ERASMUS UNIVERSITY

ERASMUS SCHOOL OF ECONOMICS

MASTER THESIS QUANTITATIVE FINANCE

An Empirical Study to Optimize Mean and Covariance Estimation in an Individual-stock and a Multi-manager Setting

Author:

Mathilde L. G. Hoving (603482)

Supervisor at APG:

Martin Prins

Supervisor:

Dr. Rasmus Lönn

Second assessor:

Dr. Onno Kleen

Date final version: June 5, 2023

Abstract

In this study, I evaluate the impact of estimation errors on the performance of APG's individual-stock and multi-manager portfolio optimization by comparing their performance with optimizations that intend to reduce the estimation errors. The empirical study entails the period October 2012 - October 2022. For the individual-stock setting the benchmark model minimizes the tracking error and hedges the factor exposure, while the multi-manager benchmark model maximizes the expected return for a given level of variance. The extensions include factor, mean, linear, and non-linear shrinkage, and dynamic covariance (DCC) models. One period out-of-sample results show that estimation errors erode the portfolio performance, and that the dynamic-covariance estimators result overall in the highest performance gain. The DCC-Factor models perform worse in the individual-stock setting when heavily accounted for factor exposure. Nevertheless, these models perform best in the multi-manager setting. Overall, none of the extensions outperforms strictly and significantly the benchmark portfolios.

Contents

1 Introduction **4**

2 Literature review **7**

2.1 Estimation of the mean vector 8

2.2 Estimation of the covariance matrix 8

3 Data **11**

4 Methodology **14**

4.1 Benchmark models 15

4.1.1 Benchmark individual-stock setting 15

4.1.2 Benchmark multi-manager setting 16

4.2 Portfolio optimization techniques 17

4.2.1 Mean estimation techniques 17

4.2.2 Factor models 18

4.2.3 Linear covariance shrinkage 19

4.2.4 Non-linear covariance shrinkage 21

4.2.5 Dynamic models 24

4.3 Portfolio performance 26

5 Results **29**

5.1 Individual-stock setting 30

5.2 Multi-manager setting 36

6 Conclusion **42**

7 Discussion **44**

A Empirical result for best hedge model with hedging coefficient 0.4 **50**

B Information ratios and CEQs for best hedge model with hedging coefficient 0.4 **51**

C Empirical result for mean-variance model with implied risk-aversion **52**

D Information ratios and CEQs for mean-variance model with implied risk-aversion **53**

E Weight statistics of individual-stock best hedge model with hedging coefficient 0.8 **54**

F Weight statistics of individual-stock best hedge model with hedging coefficient 0.4 **55**

G Weight statistics of the individual-stock minimum-variance model **56**

H Weight statistics of multi-manager model with risk-aversion of 2 **57**

I Weight statistics of multi-manager model with risk-aversion of 7 **58**

J Weight statistics of multi-manager model with implied risk-aversion **59**

1 Introduction

APG Asset Management (APG) is globally one of the largest institutional pension investors with 562 million euros asset under management (AUM) [1]. One of the key objectives of APG is to provide a "good income for today, tomorrow and beyond" to its 4.8 million Dutch participants. In this regard, APG invests 27% of its assets in equities, with 116 million euros invested in developed market equities and 44 million euros in emerging market equities. This research concentrates on optimal portfolio construction for developed market equities. Presently, APG has instituted two distinct methods for portfolio optimization. The first approach employs a minimum-variance optimization strategy with a large number of assets N relative to the number of observations T, and aims to strike a balance between limiting the exposure to factor risk relative to the benchmark model while minimizing the costs associated with hedging. This method is referred to as the best hedge model. The second optimization technique is utilized in situations where $N < T$ and involves an optimal mean-variance allocation over institutional portfolio managers (Markowitz 1952).

The mean-variance portfolio, introduced by Markowitz (1952), optimally allocates wealth across risky assets by maximizing the expected returns for a certain level of risk, referred to as the portfolio variance. The mean-variance theory aims to construct well-balanced and robust portfolios. This indicates optimal portfolios are determined by achieving the optimal return-risk trade-off and by their ability to keep this performance for varying underlying parameters such as different time periods T and estimation windows M. Markowitz (1952) constructs an optimal portfolio by maximizing the expected return minus its covariance. In order to hold this portfolio an investor needs to estimate the mean and covariance of the individual assets resulting in estimation errors in both variables. DeMiguel, Garlappi, and Uppal (2009) conclude that these estimation errors erode the portfolio's performance. As a result, the mean-variance portfolio does function as a theoretical optimal portfolio that is a foundational concept for ongoing research to construct optimally well-balanced and robust portfolios in practice.

DeMiguel, Garlappi, and Uppal (2009) empirically compare several portfolios among each other. Remarkably, they conclude the naive $1/N$ portfolio, with N the number of assets, outperforms the out-of-sample mean-variance optimization in almost all empirical settings. This raises the question of why a theoretical optimal portfolio that includes the first and second moment of assets, such as the mean-variance optimization, is outperformed by a naive portfolio. Kan and Zhou (2007) underline the results of DeMiguel, Garlappi, and Uppal (2009) and conclude that the presence of estimation risk completely erodes the theoretical optimal performance of the mean-variance portfolio. Jobson (1979) states, based on a simulation study, that the sensitivity of the mean-variance portfolio can be attributed to the estimation errors in the mean and variance. Hence, they conclude that estimated portfolios are consistently inferior to the theoretical optimal portfolio. Finally, Michaud

[1] Values as of September 1st 2022

(1989) states the mean-variance method functions as an error maximizer by over-weighting stocks with large estimated returns and small variances.

The literature has attempted to address the mean and variance estimation errors. Merton (1980) noted that when the sample size increase, the estimation errors asymptotically decrease slower in the mean than in the covariance. Thus an accurate mean estimator greatly influences the robustness of a portfolio, resulting in several studies that have attempted to address this issue. For example, Jobson (1979) introduces a shrinkage estimator based on Stein (1956) to reduce the mean estimation errors. Additionally, Pástor (2000) and Jorion (1985) further improve this approach by applying Bayesian techniques to the shrinkage estimator. An alternative approach is to avoid the estimation of the mean altogether by constraining the mean-variance optimization to a minimum-variance portfolio, which was introduced by Markowitz (1952).

To address the estimation errors in the covariance matrix, Chan, Karceski, and Lakonishok (1999) introduce factor models to capture the general covariance structure. Alternatively, Ledoit and Wolf (2004b) and Ledoit and Wolf (2012) propose linear and non-linear shrinkage covariance estimators. These estimators involve a convex combination of the sample covariance and a target matrix, with the non-linear shrinkage estimator taking a more flexible approach to the shrinkage intensity. Engle (2002) introduces a time-varying estimation of the covariance matrix in a dynamic conditional correlation (DCC) model, and Engle, Ledoit, and Wolf (2019) further combine this approach with linear and non-linear shrinkage covariance estimators. Finally, Nard, Ledoit, and Wolf (2021) blend the factor structure with the dynamic conditional correlation (DCC) models.

Despite the extensive research to reduce the mean and covariance estimation errors, there is still a lack of consensus on the optimal techniques for constructing well-balanced and robust portfolios in practice. Therefore, the primary aim of this research is to explore the consequences of sample estimation errors on the robustness of portfolios and to identify techniques that can minimize these errors. To achieve this, the study focuses on the two outlined portfolio settings: a high-dimensional individual-stock setting and a low-dimensional multi-manager setting. The study seeks to gather empirical evidence from October 2012 - October 2022 by utilizing daily individual-stock price data and monthly managed portfolio return data. The individual-stock portfolio consists of a comprehensive universe of 1197 stocks from developed markets, and the data includes information on the long positions of APG in 564 of the 1197 stocks. The benchmark portfolio, used at APG, has positive weights in the 1197 stocks from the stock universe such that it simulates the MSCI World Index. The multi-manager portfolio data consists of 8 portfolios from institutional managers active in the developed markets. The benchmark portfolio is set to the corresponding geographical MSCI Indices used at APG.

By conducting an empirical study, I examine the performance of APG's benchmark weights and optimization relative to the performance of alternative portfolio optimizations that aim to reduce estimation errors. The benchmark weights are set to the weights that APG has held over the period October 2012 - October 2022. The benchmark portfolio is set to the respective benchmark portfolios of the individual-stock and multi-manager setting, i.e., the best hedge and mean-variance model. APG's benchmark weights and APG's portfolio are not strictly similar, because the portfolio managers at APG determine the weights by not only applying an optimization model but also by studying other trends and asset characteristics. Different estimation windows are used to calculate the one-period out-of-sample performance indicators. For the individual-stock setting, the estimation window equals $M = 250$ or $M = 750$ days, while for the multi-manager setting the estimation windows equal $M = 12$, $M = 36$, or $M = 72$ months.

Several portfolio optimizations that intend to improve the mean or covariance estimation are compared with each other. For the mean estimation this paper studies the James-Stein and Pástor (2000) shrinkage estimator in the mean-variance setting, and constraining the mean-variance to a minimum-variance portfolio. For the covariance estimation, this paper studies factors models, linear and non-linear shrinkage introduced by Ledoit and Wolf (2004b), DCC models by Engle (2002), DCC shrinkage by Engle, Ledoit, and Wolf (2019) and DCC-Factor models by Nard, Ledoit, and Wolf (2021). These optimization methods are compared with each other in an empirical study. In both settings, the one-period out-of-sample excess and alpha returns with their corresponding volatility are calculated. Further performance indicators include the Sharpe and Information ratio, the certainty equivalent (CEQ), and the turnover rate. Based on the scoring of these performance indicators, this study intends to conclude whether estimation errors impact portfolio performance and which portfolio optimization techniques result in better-performing and more robust portfolios.

I find that in both settings the estimation errors in the mean and covariance erode the portfolio performance. Therefore, it is beneficial to see whether these estimation errors can be reduced by alternative optimization techniques. In the individual-stock setting, I find that depending on the hedging coefficient, DCC estimators that account for time variance in the individual assets outperform the alternative strategies. Moreover, accounting for factor exposure while minimizing the tracking error in the best hedges model leads to better-performing portfolios than simply minimizing the covariance in the minimum-variance portfolio. In the multi-manager setting, I find that mean estimation errors increase for larger estimation windows. Similar to the individual-stock setting, the dynamic covariance estimators outperform the alternative strategies, though the DCC-Factor models outperform the DCC models. Therefore, I conclude that it is favorable to account for time-variance in underlying factors in a mean-variance optimization. Nevertheless, the differences between the extensions and benchmark models do not seem to be significant at a 5% level as the differences in returns and volatilities are small.

