5.1 Conclusions

Classical mean-variance optimization, as introduced by Markowitz (1952), relies heavily on the sample covariance (VCV) matrix as an input, which is inherently prone to estimation errors. These errors often lead to overly concentrated and unstable portfolio allocations, especially when faced with a sample of assets comparable in number to the historical return observations. We examine the impact of a selection of linear shrinkage estimators on two efficient frontier portfolio strategies: the Global Minimum Variance (GMV) and Tangency (TG) portfolios. Outlined in Kwan (2017), these estimators involve shrinkage of the sample VCV matrix towards both a constant correlation and zero correlation target, with the statistical issue of bias being corrected or ignored. We also include the average of sample correlations as a predictor of individual correlations as a benchmark estimator.

Our out-of-sample empirical analysis, which spanned from 30-Dec-19 – 20-May-24, shows that the average correlation (AV) estimator can significantly reduce portfolio risk and enhance risk-adjusted return on average for the GMV and TG portfolios, respectively, during relatively stable interest rate environments. However, the four linear shrinkage estimators generally provide corresponding estimates that are indistinguishable or marginally worse to those derived from sample VCV matrix over the same initial period. This finding is especially evident in the GMV portfolio, where the AV estimator achieved the lowest volatility, highest cumulative return, and best Sharpe ratio when compared to both the sample estimator and the other linear shrinkage methods. We also find that linear shrinkage estimators are generally unable to mitigate increased estimation errors associated with significant market turbulence. In particular, the Federal Reserve's series of consecutive rate hikes in 2022 lead to rapid increases in the risk-free rate, which in turn is associated with greater underperformance of the linear shrinkage estimators relative to the sample estimator. The TG portfolio is especially vulnerable to such estimation error, where extreme and impractical estimate magnitudes are observed when attempting to maximise risk-adjusted return.

Analysing λ dynamics reveals interesting performance statistic relationships. For the GMV portfolio, greater λ is generally associated with an increase in risk and return estimates, although this could be because of risk-free rate fluctuations. The TG portfolio, however, displays more estimator-dependent patterns, with all linear shrinkage estimators displaying an absence of robust relationships between λ and annualised Sharpe ratio. Lastly, our analysis of VCV matrix eigenvalues confirms that linear shrinkage behaves as theoretically expected by reducing the range of eigenvalues, thereby improving the conditioning of the covariance matrix. This benefit does not – however – always translate into superior portfolio performance relative to the sample VCV matrix, especially when market volatility is driven by macroeconomic shocks.

5.2 Assumptions and limitations

This study is based on several simplifying assumptions that facilitate the analysis of covariance matrix shrinkage and its impact on portfolio optimisation. Firstly, we assume that asset returns are normally distributed, an assumption that underpins classical mean-variance optimisation. The analysis is conducted using fully unconstrained portfolios meaning that no short-selling or asset weighting restrictions are imposed, and transaction costs are not accounted for. This choice aims for a more foundational investigation of shrinkage behaviour in isolation, although it may limit the direct applicability of the results to real-world portfolio management, where various constraints and costs are present.

Another key assumption is related to the sample period used for estimating the VCV matrices. The selected period represents a trade-off between the stationarity assumption of the joint probability distribution of returns and the need to reduce estimation errors (Kwan, 2017). By estimating the correlation matrix using past return observations, the stationarity assumption is implicitly accepted, yet no effort has been made to determine the statistically optimal sample period length. In addition, it is assumed that any revisions to the sample standard deviations of returns – aimed at potential improvements – are made at the discretion of the financial analysts. This assumption allows the study to focus on shrinkage estimation of individual correlations without confronting the usual statistical challenges associated with estimating variances and covariances.

The study also focuses exclusively on hypothetical portfolios constructed solely of stocks, largely because the constant correlation model used as one of the shrinkage targets assumes uniform pairwise correlations among the assets. This assumption restricts the applicability of the constant correlation model to multi-asset portfolios, such as those including bonds, where inter-asset correlations typically vary. In such cases, more general shrinkage targets would be required to capture the complexity of inter-asset relationships. The use of weekly price data is justified by practical portfolio management considerations. Portfolio managers typically rebalance on a weekly or monthly rather than a daily basis to avoid excessive transaction costs and operational complexity. Weekly data strike a balance by capturing meaningful price changes and correlations while reducing the noise inherent in higher-frequency data.