The paper is organized in the following way. Section 2 reviews the existing literature, followed by Section 3 discussing the data used in this research. Section 4 outlines the benchmark portfolios, optimizations that extend these benchmark portfolios, and the performance indicators. Section 5 discusses the results, and Section 6 its conclusion followed by a discussion in Section 7.

2 Literature review

The seminal work of Markowitz (1952) introduced the concept of mean-variance optimization as a portfolio selection technique aimed at maximizing return for a given level of risk or minimizing risk for a given level of return. This study is concerned with the application of mean-variance optimization in the context of common stocks.

The empirical performance of mean-variance portfolios has been an ongoing topic of research in the literature. Jobson and Korkie (1980) and Michaud (1989) show that the mean-variance optimization is prone to produce volatile and inferior portfolio selections due to estimation errors in the mean and covariance. Michaud (1989) states a mean-variance optimizer overweights securities with large estimated returns, negative correlations, and small variances, while it underweights securities with small estimated returns, positive correlations, and large variances. The overweight securities are the most sensitive to estimation errors, making mean-variance portfolios error maximizers. The estimation errors erode the out-of-sample performance of the portfolio, making the theoretical optimal asset allocation inferior to the naive portfolios in practice (DeMiguel, Garlappi, and Uppal 2009).

Jobson and Korkie (1980) study the effect of estimation errors on a portfolio's performance by conducting a Monte Carlo simulation study. They use a known multivariate normal distribution of monthly returns for $N = 20$ stocks to set the optimal portfolio as the maximum Sharpe ratio on the frontier. The Monte Carlo simulations are used to calculate the expected returns and covariances. Ultimately, they compare the optimal Sharpe ratios to the estimated ones and find the estimated portfolios are consistently inferior. Although the performance of the estimated portfolios improves when short-sale restrictions are in place, they remain inferior. DeMiguel, Garlappi, and Uppal (2009) further substantiate the error maximization hypothesis of the mean-variance portfolio by demonstrating the out-of-sample mean-variance portfolio is empirically outperformed by the $1/N$ portfolio for the majority of the empirical settings. On the contrary, Kirby and Ostdiek (2010) find that the advantage of mean-variance portfolios over the $1/N$ portfolio is easily eroded after accounting for transaction costs.

This research entails distinct benchmark models for individual-stock and the multi-manager setting. The benchmark model for the individual-stock setting only involves covariance estimation, while the benchmark model for the multi-manager setting involves mean and covariance estimation. Therefore, the previous papers that aim to reduce mean estimation errors and use factor models are only applicable to optimizations in the multi-manager setting. On the contrary, the papers that reduce estimation errors in the covariance matrix are applicable to both the individual-stock and multi-manager settings.

2.1 Estimation of the mean vector

Merton (1980) highlights the significance of accurately estimating the mean vector versus the covariance matrix in portfolio theory. The author notes that the covariance estimator tends to have better accuracy in high-dimensional settings compared to the mean estimator, due to its more persistent time-variation. Consequently, various studies have sought to enhance the accuracy of mean estimation.

Stein (1956) and Pástor (2000) both introduce an estimator that shrinks the mean. Jobson (1979) introduces a James-Stein shrinkage estimator that minimizes the expected mean squared error. Thus the estimator shrinks the mean towards a more central vector by taking a convex combination of the sample mean $\hat{\mu}$ and a target vector μ_y. The target vector is usually assumed to be the same for all the assets. On the contrary, Pástor (2000) introduces a mean estimator that shrinks the sample means towards its prior variances given on the diagonal of the covariance matrix. This results in different target mean for each asset as the variance tends to differ across the assets. Moreover, the shrinkage estimator of Pástor (2000) involves an informative prior that allows a Bayesian investor to incorporate prior beliefs about the mean estimation.

Jorion (1986) introduces the Bayes-Stein estimator, which incorporates a prior distribution of returns into the target vector and computes weights from the data. Although it is biased, Jorion (1986) finds that it outperforms the classical sample mean estimator. Moreover, Jorion (1991) concludes that the Sharpe-Lintner CAPM outperforms the Bayes-Stein and sample mean estimators. Nevertheless, the minimum-variance portfolio, which minimizes the portfolio's volatility and does not involve a mean estimation, outperforms both shrunk mean portfolios. This results in the suggestion that avoiding estimation of the mean might be optimal (Jorion 1991). DeMiguel, Garlappi, and Uppal (2009) further support this finding by concluding that the minimum-variance portfolio consistently outperforms the mean-variance and Bayes-Stein portfolios across a range of empirical data sets.

2.2 Estimation of the covariance matrix

The estimation of the covariance matrix is an important aspect of portfolio optimization. Despite being unbiased, the sample covariance matrix can lead to high estimation errors, which can be exacerbated by mean-variance optimization. Moreover, the optimal weights in the mean-variance optimization as well as the minimum-variance optimization are a function of the inverse covariance matrix. For that reason, one could, in contrast to the mean, not ignore the covariance estimation by restricting a mean-variance portfolio to a minimum-variance portfolio. Thus much research has been dedicated to improving covariance estimation, with several methods being researched in this study.

One such method is factor models, used for covariance estimation in the paper of Chan, Karceski, and Lakonishok (1999). In this method, a few factors are used to capture the general covariance structure, leading to a more robust estimation of the covariance matrix and a higher Sharpe ratio. These factors can be defined upfront, for instance by including three Fama-French factors, but they can also be derived from the data following a principal component analysis (PCA). Chan, Karceski, and Lakonishok (1999) empirically study PCA for portfolio construction with domestic stocks issued on the New York Stock Exchange (NYSE) and American Stock Exchange (AMEX). They find the first factor in PCA of Connorand and Korajczyk (1988) already explains 75% of the variability of the assets. In general, Chan, Karceski, and Lakonishok (1999) conclude the estimation of covariance matrices using factor models does not necessarily improve by including more factors. Especially the market factor, equal to the "MKT-RF" in the Fama-French data library [2], explains a large part of the co-variation in the return data.

The optimal weights of the mean-variance and minimum-variance optimization are a function of the inverse covariance matrix. Stevens (1998) states the sample covariance matrix especially becomes unstable in high-dimensional settings when the number of assets exceeds the moments in time, i.e., $N > T$. Therefore, he characterizes the inverse covariance matrix directly, referred to as a node-wise regression in Callot et al. (2021). Stevens (1998) regresses the asset's excess return on a set of excess returns of other risky assets. Next, he constructs the inverse covariance matrix with these regression coefficients and residual variances. The residual variances are interpreted as the unhedgeable risk of the assets. Similarly, Meinshausen and Bühlmann (2006) estimate the inverse covariance matrix by using neighborhood selection. They use a Lasso regression to regularize the inverse covariance matrix by imposing a penalty term on the absolute value of the coefficients. Alternatively, one could use Ridge or Elastic net regularization. Ridge regularization imposes a penalty on the square of the coefficient. As a result, the Ridge regression shrinks the coefficients towards zero, while the Lasso regression sets coefficients equal to zero. The latter results in variable selection. Callot et al. (2021) empirically study the performance of node-wise regressions in comparison to factor models and shrinkage methods. They encourage direct covariance estimation as the node-wise regression approach performs well in comparison to the alternative methods.

Another approach to improving covariance estimation is through shrinkage, introduced by Ledoit and Wolf (2003). Shrinkage involves taking a convex combination of the sample covariance matrix and a shrinkage target matrix. Ledoit and Wolf (2004b) outline different target matrices, including identity, single-factor, and second-moment matrices. Linear shrinkage of the covariance matrix involves a fixed shrinkage intensity that determines the reliance on the target covariance matrix. Ledoit and Wolf (2003) explain that by taking a convex combination of a sample and target covariance matrix, the impact of estimation errors on the portfolio allocation reduces as the excessive covariance estimators are shrunk towards to the corresponding target matrix. The

[2] Kenneth R. French Data Library

results in Ledoit and Wolf (2003, 2004b) show that shrunk covariance matrices can outperform traditional sample covariance matrices, as well as factor models and covariance estimators that rely on PCA.

A challenge of linearly shrinking the covariance matrix is determining the shrinkage intensity. Ledoit and Wolf (2004b), Ledoit and Wolf (2004a), and Ledoit and Wolf (2003) introduce distinct shrinkage intensities for the different target matrices. Parameters such as the time period, sample and factor correlation and variance are used for calculating the shrinkage factor. Nonetheless, the shrinkage intensity equals the same value for all different assets. As a result, the performance gain of linearly shrunk covariance matrices erodes in high-dimensional settings where the number of assets N is large.

Ledoit and Wolf (2012) introduce a non-linear shrinkage estimator of the covariance matrix. This covariance estimator uses spectral decomposition of the sample covariance matrix to shrink the eigenvalues with varying intensities. As a result, the non-linear shrinkage estimator allows the shrinkage intensity to impose individual asset shrinkage intensities. Therefore, Ledoit and Wolf (2012) state the non-linear shrinkage estimator outperforms the linear shrinkage estimator in high-dimensional settings. However, non-linear shrinkage as in Ledoit and Wolf (2012) involves numerical estimation that can be computationally intensive. To overcome this, Ledoit and Wolf (2020) introduce an analytical method for nonlinear covariance shrinkage, which is as accurate as numerical methods while being computationally more feasible.

Further studies of Ledoit and Wolf (2022a) derive different techniques that non-linearly shrink the covariance matrix. The linear-inverse shrinkage estimator applies a smoothing of the Stein shrinkage to the eigenvectors of the sample covariance matrix. Likewise, the quadratic-inverse shrinkage estimator shrinks the eigenvectors of the sample covariance matrix with the inverse Stein and minimum-variance loss function. Finally, Ledoit and Wolf (2022a) introduce a geometric average of the linear- and quadratic-inverse shrinkage covariance estimation. All these estimators are analytical solutions and are researched in this study. The numerical estimations are not included in this study as Ledoit and Wolf (2020) find the analytical estimator is as accurate as the numerical estimator but computationally less heavy.

Alternatively, Jagannathan and Ma (2003) show that imposing short-sale constraints on the optimal weights is equivalent to shrinking the covariance matrix. They find that constrained portfolios perform similarly to portfolios constructed with factor models. In this study, all the researched portfolios involve short-sale constraints as APG only takes long positions in stocks and institutional-managed portfolios.