5.3 Future work

Future research may investigate the application of more sophisticated shrinkage estimators during the sample period of this study, such as nonlinear shrinkage techniques posited by Ledoit and Wolf (2012, 2015, 2017, 2020a, 2020b). Such extensions could provide valuable insights into whether these methods can achieve improved and more stable performance estimates relative to the sample VCV matrix during periods of market turbulence and structural shifts in risk. Future work should consider extending these findings to constrained portfolio settings by incorporating practical considerations, such as regulatory limits, long-only constraints, liquidity restrictions, and position sizing. Such extensions would more closely mirror real-world portfolio management practices and allow for a comprehensive evaluation of shrinkage techniques within applied contexts. This would ultimately contribute to developing robust portfolio optimisation strategies that are both theoretically sound and practically implementable across a range of market environments. The scope of this study may be extended beyond US equity markets to derive more robust empirical conclusions regarding the effectiveness of VCV matrix shrinkage. Investigating international markets or domestic markets such as South Africa – and expanding the analysis across a broader range of asset classes – would enhance general understanding of shrinkage methods under diverse market conditions.

6 References

Bauwens, L., Laurent, S. & Rombouts, J.V.K. 2006. Multivariate GARCH models: a survey. *Journal of Applied Econometrics*. 21(1): 79–109. DOI: https://doi.org/10.1002/jae.842.

Beerenwinkel, N., Antal, T., Dingli, D., Traulsen, A., Kinzler, K.W., Velculescu, V.E., Vogelstein, B. & Nowak, M.A. 2007. Genetic progression and the waiting time to cancer. *PLoS Computational Biology*. 3(11): 2239–2246. DOI: https://doi.org/10.1371/journal.pcbi.0030225.

Chan, L.K.C., Karceski, J. & Lakonishok, J. 1999. On portfolio optimization: forecasting covariances and choosing the risk model. *Review of Financial Studies*. 12(5): 937–974. DOI: https://doi.org/10.1093/rfs/12.5.937.

Chow, G., Jacquier, E., Kritzman, M. & Lowry, K. 1999. Optimal portfolios in good times and bad. *Financial Analysts Journal*. 55(3): 65–73. Available: https://www.jstor.org/stable/4480169 [2024, February 01].

DeMiguel, V., Garlappi, L. & Uppal, R. 2009. Optimal versus naive diversification: how inefficient is the 1/N portfolio strategy? *Review of Financial Studies*. 22(5):1915–1953. DOI: https://doi.org/10.1093/rfs/hhm075.

Dobbin, K.K. & Simon, R.M. 2006. Sample size planning for developing classifiers using high-dimensional DNA microarray data. *Biostatistics*. 8(1): 101–117. DOI: https://doi.org/10.1093/biostatistics/kxj036.

Dom, M.S., Howard, C., Jansen, M. & Lohre, H. 2024. Beyond GMV: Raising the bar for evaluating covariance matrix estimators. *SSRN Electronic Journal*. DOI: https://doi.org/10.2139/ssrn.4684926.

Elton, E.J. & Gruber, M.J. 1973. Estimating the dependence structure of share prices – implications for portfolio selection. *The Journal of Finance*. 28(5):1203. DOI: https://doi.org/10.2307/2978758.

Elton, E.J., Gruber, M.J. & Urich, T.J. 1978. Are betas best? *The Journal of Finance*. 33(5):1375. DOI: https://doi.org/10.2307/2327272.

Frost, P.A. & Savarino, J.E. 1986. An empirical Bayes approach to efficient portfolio selection. *The Journal of Financial and Quantitative Analysis*. 21(3):293. DOI: https://doi.org/10.2307/2331043.

Jagannathan, R. & Ma, T. 2003. Risk reduction in large portfolios: why imposing the wrong constraints helps. *The Journal of Finance*. 58(4):1651–1683. DOI: https://doi.org/10.1111/1540-6261.00580.

James, W. & Stein, C. 1961. Estimation with quadratic loss. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*. 4(1):361–379.

Jiang, B., Liu, C. & Tang, C.Y. 2023. Dynamic covariance matrix estimation and portfolio analysis with high-frequency data. *Journal of Financial Econometrics*. 22(2). DOI: https://doi.org/10.1093/jjfinec/nbad003.

Jobson, J.D. & Korkie, B. 1980. Estimation for Markowitz efficient portfolios. *Journal of the American Statistical Association*. 75(371): 544–554. DOI: https://doi.org/10.2307/2287643.

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

Jorion, P. 2003. Portfolio Optimization with tracking-error constraints. *Financial Analysts Journal*. 59(5):70–82. DOI: https://doi.org/10.2469/faj.v59.n5.2565.

Kwan, C.C.Y. 2008. Estimation error in the average correlation of security returns and shrinkage estimation of covariance and correlation matrices. *Finance Research Letters*. 5(4):236–244. DOI: https://doi.org/10.1016/j.frl.2008.09.001.

Kwan, C.C.Y. 2010. The requirement of a positive definite covariance matrix of security returns for mean-variance portfolio analysis: a pedagogic illustration. *Spreadsheets in Education*. 4(1). Article 4. Available: https://sie.scholasticahq.com/article/4577 [2024, February 25].