In the realm of financial returns, the assumption that returns are independent and identically distributed (i.i.d.) is often deemed inappropriate due to the inherent dependence and heterogeneity of the data. To address this limitation, Engle (1982) propose the utilization of auto-regressive moving average (ARMA) models for variance estimation, leading to the development of multivariate GARCH models. To further improve upon this methodology, Bauwens, Laurent, and Rombouts (2006) introduce the Dynamic Conditional Correlation (DCC) models, which offer greater flexibility in modeling the correlations and constancy between variances. The DCC model, introduced by Engle (2002), represents a multivariate extension of the GARCH model and incorporates a large-dimensional static or time-varying covariance matrix (Ledoit and Wolf 2022b). The estimation follows a two-step approach that first fits a GARCH model to univariate data and next estimates the covariance matrix based on the chosen multivariate distribution.

Similar to sample covariance estimators, computational difficulties arise by estimating the dynamic conditional correlation matrix in high-dimensional settings. Therefore, Engle, Ledoit, and Wolf (2019) introduce a linear and non-linear shrinkage estimator in the DCC model. The DCC-LS and DCC-NLS estimators first fit a univariate GARCH model. Next, linear or non-linear shrinkage is applied to the unconditional correlation matrix. Engle, Ledoit, and Wolf (2019) find the superior performance of the DCC-LS and DCC-NLS estimators in comparison to alternative estimators in high-dimensional settings where $N = 500, 1000$. Hence, the DCC-shrinkage estimation tends to improve a portfolio's robustness when the number of assets exceeds the moments in time, $N > T$. Therefore, Engle, Ledoit, and Wolf (2019) conclude dynamic covariance estimation works especially well for daily or weekly return data.

Finally, Nard, Ledoit, and Wolf (2021) propose a covariance estimator that blends the factor structure with the time-varying conditional heteroskedasticity of residuals in high-dimensional settings. This estimator accounts for time-variance over the factors and/or over the residual matrix of regressing the asset returns on the factors. Though, the residual matrix can also be set to a time-invariant residual matrix over the entire time period. The results showed that this model outperforms a number of existing models, including the non-linear shrinkage estimators of Ledoit and Wolf (2015) and the DCC-shrinkage estimator of Engle, Ledoit, and Wolf (2019).

3 Data

The data employed in this research is sourced from APG and encompasses daily stock-level and monthly multi-manager portfolio information over the 10-year time frame from October 16, 2012, to October 13, 2022. The stock-level data includes a benchmark universe of 1410 developed market stocks with 2509 trading days of stock price information, corresponding to 121 months. APG invests in a subset of this benchmark universe equal to 663 stocks. Meanwhile, the multi-manager data consists of 10 portfolios managed by institutional investors, with 121 months of portfolio return data.

The data of the individual-stock and multi-manager contain missing values which hinder optimizing over these historical returns. Therefore, I omit the missing values from the data set. This results in a benchmark universe of 1197 stocks and a portfolio universe of 564 stocks with 2509 trading days of stock price information.

In order to get the daily returns, I calculate the daily returns from the daily price information. The multi-manager data is reduced to 8 institutional-managed portfolios with 121 months of portfolio return data. The monthly returns given in the multi-manager setting are alpha returns, which means nominal monthly returns minus the return on a benchmark portfolio. In this study, the benchmark portfolio for the multi-manager setting equals geographical MSCI indices used at APG.

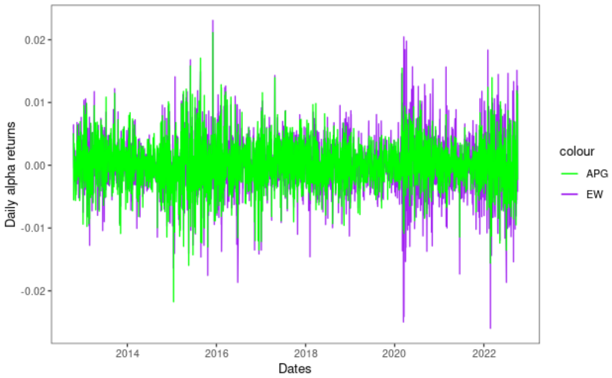
On October 12, 2022, APG's individual-stock portfolio comprised long positions in 1197 of the stocks with weights ranging from 0.002% to 9.7%. This portfolio is referred to as bmk weights, as it serves as the MSCI World Index portfolio at APG [3]. The official MSCI World Index entails long positions in 1,396 stocks with similar weight ranges, though APG excludes some stocks because of its investment beliefs. The equally weighted (EW) portfolio considers the entire stock universe of 1197 stocks. The weights in the EW portfolio are determined by the number of assets, N, and thus, these weights remain constant over time, T.

Similarly, the bmk weight portfolio in the multi-manager setting equals APG's historical weights. The weights are updated on a monthly basis and thus differ over the time period October 2012 - October 2022. The weights range from 9% to 18% over time. Furthermore, we include the EW portfolio which equally allocates the weights over the N assets for all the time periods T.

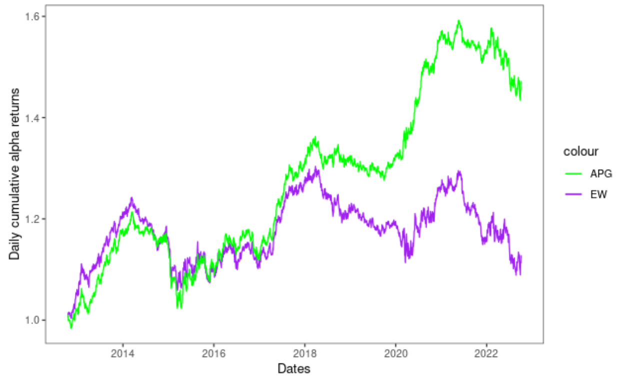
The daily and monthly (cumulative) returns for the individual-stock and multi-manager portfolios are depicted in Figures 1a and 1c. For both situations, the figure shows alpha returns, which equals the nominal returns minus the benchmark returns. For the individual-stock portfolio, this indicates I subtract the daily MSCI World Index return. The multi-manager returns are given as alpha returns. Figures 1a and 1c show a similar bandwidth for the daily and monthly returns, though I would expect a greater bandwidth for the monthly returns. Moreover, I see in figures 1b and 1d the individual-stock portfolios obtain a higher cumulative return than the multi-manager portfolios. Therefore, I state that the individual-stock setting obtains higher alpha returns than the multi-manager setting.

Another difference that I see in figures 1b and 1d is the difference between the APG and EW portfolios. This difference is larger in the individual-stock setting than in the multi-manager setting. This is likely caused by the number of assets in the two different settings. For the individual-stock setting, I deal with $N = 1197$ assets, while in the multi-manager setting, I only deal with $N = 8$ assets. As a result, the difference between APG's portfolio and the EW portfolio enlarges in the individual-stock setting.

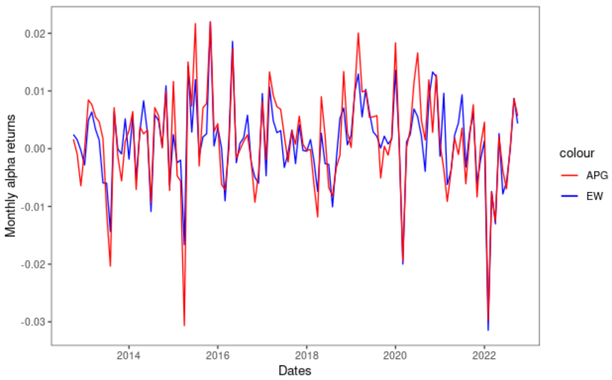
[3] https://www.msci.com/World



(a) Daily alpha returns individual-stock portfolio



(b) Cumulative daily alpha returns individual-stock portfolio



(c) Monthly alpha returns multi-manager portfolio

(d) Cumulative alpha returns multi-manager portfolio

Figure 1: Alpha returns and cumulative alpha returns of the EW and APG individual-stock and multi-manager portfolio over the period October 2012 - October 2022. The EW portfolio equally assigns weights over the assets N which equals 1197 in the individual-stock and 8 in the multi-manager setting. The APG portfolios equal the benchmark weight portfolios used at APG. The alpha returns equal the nominal returns minus the MSCI World Index in the individual-stock setting, and minus the geographical MSCI Index returns used at APG in the multi-manager setting.

Table 1 shows the APG portfolio slightly outperforms the EW portfolio indicating APG has invested in a better-performing portfolio from October 2012 - October 2022. Notably, the Information ratio of the EW portfolio in the individual-stock setting is much lower than the Information ratio of APG's portfolio. This suggests APG's portfolio outperforms the $1/N$ portfolio. Moreover, table 1 shows higher alpha returns for the individual-stock portfolio than for the multi-manager portfolio. The same results are observed in Figure 1 which shows a higher cumulative return for the individual-stock portfolio than for the multi-manager portfolio. The Sharpe ratios of the APG and EW portfolio in the individual-stock setting are close to 1. This indicates a portfolio obtains an excess return equal to its standard deviation and is seen as a balanced trade-off between the expected return and its risk level. The information ratios of the individual-stock and multi-manager portfolios are close to 0.6, indicating a lower return for the risk taken.

	Individual-stock portfolio				**Multi-manager portfolio**	
	Excess return	Sharpe ratio	Alpha return	Info ratio	Alpha return	Info ratio
APG (bmk weights)	0.1575 (0.1491)	1.0563	0.0401 (0.0603)	0.6645	0.0171 (0.0300)	0.5706
EW	0.1256 (0.1342)	0.9360	0.0142 (0.0732)	0.1941	0.0142 (0.0252)	0.5630

Table 1: Performance measures of APG and EW portfolios in the individual-stock and multi-manager setting over the time period October 2012 - October 2022. The excess return equals the nominal returns minus the risk-free rate, and the alpha return is the nominal returns minus benchmark portfolio returns. For the individual-stock the benchmark returns equal the MSCI World Index returns, for the multi-manager setting the benchmark returns equal the geographical MSCI Index returns used at APG. The Sharpe and Information ratio equal the respective return over its volatility.

4 Methodology

In order to answer the research question, I evaluate the effect of estimation techniques on the robustness of portfolios. Robustness refers to the ability of a portfolio to withstand and perform well under shocks or changes in the underlying parameters. Therefore, this study calculates the performance measures in an empirical study. We compare several portfolio optimization techniques to the bmk weights and benchmark portfolios. The bmk weights equal the APG (bmk weights) from table 1. These bmk weights equal positive positions on 13 October 2022 over the 1197 stocks in the individual-stock setting, and positive positions over 8 institutional managers that are monthly updated between October 2012 - October 2022. These allocations are used as a benchmark at APG.

In addition to the benchmark weights, this study deals with benchmark portfolios. The benchmark portfolio equals the model where the extensions are implemented, therefore they function as a validation to interpret the performance gain of the respective extension. For the individual-stock setting the benchmark model equals the best hedges optimization, while for the multi-manager setting, the benchmark model equals the mean-variance optimization. Furthermore, I use the minimum-variance optimization as a benchmark model in the individual-stock setting. This study compares extensions that aim to reduce estimation errors in the mean and the variance with its respective benchmark model, by looking at the Sharpe and Information ratio, Certainty Equivalent (CEQ), and turnover ratio.

The expected excess and alpha returns are the 1-period out-of-sample historical returns times the optimized weight vector. The estimation window determines the number of historical returns M that are used to estimate the portfolio return at $M+1$. The total estimation period equals $M+1,\ldots,T$. For the individual-stock portfolio, I want to reduce the computation time. Therefore, I calculate the daily out-of-sample returns based on the past M days once every month, assuming each month contains 21 trading days. I assume that the investor holds the optimized weights for the following 20 trading days. Due to changes in the stock returns, the proportion of the weights will change as well. In order to hold the same optimized weight proportions, I assume a certain trading volume over these 20 trading days. For the multi-manager setting, I do not make such an assumption as I deal with monthly returns, thus I test based one-period out-of-sample monthly returns.

4.1 Benchmark models

4.1.1 Benchmark individual-stock setting

APG currently optimizes the weights w in its individual-stock portfolio by limiting the amount of factor risk versus the benchmark model (2) and by minimizing the trading costs by imposing a turnover restriction (5). This translates into the best hedge function given by the objective

$$\min\ (w-w_p)'\hat{\Sigma}(w-w_p) \tag{1}$$

$$\text{s.t. } (w-w_b)'B'\Sigma_f B(w-w_b) \leq \sigma^2, \tag{2}$$

$$w \geq 0, \tag{3}$$

$$\iota' w = 1, \tag{4}$$

$$\iota'|w-w_p| \leq \tau, \tag{5}$$

where w are the weights in the constructed portfolio, w_p in APG's current portfolio, and w_b in the benchmark portfolio, referred to as bmk weights. APG's current portfolio consists of positive positions in 564 assets. The benchmark portfolio consists of positive positions in 1197 assets. Both portfolios sum up to one, meaning the full-investment constraint holds. Further, $\hat{\Sigma}$ is the sample covariance matrix of returns, Σ_f the covariance matrix of factors, B a matrix of factor exposures to these factors. Finally, σ^2 stands for the maximum amount of factor risk and ι is a vector of ones.

Equation 1 contains a covariance matrix Σ that combines factors and stock returns such that $\Sigma = B'\Sigma_f B + \Sigma_\varepsilon$ where Σ_ε is the covariance matrix of the residuals. This matrix is a diagonal matrix with the residuals on its diagonal. The first restriction in equation 2 restricts the constructed portfolio weights on deviating too much from the benchmark. Equation 3 imposes a non-negativity constraint that hinders short-selling. Equation 4 ensures the weights sum op to 1 such that all the capital is invested in stocks and equation 5 restricts on the trading volume relative to the portfolio in place at APG. In this research the maximal turnover is set to 1. This can be interpreted as a maximum trading volume of 100% to obtain the weights w compared to the current weight allocation w_p.

The output of the best hedges optimization gives a weight vector with the same dimensions as the w_p vector, thus 1×564. Hence, the optimization can only re-allocate the over APG's weights but not allocate to assets that are in the benchmark portfolio and not in the portfolio of APG. Note that the APG's portfolio is a subset of the benchmark portfolio bmk weights. As a result, the effective assets space equals $N = 564$ assets.

I can rewrite the best hedges function to the objective:

$$\min\ (1-\nu)(w-w_p)'\hat{\Sigma}(w-w_p) + \nu(w-w_b)'B'\Sigma_f B(w-w_b) \tag{6}$$

$$\text{s.t. } w \geq 0, \tag{7}$$

$$\iota' w = 1, \tag{8}$$

$$\iota'|w-w_p| \leq \tau. \tag{9}$$

The variables and restrictions are defined similarly as in equation 1, though now I account for the tracking error by defining a hedging parameter ν. This parameter measures the relative importance that is attached to the hedging objective which equals the factor exposure. When $\nu = 0.8$ imposes factor hedging with a weight of 0.8 and a current portfolio with a weight of 0.2. Shifting the hedging parameter towards zero, I allow for a larger deviation of the current portfolio and for lower factor hedging. In this study, I set $\nu = 0.8$ and $\nu = 0.4$. The hedging parameter is set to $\nu = 0.8$ in the default setting at APG. Nonetheless, the portfolio managers are assumed to have a view on their wanted hedging level and thus can change the hedging parameter accordingly.

In addition to the best hedge function, I use the minimum-variance model as a benchmark model in the individual-stock setting. The minimum-variance optimization is further explained in equation 17. Inserting extensions in the best hedge model leads to a comparison in a setting that surely incorporates factor exposure in its optimization which might impact the results. Therefore, I test the same extensions also in a minimum-variance model that simply minimizes the portfolio variance.

4.1.2 Benchmark multi-manager setting

For the multi-manager portfolio, I study a portfolio w following the mean-variance optimization introduced by Markowitz (1952). This translates into the following objective

$$\max\ w'\hat{\mu} - 0.5\gamma w'\hat{\Sigma}w \tag{10}$$

$$\text{s. t. } w \geq 0, \tag{11}$$

$$\iota' w = 1, \tag{12}$$

where w are the weights in the constructed portfolio, $\hat{\mu}$ a vector of expected returns of the assets, $\hat{\Sigma}$ the corresponding sample covariance matrix and γ the investor's level of risk aversion. Similar to equation 3 and 4, equation 11 and 12 ensure the weights sum up to one and prohibit short-selling.

I calculate the level of risk-aversion γ by considering historical returns such that

$$\gamma = \frac{w_m' \hat{\mu}}{w_m' \hat{\Sigma} w_m}, \tag{13}$$

with w_m APG's portfolio weights in the multi-manager setting, $\hat{\mu}$ and $\hat{\Sigma}$ the sample mean and covariance. However, calculating the implied level of risk aversion γ for the monthly multi-manager returns, the risk-aversion should be set $\gamma \simeq 22$. An investor with a risk-aversion of $\gamma = 1$ is considered risk-neutral. For larger values of γ, the investor is considered more risk-averse. Previous papers (DeMiguel, Garlappi, and Uppal 2009) tend to use a $\gamma = 2, ..., 10$. Therefore, $\gamma \simeq 22$ is considered extreme risk-averse. In this study, I include the settings $\gamma = 2, 7, 22$, where the latter equals the implied risk-aversion from equation 13.

Altogether, the optimizations for the individual-stock and multi-manager settings contain similarities and differences. They similarly intend to maximize the expected returns given a predetermined level of risk, constrain short-selling and sum the weights to one. The individual-stock differs from the multi-manager model in constraining the mean-variance to minimum-variance setting. Additionally, the individual-stock optimization imposes a limitation on the quantity of factor risk relative to the benchmark model and it restricts the turnover relative to the portfolio in place at APG. These additional restrictions are not enforced in the multi-manager setting.

4.2 Portfolio optimization techniques

4.2.1 Mean estimation techniques

The estimation of the mean can be improved by shrinking the sample mean to a target mean vector. Stein (1956) proposes to shrink the mean towards a target vector μ_y with a shrinkage factor δ such that

$$\mu = \hat{\mu} + \delta(\mu_y - \hat{\mu}) = (1 - \delta)\hat{\mu} + \delta\mu_y, \tag{14}$$

$$\delta_{js}^* = \min\left(1, \frac{(N-2)/T}{(\hat{\mu} - \mu_y)' \hat{\Sigma}^{-1} (\hat{\mu} - \mu_y)}\right), \tag{15}$$

where $\hat{\mu}$ equals the sample mean estimations, μ_y the target mean vector and δ the shrinkage factor. This study sets the target mean vector μ_y to the average sample mean. As a result, the volatility in the mean estimation will decline. For the shrinkage factor holds $0 < \delta < 1$. Following Brandt (2010), δ_{js}^* is the optimal shrinkage factor for the James-Stein estimator with N the number of assets, T the time period and $\hat{\Sigma}^{-1}$ the inverse sample covariance matrix.

Another method to shrink the mean estimation is using Bayesian methods. Bayesian shrinkage estimators incorporate a prior distribution of the return to determine the target mean vector. They involve an informative prior that can be based on the data or an expert belief. Pástor (2000) explains that it is likely that for smaller sample sizes the sample mean estimates obtain a higher variability. Therefore, shrinking the sample mean estimates has more effect in small sample sizes. He uses the information prior $p(\alpha) \sim N(0, \sigma_\alpha^2 I_N)$ with $\sigma_\alpha^2 I_N$ covariance matrix of the miss-pricing.

In this study, I assume α follows a normal distribution around $\alpha = 0$ and with $\sigma_\alpha = 1\%$. This results in a target mean vector equal to $E(\alpha|Y) = (1-\delta)\hat{\alpha}_{OLS}$, where δ again equals the proportion of wealth invested in riskless assets which can be interpreted as the shrinkage factor. The optimal shrinkage factor δ_p^* is given by

$$\delta_p^* = 1 - \frac{\lambda \hat{\sigma}^2}{\frac{1}{N}\sum_{i=1}^N \hat{\sigma}_i + \hat{\mu}^2} \text{ with } \lambda = 1 - \frac{N-2}{\sum_{i=1}^N \hat{\mu}_i^2}, \tag{16}$$

with N the number of assets, $\hat{\mu}$ and $\hat{\sigma}^2$ the sample mean and variance of the returns, and $\hat{\mu}_i$ the sample mean of asset i.