Kwan, C.C.Y. 2011. An introduction to shrinkage estimation of the covariance matrix: a pedagogic illustration. *Spreadsheets in Education (eJSiE)*. 4(3). Article 6. Available: https://sie.scholasticahq.com/article/4569 [2024, February 25].

Kwan, C.C.Y. 2017. Shrinkage of the sample correlation matrix of returns towards a constant correlation target: a pedagogic illustration based on Dow Jones stock returns. *Spreadsheets in education (eJSiE)*. 10(1). Article 3. Available: https://sie.scholasticahq.com/article/4665 [2024, February 25].

Ledoit, O. & Péché, S. 2011. Eigenvectors of some large sample covariance matrix ensembles. *Probability Theory and Related Fields*. 151(1-2):233–264. DOI: https://doi.org/10.1007/s00440-010-0298-3.

Ledoit, O. & Wolf, M. 2003. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance*. 10(5):603–621. DOI: https://doi.org/10.1016/s0927-5398(03)00007-0.

Ledoit, O. & Wolf, M. 2004a. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*. 88(2):365–411. DOI: https://doi.org/10.1016/s0047-259x(03)00096-4.

Ledoit, O. & Wolf, M. 2004b. Honey, I shrunk the sample covariance matrix. *The Journal of Portfolio Management*. 30(4):110–119. DOI: https://doi.org/10.3905/jpm.2004.110.

Ledoit, O. & Wolf, M. 2012. Nonlinear shrinkage estimation of large-dimensional covariance matrices. *The Annals of Statistics*. 40(2). DOI: https://doi.org/10.1214/12-aos989.

Ledoit, O. & Wolf, M. 2015. Spectrum estimation: A unified framework for covariance matrix estimation and PCA in large dimensions. *Journal of Multivariate Analysis*. 139:360–384. DOI: https://doi.org/10.1016/j.jmva.2015.04.006.

Ledoit, O. & Wolf, M. 2017. Nonlinear Shrinkage of the Covariance Matrix for Portfolio Selection: Markowitz Meets Goldilocks. *The Review of Financial Studies*. 30(12):4349–4388. DOI: https://doi.org/10.1093/rfs/hhx052.

Ledoit, O. & Wolf, M. 2020a. Analytical nonlinear shrinkage of large-dimensional covariance matrices. *Annals of Statistics*. 48(5):3043–3065. DOI: https://doi.org/10.1214/19-aos1921.

Ledoit, O. & Wolf, M. 2020b. The Power of (Non-)Linear Shrinking: A Review and Guide to Covariance Matrix Estimation. *Journal of Financial Econometrics*. 20(1):187–217. DOI: https://doi.org/10.1093/jjfinec/nbaa007.

Markowitz, H. 1952. Portfolio Selection. *The Journal of Finance*. 7(1):77–91. DOI: https://doi.org/10.2307/2975974.

Merton, R.C. 1972. An Analytic Derivation of the Efficient Portfolio Frontier. *The Journal of Financial and Quantitative Analysis*. 7(4):1851–1872. DOI: https://doi.org/10.2307/2329621.

Michaud, R.O. 1989. The Markowitz optimization enigma: Is "optimized" optimal? *Financial Analysts Journal*. 45(1):31–42. Available: https://www.jstor.org/stable/4479185 [2024, February 13].

Olkin, I. & Pratt, J.W. 1958. Unbiased Estimation of Certain Correlation Coefficients. *Annals of Mathematical Statistics*. 29(1):201-211.

Schäfer, J. & Strimmer, K. 2005. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Statistical Applications in Genetics and Molecular Biology*. 4(1). DOI: https://doi.org/10.2202/1544-6115.1175.

Sharpe, W.F. 1963. A Simplified Model for Portfolio Analysis. *Management Science*. 9(2):277–293. DOI: https://doi.org/10.1287/mnsc.9.2.277.

Silvennoinen, A. & Teräsvirta, T. 2009. Multivariate GARCH Models. *Handbook of Financial Time Series*. 201–229. DOI: https://doi.org/10.1007/978-3-540-71297-8_9.

Stein, C. 1956. Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*. 3(1):197–206. DOI: https://doi.org/10.1525/9780520313880-018.

vom Lehn, C. & Winberry, T. 2021. The Investment Network, Sectoral Comovement, and the Changing U.S. Business Cycle. *The Quarterly Journal of Economics*. 137(1):387–433. DOI: https://doi.org/10.1093/qje/qjab020.

Yao, J., Chang, C., Salmi, M.L., Hung, Y., Loraine, A. & Roux, S.J. 2008. Genome-scale cluster analysis of replicated microarrays using shrinkage correlation coefficient. *BMC Bioinformatics*. 9(1): 9–288. DOI: https://doi.org/10.1186/1471-2105-9-288.

Zimmerman, D.W., Zumbo, B.D. & Williams, R.H. 2003. Bias in estimation and hypothesis testing of correlation. *Psicológica*. 24(1):133-158.