Instead of shrinking the mean estimation, one can restrict the mean-variance optimization to minimum-variance optimization. Markowitz (1952) first introduced the minimum-variance portfolio. Further research, including DeMiguel, Garlappi, and Uppal (2009), shows the minimum-variance portfolio outperforms the mean-variance portfolios in empirical studies. The minimum-variance model is formulated as a quadratic optimization problem subject to constraints that ensure long positions in the stocks and that the sum of the weights equal one such that

$$\min w'\hat{\Sigma}w \tag{17}$$

$$\text{s.t. } w \geq 0, \tag{18}$$

$$\iota' w = 1, \tag{19}$$

with $\hat{\Sigma}$ the sample covariance matrix and w the portfolio weights. Because the minimum-variance portfolio minimizes the portfolio variance it does not require a level of risk-aversion γ.

4.2.2 Factor models

Sharpe (1963) follows the Capital Asset Pricing Model (CAPM) to explain the variation in the assets. They regress the asset returns on just the market risk premium, while other papers include a range of factors. The idea is that many individual assets entail similar variations coming from certain factors, for example the market factor. Explaining this common variation by the corresponding factors could improve the mean and covariance estimation. The resulting estimation is given by

$$\mu = a + B\mu_f \text{ and } \Sigma = B\Sigma_f B' + \Sigma_\varepsilon, \tag{20}$$

where B equals the factor loadings, μ_f and Σ_f the estimated factor mean and covariance, and Σ_ε the residual covariance matrix. This study considers the 1-Factor and 3-Factor model introduced by Fama and French (1993). The 1-Factor model entails the market (MKT-RF) factor. The 3-Factor model entails the market (MKT-RF), size

(SMB), and value (HML) factors [4]. The market factor equals the excess return of the overall market portfolio. The size factor subtracts the returns of large-cap stocks from the returns of small-cap stocks. The value factor subtracts returns from low-value stocks from the returns of high-value stocks (Fama and French 1993).

4.2.3 Linear covariance shrinkage

Covariance estimation is sensitive to estimation errors, especially in high dimensional settings where N/T is larger than or close to one. Ledoit and Wolf (2003) employ several shrinkage methods for the covariance matrix. Inspired by Stein (1956)'s shrunk mean estimator, Ledoit and Wolf (2003) take a convex combination of the sample covariance estimator and a target matrix. Therefore, depending on the covariance target matrix, the covariance estimation equals

$$\Sigma = (1-\delta)\hat{\Sigma} + \delta\Sigma_y, \tag{21}$$

with $\hat{\Sigma}$ the sample covariance matrix, Σ_y the target covariance matrix and δ the shrinkage factor. Ledoit and Wolf (2003) explore several target matrices. In this study I include five of them: 1) one-parameter (cov1Para), 2) two-parameter (cov2Para), 3) constant correlation (covCor), 4) diagonal correlation (covDiag), and 5) one-factor covariance matrix that the single-index covariance matrix (covMarket).

Ledoit and Wolf (2004a) introduce a one-parameter matrix as covariance target matrix referred to as cov1Para. This target matrix assumes the variances of the assets are the same and the covariances are zero. As a result, the diagonal elements of the target matrix equal the common variance, and the off-diagonal elements are set to zero. The shrinkage factor equals

$$\delta^*_{para} = \max\left(0, \min\left(1, \frac{\kappa}{T}\right)\right), \tag{22}$$

where κ equals the tuning parameter and determines the extent of shrinkage applied to the sample covariance matrix $\hat{\Sigma}$. Furthermore, T equals the sample size. Instead of shrinking towards zero covariances in cov1Para, cov2Para shrinks the covariance matrix towards a matrix with constant variances and covariances. Thus, the shrinkage target Σ_y has one value for the variances on the diagonal and one value for the covariances on the off-diagonal.

In further studies, Ledoit and Wolf (2004b) introduce a constant-correlation matrix as shrinkage target Σ_y. The shrinkage target Σ_y has sample variances on its diagonal elements and constant correlations on its off-diagonal elements such that the asset's variances equal their sample variances $\hat{\sigma}^2$ and the asset's covariances equal a common correlation coefficient $\hat{\rho}_{ij}$. The shrinkage intensity equals

$$\delta^*_{cor} = \max\left(0, \min\left(1, \frac{\sum_{i\neq j} \hat{\rho}^2_{ij})/\hat{\sigma}^2}{T}\right)\right), \tag{23}$$

[4] Kenneth R. French - Data Library

where $\hat{\rho}_{ij}$ is an estimator of the population correlation coefficient between assets i and j, $\hat{\sigma}^2$ is an estimator of the sample variance, and T is the time period. The population correlation estimator $\hat{\rho}_{ij}$ is given by:

$$\hat{\rho} ij = \frac{\sum_{t=1}^{T}(r_{ti}-\hat{\mu}_i)(r_{tj}-\hat{\mu}_j)}{\sqrt{\sum_{t=1}^{T}(r_{ti}-\hat{\mu}_i)^2}\sqrt{\sum_{t=1}^{T}(r_{tj}-\hat{\mu}_j)^2}}, \tag{24}$$

where r_{ti} is the return of the ith asset at time t, $\hat{\mu}_i$ is the sample mean of the ith asset, and T equals the time period.

Similarly to the covCor method, the covDiag shrinkage target contains the sample variances on the diagonal and zeros on the off-diagonal elements of the shrinkage target Σ_y. Thus, the covDiag method shrinks the sample covariance matrix to a diagonal matrix with sample variances on the diagonal, while the covCor additionally entails constant correlations on the off-diagonal elements of the target matrix Σ_y.

Finally, Ledoit and Wolf (2003) introduce a one-factor covariance estimator as target matrix Σ_y, given by the single-index covariance matrix. The estimator is referred to as covMarket. The shrinkage estimator can be interpreted as a way to account for market exposure without the need to specify an arbitrary multi-factor structure. By following Sharpe (1963)'s single-index model, Ledoit and Wolf (2003) define the covariance matrix of the stock returns as:

$$\Phi = \sigma^2_{\text{mkt}} \beta \beta' + \Pi, \tag{25}$$

where σ^2_{mkt} equals the variance of the market returns, β the coefficient of regressing the returns on the market factor, and Π the diagonal matrix of the residuals. The estimation of Φ by running the regression of the returns on the market factor results in the estimated covariance matrix of the stock returns $\hat{\Phi}$. Ledoit and Wolf (2003) assume Φ differs from the true covariance matrix Σ.

Next, Ledoit and Wolf (2003) consider the Frobenius norm of the difference between the shrinkage estimator and the true covariance matrix, with Φ the covariance matrix of the stock returns as in equation 25, Σ the true covariance matrix, and $\hat{\Phi}$ and $\hat{\Sigma}$ their respective estimates. By taking the derivative, they find the optimal shrinkage intensity δ^*_{mkt}. The covariance estimation and its shrinkage factor are given by

$$\Sigma = (1-\delta)\hat{\Sigma} + \delta\hat{\Phi} = \sum_{i=1}^{N}\sum_{j=1}^{N}(1-\delta)\hat{\sigma}_{ij} + \delta\hat{\phi}_{ij}, \tag{26}$$

$$\delta^*_{\text{mkt}} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left[\text{Var}\left(\hat{\sigma}_{ij}\right) - \text{Cov}\left(\hat{\phi}_{ij}, \hat{\sigma}_{ij}\right)\right]}{\sum_{i=1}^{N}\sum_{j=1}^{N}\left[\text{Var}\left(\hat{\phi}_{ij} - \hat{\sigma}_{ij}\right) + (\phi_{ij} - \sigma_{ij})\right]}, \tag{27}$$

where $\hat{\Sigma}$ equals the sample covariance matrix and $\hat{\sigma}_{ij}$ its elements, $\hat{\Phi}$ the sample one-factor target covariance matrix and $\hat{\phi}_{ij}$ its elements, Φ the implied one-factor covariance matrix and ϕ_{ij} its elements, and finally Σ the covariance matrix with its elements σ_{ij}.

4.2.4 Non-linear covariance shrinkage

Ledoit and Wolf (2020) propose a non-linear shrinkage estimator for the covariance matrix. They show that the non-linear shrinkage is an improvement over the linear shrinkage method introduced by Ledoit and Wolf (2003). While the linear shrinkage method requires careful selection of the optimal shrinkage target, the non-linear shrinkage method uses the spectral decomposition of the sample covariance matrix to shrink the eigenvalues, λ_i, with varying intensities:

$$\widetilde{\Sigma} = U\Delta U', \text{ where } \Delta := \text{Diag}(\zeta_1, ..., \zeta_N), \zeta_i^* := \delta^* \hat{\mu} + (1-\delta^*)\lambda_i \tag{28}$$

with $\widetilde{\Sigma}$ the non-linear shrunk covariance matrix, U the matrix of eigenvectors u_i, Δ the diagonal matrix whose elements ζ_i are a function of the i-the sample eigenvalue λ_i, sorted in non-decreasing order. Equation 21 and 28 both intend to shrink the covariance matrix towards its grand mean, but the first with a linear and the second with a non-linear shrinkage technique. The linear shrinkage technique in equation 21 imposes a convex combination between the sample and the target matrix, while the non-linear shrinkage technique in equation 28 implies the convex relation on the eigenvalues of the covariance matrix. As a result, the non-linear shrinkage technique imposes an individual shrinkage intensity, in contrast to the general shrinkage intensity in the linear shrinkage estimator. The individual shrinkage factor increases the shrinkage for small eigenvalues and reduces it for large eigenvalues. This makes non-linear covariance shrinkage especially useful for large-dimensional covariance matrices where N/T is close to one or even larger than one because the eigenvalues are expected to diverge more in that case. The paper of Ledoit and Wolf (2020) outlines an analytical solution for the non-linear shrinkage which makes the optimization computationally more convenient.

Ledoit and Wolf (2022a) deduce the non-linear shrinkage estimators from the Stein (1956) loss function. The resulting optimization is given by

$$\underset{\widetilde{\Delta} \text{ diagonal}}{\text{argmin}} \mathscr{L}\left(\Sigma, U\Delta U'\right) \tag{29}$$

$$\text{with Stein's Loss function } \mathscr{L}^{\text{S}}\left(\Sigma, \widetilde{\Sigma}\right) := \frac{1}{N}\text{Tr}\left(\Sigma^{-1}\widetilde{\Sigma}\right) - \frac{1}{N}\log\det\left(\Sigma^{-1}\widetilde{\Sigma}\right) - 1, \tag{30}$$

where Σ equals the true covariance matrix and $\widetilde{\Sigma} = U\Delta U'$ the non-linear shrunk covariance matrix. From equation 28 I derive $\widetilde{\Sigma} = \sum_{i=1}^{N} \zeta_i u_i u_i'$. Therefore, the solution of the optimization problem is given by

$$\Delta := \text{Diag}\left(\tilde{\zeta}_1, \ldots, \zeta_N\right) \text{ where } \tilde{\zeta}_i := \frac{1}{u_i'\Sigma^{-1}u_i} \text{ for } i = 1, \ldots, N, \tag{31}$$

with u_i the eigenvector corresponding to asset i. Nonetheless, this solution is not observable as the true covariance and its inverse Σ^{-1} are not observable. As a result, the elements ζ_i are estimated with a bona fide estimator. The linear-, quadratic- and geometric-inverse shrinkage estimators follow from Stein's loss function as formulated in equation 30, but have their distinct bona fide estimator for ζ_i.

Ledoit and Wolf (2022a) introduce a linear-inverse shrinkage (LIS) estimator. This estimator is a simple smoothing of the Stein shrinker that linearly shrinks the eigenvalues of the covariance matrix. The shrinkage factor ζ_i for $i = 1 \dots N$ equals

$$\left(\hat{\zeta}_i^{\text{LIS}}\right)^{-1} := \max\left[\lambda_N^{-1}, \frac{1}{N}\sum_{j=1}^{N}\lambda_j^{-1}\frac{\lambda_j^{-1}-x}{\left(\lambda_j^{-1}-x\right)^2+h^2\lambda_j^{-2}}\right] \text{ and } \widehat{\Delta}^{\text{LIS}} := \sum_{i=1}^{N}\widehat{\zeta}_i^{\text{LIS}} \cdot u_i u_i', \tag{32}$$

with h the regularization parameter which controls the degree of smoothing. I require h to be strictly positive as for $h = 0$ the fraction equals $\frac{1}{\lambda_i^{-1}-x}$. The maximum function in 32 implies all $\hat{\zeta}_i^{\text{LIS}}$ should be greater than or equal to $\lambda_N^{-1} = \min_{i=1,\dots,N}(\lambda_i)$ which is strictly positive. Moreover, the local shrinkage works by attracting the inverse eigenvalues toward each other. If $\lambda_i^{-1} < 1$, λ_i^{-1} is slightly below λ_j^{-1} for $i \neq j$. As a result, $\hat{\zeta}_i^{-1}$ tends to go up. In the opposite way, $\hat{\zeta}_i^{-1}$ tends to go down if $\lambda_i^{-1} > 1$. Altogether, the LIS estimator combines linear and non-linear shrinkage optimization techniques. The sample covariance is scaled with an identity matrix, resulting in linearly shrunk diagonal elements in the estimated covariance matrix. The off-diagonal elements are estimated with non-linear shrinkage and varying shrinkage intensities.

Section 4 of Ledoit and Wolf (2022a) introduces the quadratic-inverse shrinkage (QIS) estimator. The QIS estimator assumes the covariance matrix is a weighted combination of the sample covariance matrix and a target covariance matrix. The optimal weights are determined by minimizing the quadratic loss function. Ledoit and Wolf (2022a) state the QIS estimator can be seen as a more advanced version of the LIS estimator because it incorporates the estimated covariance structure by estimating the off-diagonal elements of the covariance matrix.

In order to derive a shrinkage estimator that depends on observable return data, I need two concepts: the conjugate and the amplitude. The conjugate, derived by Gabor (1946), is anti-involutive. This means that the conjugate of the conjugate is simply the original function. For the QIS function Ledoit and Wolf (2022a) derive the conjugate of the smoothed Stein shrinker. The smoothed Stein shrinker $\widehat{\theta}(x)$ and its conjugate $\widehat{\theta}^*(x)$ equal:

$$\widehat{\theta}(x) := \frac{1}{N}\sum_{j=1}^{N}\lambda_j^{-1}\frac{\lambda_j^{-1}-x}{\left(\lambda_j^{-1}-x\right)^2+h^2\lambda_j^{-2}} \quad , \tag{33}$$

$$\widehat{\theta}^*(x) := \frac{1}{N}\sum_{j=1}^{N}\lambda_j^{-1}\frac{h\lambda_j^{-1}}{\left(\lambda_j^{-1}-x\right)^2+h^2\lambda_j^{-2}} \quad , \tag{34}$$

with h the bandwidth parameter that controls for the degree of smoothing. Gabor (1946) defines the amplitude $\mathscr{A}_{\widehat{\theta}}^2(x)$ by combining the smoothed Stein shrinker with its conjugate in a quadratic way resulting in

$$\mathscr{A}_{\widehat{\theta}}^2(x) = \left[\frac{1}{N}\sum_{j=1}^{N}\lambda_j^{-1}\frac{\lambda_j^{-1}-x}{\left(\lambda_j^{-1}-x\right)^2+h^2\lambda_j^{-2}}\right]^2 + \left[\frac{1}{N}\sum_{j=1}^{N}\lambda_j^{-1}\frac{h\lambda_j^{-1}}{\left(\lambda_j^{-1}-x\right)^2+h^2\lambda_j^{-2}}\right]^2 . \tag{35}$$

The quadratic loss functions can be derived from Frobenius, Inverse Stein's, and Minimum-variance loss functions. This study focuses on the Inverse Stein's $\mathscr{L}_t^{IS}$ loss function given by

$$\mathscr{L}^{IS}\left(\Sigma,\widetilde{\Sigma}\right) := \frac{1}{N}\operatorname{Tr}\left(\Sigma\widetilde{\Sigma}^{-1}\right) - \frac{1}{N}\log\det\left(\Sigma\widetilde{\Sigma}^{-1}\right) - 1, \tag{36}$$

with Σ and $\widetilde{\Sigma}$ the true and non-linear shrunk covariance matrix, respectively. To estimate the QIS estimator, I make use of the convergence in probability to a nonrandom limit as the number of assets N goes to infinity. This limit is minimized if $\widetilde{\zeta}_n(\lambda_{n,i}) = \widehat{\zeta}_i$. The estimated shrinkage factor ζ_i for $i = 1,\dots,N$ equals

$$\left(\widehat{\zeta}_i^{\text{QIS}}\right)^{-1} = \left(1-\frac{N}{T}\right)^2 \lambda_i^{-1} + 2\frac{N}{T}\left(1-\frac{N}{T}\right)\lambda_i^{-1}\widehat{\theta}_t\left(\lambda_i^{-1}\right) + \left(\frac{N}{T}\right)^2 \lambda_i^{-1}\mathscr{A}_{\widehat{\theta}}^2\left(\lambda_i^{-1}\right), \tag{37}$$

with T the moments in time, N the number of assets with $i = 1,...,N$, $\widehat{\theta}(x)$ the smoothed Stein shrinker, and $\mathscr{A}_{\widehat{\theta}}^2(x)$ the squared amplitude.

Ledoit and Wolf (2022a) define the concentration ratio as N/T. They state that the LIS estimator assumes the concentration converges to a limit $c \in (0,1)$, indicating the number of assets N cannot exceed the number of periods in time T. On the contrary, the QIS estimator is defined for settings where $N > T$. They show graphically that the Stein shrinker works on the mid range target and the amplitude on the high range target. Therefore, the quadratic weights in the QIS estimator function well in high-dimensional settings where the number of assets equals or is larger than the moments in time.

The final non-linear shrinkage covariance estimator is derived under the Symmetrized Kullback-Leibler loss function. It can be viewed as geometrically averaging linear-inverse shrinkage (LIS) with quadratic-inverse shrinkage (QIS) which results in

$$\widehat{\Sigma}^{\text{GIS}} := \sum_{i=1}^{N} \sqrt{\widehat{\Sigma}^{\text{LIS}} \times \widehat{\Sigma}^{\text{QIS}}} \cdot u_{t,i}u_{t,i}', \tag{38}$$

with $\widehat{\Sigma}^{\text{LIS}}$ and $\widehat{\Sigma}^{\text{QIS}}$ the linear and quadratic inverse shrinkage estimator, respectively.

Section 5 in Ledoit and Wolf (2022a) explains the LIS and GIS function only functions in settings where $N < T$. For $N > T$ one needs to estimate the sample eigenvalues which is only feasible through a numerical approach (O. Ledoit and Péché 2011). The study explains that if $N > T$ the first $N - T$ null sample eigenvalues are undetectable. These eigenvalues are only detectable if there are $N - T + 1$ null sample eigenvalues which is not the case. As a result, the estimator of ζ_i for the first $N - T$ eigenvalues is undetectable.

$$\hat{\zeta}_i^{-1} = \begin{cases} \left(\frac{N}{T}-1\right) \times \frac{1}{n}\sum_{j=(N-T+1)}^{N} \lambda_j^{-1} & \text{for } i = 1,\dots,N-T \\ \lambda_j^{-1}\mathscr{A}_{\widehat{\theta}}^2(\lambda_j^{-1}) & \text{for } i = N-T+1,\dots,N \end{cases} \tag{39}$$

The problem arises for $i = 1,\dots,N-T$ as $\hat{\zeta}_i^{-1}$ is a linear function function λ_j^{-1}. The QIS estimator is not a linear function in λ_j^{-1}, therefore able to shrink matrices for $N > T$.

4.2.5 Dynamic models

So far the covariance estimators assume constant estimation over time. However, constant covariance estimation might be a too strong assumption, especially for weekly or daily returns. Engle (2002) studies the dynamic conditional correlation (DCC) model as a covariance estimator. The DCC model is a generalization of the constant conditional correlation (CCC) model introduced by Bollerslev (1990). He defines the CCC estimator as

$$r_t|F_{t-1} \sim N(0,\Sigma_t), \tag{40}$$

$$\Sigma_t = D_t R D_t, \tag{41}$$

where Σ_t equals the conditional covariance matrix, R a correlation matrix containing the conditional correlations, and D_t a diagonal matrix containing the asset's volatility. Both Σ_t and D_t are time-variant. Engle (2002) generalizes equation 40 by allowing the correlation matrix R_t to vary over time. This means the DCC model is given by

$$\Sigma_t = D_t R_t D_t, \tag{42}$$

$$D_t = \text{diag}(\omega_i) + \text{diag}(\upsilon_i) r_{t-1} r'_{t-1} + \text{diag}(D^2_{t-1}), \tag{43}$$

$$R_t = \text{diag}(Q_t)^{-1/2} Q_t \text{diag}(Q_t)^{-1/2}, \tag{44}$$

$$Q_t = C(1-\alpha-\beta) + \alpha s_{t-1} s'_{t-1} + \beta Q_{t-1}, \tag{45}$$

$$s_t = D_t^{-1} r_t \tag{46}$$

where C equals the covariance matrix of s_t, the unconditional correlation matrix of the returns r_t. Moreover, (α, β) are the DCC parameters that are related to the univariate GARCH(1,1) model. The DCC model is estimated in three steps. First, the vector s_{t-1} is estimated by dividing the individual asset returns at time $t-1$ by their estimated conditional standard deviations. Secondly, the correlation matrix C is estimated. In the third and final step, the DCC parameters (α, β) are estimated following equation 45.

An occurring problem in the DCC model of Engle (2002) is the curse of dimensionality that arise in steps 2 and 3, because C is a large-dimensional $N \times N$ matrix. Therefore, shrinkage is applied to the unconditional correlation matrix C in step 2 of the DCC estimation. Ledoit, Wolf, and Zhao (2019) apply the linear shrinkage estimator of Ledoit and Wolf (2004a) to the unconditional correlation matrix C.

$$\widehat{C} := \sum_{i=1}^{N} \left[\delta\bar{\lambda} + (1-\delta)\lambda_i\right] u_i u_i' \text{ with } \bar{\lambda} := \frac{1}{N}\sum_{i=1}^{N}\lambda_i, \tag{47}$$

where λ_i equals the eigenvalue of asset $i = 1,\ldots,N$, and δ the shrinkage factor. Similar to the covariance estimator, the shrinkage intensity δ is constant over all the assets for the linear shrinkage estimators.

On the contrary, the non-linear shrinkage estimator uses different shrinkage intensities for all the assets. Ledoit, Wolf, and Zhao (2019) study the dynamic constant correlation non-linear shrinkage (DCC-NLS) covariance estimator.

$$\hat{C} = \sum_{i=1}^{N} \tilde{\lambda}_i(\tilde{\tau}) u_i u_i', \tag{48}$$

where I follow Ledoit and Wolf (2020) to compute the shrunk eigenvalues $\tilde{\lambda}(\tilde{\tau}) = (\tilde{\lambda_1}(\tilde{\tau}), ..., \tilde{\lambda_N}(\tilde{\tau}))$, where $\tilde{\tau} = \text{argmin}_{t \in [0,\infty)} \frac{1}{N} \sum_{i-1}^{N} [q_{N,T}^i(t) - \lambda_i]^2$. Two important advantages of the DCC-NLS model are (1) it does not require a normality condition and (2) it can estimate covariance matrices for $N > T$.

Estimating time-variant covariance matrices for portfolios that contain many assets N becomes computationally heavy. Nard, Ledoit, and Wolf (2021) blend the time-varying covariance estimation with a factor model to reduce this computation time. Taking the same setting as equation 20, the factor model explains the variation in returns with the three Fama-French factors. The factor exposure B and the covariance matrices Σ_ε are both time-invariant. The dynamic conditional correlation factor (DCC-FF) model starts by calculating the factor exposures B and the corresponding residual matrix Σ_ε such that

$$\Sigma_t = B' \Sigma_f B + \Sigma_{\varepsilon,t}, \tag{49}$$

with $\Sigma_{\varepsilon,t}$ a diagonal time-invariant covariance matrix of the residuals. Nard, Ledoit, and Wolf (2021) assume a time-variant residual matrix by calculating Σ_t, though I assume a constant covariance matrix of residuals Σ_ε. Nard, Ledoit, and Wolf (2021) even implement time-varying factor exposure, though they find that this implementation does not improve the estimation performance.

This study applies the factor models in the DCC models by first estimating the exposure of the individual-stock returns to the three Fama and French factors. The factor exposure and the residuals are assumed to be constant over the time period T. Next, the DCC model is used to estimate the time-varying covariance matrix of the three factors which yields a $3 \times 3 \times N$ dynamic covariance matrix. The return data accounts for these dynamics by multiplying the DCC-FF covariance matrix by the time-invariant factor exposure B and by adding the time-invariant residual matrix. The dynamic conditional correlation models used to estimate the time-variant covariance matrix of the factors use no, linear or non-linear shrinkage. This results in the DCC-FF, DCC-FF-LS, and DCC-FF-NLS estimators.

#	Method	Description
Benchmark models		
Both benchmark models use the sample mean $\hat{\mu}$ and covariance matrix $\hat{\Sigma}$		
1	Best hedges	Minimizes the tracking error while accounting for factor risk and restricting the turnover
2	Mean-variance	Maximizes the expected return, while reducing the portfolio variance for a given level of risk-aversion
Mean shrinkage		
3	Minimum-variance	Minimizes the portfolio variance
4	James-Stein	Shrinks sample means to target mean μ_y given by the mean of the sample means
5	Pastor	Shrinks sample means using an informative prior that entails a level of misspricing α
Factor model		
6	1FF	Estimates the mean and covariance by regressing the returns on the Fama-French MKT-RF factor
7	3FF	Estimates the mean and covariance by regressing the returns on the 3 Fama-French factors (MKT-RF, SMB, HML)
Linear shrinkage		
8	cov1Para	Shrinks towards matrix with constant variances on the diagonal and zero covariances on the off-diagonal
9	cov2Para	Shrinks towards matrix with constant variances on the diagonal and constant covariances on the off-diagonal
10	covCor	Shrinks towards matrix with sample variances on diagonal and constant covariances on the off-diagonal
11	covDiag	Shrinks towards matrix with sample variances on diagonal and zero covariances on the off-diagonal
12	covMarket	Shrinks towards single-index covariance matrix
Non-linear shrinkage		
13	LIS	Linear-inverse shrinkage estimator using Stein's loss function
14	QIS	Quadratic-inverse shrinkage estimator using the Inverse Stein's loss function
13	GIS	Geometric average of the LIS and QIS shrinkage estimators
Dynamic models		
16	DCC	Dynamic covariance estimation with sample correlation estimation
17	DCC-LS	Dynamic covariance estimation with linear shrinkage correlation estimation
18	DCC-NLS	Dynamic covariance estimation with non-linear shrinkage correlation estimation
19	DCC-3FF	Dynamic 3-Factor model with sample correlation estimation
20	DCC-3FF-LS	Dynamic 3-Factor model with linear shrinkage correlation estimation
21	DCC-3FF-NLS	Dynamic 3-Factor model with non-linear shrinkage correlation estimation

Table 2: Overview of the benchmark models and the optimization models that aim to reduce the mean and covariance estimation errors. All the portfolios incorporate the full-investment and non-negativity constraint.

4.3 Portfolio performance

The goal of this study is to assess the performance of each of the discussed portfolio optimization techniques in an empirical setting. From the several optimization techniques, I obtain optimal weight matrices with the dimensions $(T-M)\times N$. I calculate the portfolio returns by multiplying the weight matrix by the historical returns such that I obtain a $(T-M)\times 1$ vector of expected portfolio returns. In the individual-stock setting I obtain daily portfolio returns and in the multi-manager setting monthly portfolio returns. Both are annualized by multiplying the mean return by 250 or 12, respectively. I use the assumption that one year contains 250 tradings days and 12 months. The annualized return is used to calculate the excess and alpha returns of the portfolios with their corresponding volatility. To compare the models among each other, four performance indicators are studied.

First, the Sharpe ratio is a widely used performance measure in portfolio theory. It is defined as the mean of excess portfolio returns divided by its standard deviation. The excess return is calculated by taking the risk-free rate from the Kenneth French Data Library [5]. To compare to investment strategies, this study considers the difference between the Sharpe ratio of the optimized portfolio and the Sharpe ratio of APG's asset allocation, referred to as bmk weights. The difference is defined as

$$\hat{\Delta} = \hat{SR}_k - \hat{SR}_l \quad \text{with } \hat{SR}_i = \frac{\mu_i^*}{\sigma_i^*}, \tag{50}$$

with $\hat{SR}_k$ and $\hat{SR}_l$ the Sharpe ratios of investment strategy k and l, respectively. The Sharpe ratio of portfolio i is defined as the excess return μ_i^* over its standard deviation σ_i^*.

Ledoit and Wolf (2008) propose a method to improve the traditional Sharpe ratio calculation for high-dimensional data, by introducing a heteroskedasticity and autocorrelation consistent (HAC) kernel estimator. They start by assuming that the observed portfolio return series constitute a strictly stationary time series such that the bivariate return distribution does not change over time. This distribution has mean vector μ^{HAC} and covariance matrix Σ^{HAC} for portfolio k and l given by

$$\mu^{\text{HAC}} = \begin{pmatrix} \mu_k^* \\ \mu_l^* \end{pmatrix} \quad \text{and} \quad \Sigma^{\text{HAC}} = \begin{pmatrix} {\sigma^*}_k^2 & {\sigma^*}_{kl} \\ {\sigma^*}_{lk} & {\sigma^*}_l^2 \end{pmatrix}. \tag{51}$$

The sample estimators of the mean and variance are given by $\hat{\mu}^*{}_k$, $\hat{\mu}^*{}_l$, $\hat{\sigma}^*{}_k^2$, and $\hat{\sigma}^*{}_l^2$.

Ledoit and Wolf (2008) define $u = (\mu_k^*, \mu_l^*, \xi_k, \xi_l)$‘ and $\hat{u} = (\hat{\mu}^*{}_k, \hat{\mu}^*{}_l, \hat{\xi}_k, \hat{\xi}_l)$‘. The second moments of the excess returns of the assets equal $\xi_i = \mathrm{E}[{r^*}_{1i}^2]$ for asset i with $i = 1, \ldots, N$. I can write the difference in Sharpe ratio as $\Delta = f(u)$ and $\hat{\Delta} = f(\hat{u})$, where

$$f(a,b,c,d) = \frac{a}{\sqrt{c-a^2}} - \frac{b}{\sqrt{d-a^2}}. \tag{52}$$

They assume the data to be i.i.d. implying

$$\sqrt{T}\left(\hat{\Delta} - \Delta\right) \xrightarrow{\text{d}} N\left(0, \nabla' f(v) \Psi \nabla f(v)\right) \tag{53}$$

with Ψ an unknown symmetric positive semi-definite matrix and T the time period. The gradient function $\nabla' f(v)$ equals

$$\nabla' f(v) = \left(\frac{c}{(c-a^2)^{1.5}}, -\frac{d}{(d-b^2)^{1.5}}, -\frac{1}{2}\frac{a}{(c-a^2)^{1.5}}, \frac{1}{2}\frac{b}{(d-b^2)^{1.5}} \right). \tag{54}$$

[5] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/datalibrary.html

Now I can define the standard error of $\hat{\Delta}$, if a consistent estimator $\hat{\Psi}$ is available, given by

$$s\left(\hat{\Delta}\right) = \sqrt{\frac{\nabla' f(v)\hat{\Psi}\nabla f(v)}{T}}. \tag{55}$$

Andrews (1991) uses heteroskedasticity and autocorrelation robust (HAC) kernel estimation to obtain a consistent estimator $\hat{\Psi}$ such that

$$\hat{\Psi} = \hat{\Psi}_T = \frac{T}{T-4}\sum_{j=-T+1}^{T-1} k\left(\frac{j}{S_T}\right)\hat{\Gamma}_T(j), \tag{56}$$

where $k()$ equals the kernel function, S_T the bandwidth, and $\hat{\Gamma}_T(j)$ a limiting covariance matrix that is given by

$$\Gamma_T(j) = \begin{cases} \frac{1}{T}\sum_{t=j+1}^{T} \mathrm{E}\left[y_t y'_{t-j}\right] & \text{for } j \geq 0 \\ \frac{1}{T}\sum_{t=-j+1}^{T} \mathrm{E}\left[y_{t+j} y'_t\right] & \text{for } j < 0 \end{cases} \quad \text{with } y'_t = (r_{ti} - \mu_1^*, r_{tn} - \mu_n^*, r_{ti}^2 - \xi_i, r_{tn}^2 - \xi_n). \tag{57}$$

Though several kernels can be applied, this study utilizes the pre-whitened Parzen kernel as proposed by Andrews (1991). Several kernels automatically pick the bandwidth S_T. When the kernel and bandwidth are in place, a two-sided p-value for the null hypothesis $H_0 : \Delta = 0$ is given by

$$\hat{p} = 2\Phi\left(-\frac{|\hat{\Delta}|}{s(\hat{\Delta})}\right). \tag{58}$$

Next to the Sharpe ratio, this study also researches the Information ratio. This performance measure subtracts the benchmark portfolio returns from the portfolio returns. This gives us the alpha returns such that

$$\hat{\text{I}} = \frac{\alpha_i}{\sigma_{i,\alpha}} \quad \text{with } \alpha_i = \mu_i^* - \mu_b^*. \tag{59}$$

with μ_i^* and μ_b^* the excess returns of portfolio i and the benchmark portfolio. The volatility of the alpha returns is captured by $\sigma_{i,\alpha}$. In the individual-stock setting, the benchmark portfolio equals the allocation over the 1197 assets. In the multi-manager setting, the returns are alpha returns meaning APG already subtracts the geographical MSCI Index from the institutional-managed portfolio returns. In summary, the Sharpe and Information ratios are both calculated following the HAC inference estimator with a Parzen kernel.

In addition, this study uses certainty-equivalent (CEQ) returns to compare different investment strategies. The CEQ return equals the risk-free rate an investor is willing to accept rather than adopting a particular risky portfolio strategy. Specifically, for strategy i, the CEQ is defined as:

$$\widehat{\text{CEQ}}_i = \hat{\mu}_i^* - \frac{\gamma}{2}(\hat{\sigma}_i^*)^2, \tag{60}$$

where $\hat{\mu}_i^*$ and $\hat{\sigma^*}_i$ the sample mean and standard deviation of the portfolio excess returns for investment strategy i, and γ the level of risk-aversion. This study uses the implied level of risk-aversion γ to calculate the CEQ in the individual-stock setting. In the multi-manager setting, I use the risk-aversion levels 2, 7, and the implied level of risk-aversion $\gamma \simeq 22$. To test whether the CEQ is significantly different from the bmk weight portfolio I simulate 10.000 CEQ calculations. The p-value is calculated by taking the sum of differences that is larger than the sample difference of the respective CEQ. This approach is consistent with that of DeMiguel, Garlappi, and Uppal (2009).

A higher CEQ means an investor demands a higher return on a risk-free asset. Thus a strategy with a lower CEQ is considered inferior to a strategy with a higher CEQ. This makes the CEQ a performance indicator especially relevant for pension funds as their overall aim is to pay a certain return to their participants by retirement.

Fourthly, the turnover ratio is a relevant performance metric to assess the efficiency of portfolio construction strategies. The turnover ratio is defined as the sum of absolute differences between the portfolio weight under strategy j at time $t+1$ versus the portfolio weight at time t, i.e.,

$$\text{Turnover} = \frac{1}{T-M} \sum_{t=1}^{T-M} \sum_{j=1}^{N} \left(\left| \hat{w}j,t+1 - \hat{w}j,t^{+} \right| \right), \tag{61}$$

where N is the number of assets and t the moment in time $M+1,\ldots,T$. The turnover ratio can be interpreted as the percentage of wealth that is traded in each period. Thus the percentage of wealth that is traded per day in the individual-stock setting, and per month in the multi-manager setting. Altogether, this results in four performance measures: 1) the Sharpe ratio with HAC inference, 2) the Information ratio with HAC inference, 3) the CEQ, and 4) the turnover ratio

5 Results

This section presents the empirical results of portfolio optimization techniques applied in the individual-stock and multi-manager settings. Both settings are evaluated empirically over a time period from October 2012 to October 2022. The returns in the individual-stock setting are based on a daily frequency, while the returns in the multi-manager setting are based on a monthly frequency. The four performance indicators for the different portfolio optimizations are presented in tables 3, 4, 5, 6, and tables 7, 8 in appendices A and C, respectively. The extensions that aim to improve mean or covariance estimation are inserted in the benchmark model. For the individual-stock setting, the best hedge model (as in equation 1) and the minimum-variance function are used, while the mean-variance portfolio is used for the multi-manager setting. All the benchmark optimizations entail a sample mean and covariance estimation. The tables also show the bmk weights, which represent the weights that APG holds over the time period October 2012 to October 2022.

In addition, I present information ratios and CEQ in figures 2, 3, and figures 4 and 5 in appendices B and D, respectively. The information ratio is a widely used metric in the industry to evaluate a portfolio's additional return compared to the benchmark portfolio over its volatility. The CEQ is a frequently used metric for pension funds as they are interested in the risk-free return that simulates the portfolio's return. Therefore, these two performance indicators are not only given in tables but also visualized in figures.

5.1 Individual-stock setting

Tables 3 and 7 give the results of the best hedge models for the hedging coefficients $\nu = 0.8$ and $\nu = 0.4$, respectively. Moreover, table 4 shows the results of the minimum-variance model. The p-values of the Sharpe ratio, Info ratio, and CEQ test whether the alternative portfolios obtain a value significantly different from APG's benchmark portfolio, the bmk weights. Owing to the overall small differences between the various optimization techniques, only a few p-values are significant at a 10% level.

First, I compare the best hedge models to the minimum-variance models. Tables 3, 4, 7 show that the best hedge models outperform the minimum-variance models for both hedging coefficients. The best hedge optimizations obtain higher Sharpe and Information ratios, CEQ, and lower turnover ratios. This indicates that accounting for minimizing the tracking error relative to the benchmark weights, while accounting for factor risk leads to better investments than simply minimizing the volatility of the portfolio. I see a similar result when I compare the best hedge and minimum-variance models to the benchmark weights, referred to as bmk weights. The best hedge optimizations perform similarly or outperform the benchmark weights, while the minimum-variance portfolios are outperformed by the benchmark weights. This indicates APG is better off by investing in the benchmark portfolio or optimizing according to the best hedge model than following the minimum-variance portfolio.

Table 4 shows that not only the minimum-variance portfolio with the sample estimates is outperformed by the benchmark weight portfolio, but also the extensions that entail factor, linear shrinkage, non-linear shrinkage, and dynamic factor models are outperformed. Only the dynamic models outperform the benchmark weights. When I compare the benchmark weights to the best hedge models, I see that dynamic-factor models obtain lower Sharpe ratios than the benchmark weights. On the contrary, the linear and non-linear shrinkage estimators obtain higher Sharpe ratios than the benchmark weights. These results indicate that the benchmark weights outperform most of the optimizations in the minimum-variance model, while the opposite is the case for the best hedge models. This further underlines the result that the best hedges model is beneficial compared to the minimum-variance optimization.

Table 3: Annualized empirical excess and alpha returns of the best hedge model (equation 1) with $\nu = 0.8$ from Oct 2012 - Oct 2022 such that $T = 2509$ days. The number of assets equals 1197 for the benchmark portfolio and 564 for the portfolio. The estimation windows differ between $M = 250$ and $M = 750$ days. The returns are the one-period out-of-sample returns, and the optimal weights are updated once every 21 trading days. For the excess returns I subtract the risk-free rate, and for alpha returns the bmk weight returns. All the models are different estimations of the covariance matrix in the best hedge model, where 'best hedges' uses the sample covariance matrix. If CEQ equals 0.01, it indicates an investor values the portfolio at a 1% risk-free return. A turnover of 0.01 indicates 1% of the wealth is traded on a daily basis.

	Benchmark		*Linear shrinkage*					*Non-linear shrinkage*			*DCC*					
	bmk	best hedges	LS cov1Para	LS cov2Para	LS covCor	LS covDiag	LS covMkt	NLS LIS	NLS QIS	NLS GIS	DCC	DCC-LS	DCC-NLS	DCC-3FF	DCC-3FF-LS	DCC-3FF-NLS
								M = 250								
Excess return	0.1446	0.1523	0.1545	0.1548	0.1525	0.1537	0.1542		0.1546		0.1601	0.1601	0.1602	0.1535	0.1535	0.1535
Std dev Excess return	0.1532	0.1601	0.1610	0.1618	0.1611	0.1604	0.1611		0.1628		0.1704	0.1704	0.1704	0.1678	0.1678	0.1678
Sharpe ratio	0.9441	0.9511	0.9594	0.9564	0.9465	0.9580	0.9569		0.9496		0.9397	0.9397	0.9401	0.9147	0.9147	0.9147
p-value Sharpe ratio		0.8790	0.7281	0.7773	0.9567	0.7542	0.7737		0.9002		0.9202	0.9213	0.9286	0.5724	0.5724	0.5724
Alpha return		0.0077	0.0099	0.0101	0.0079	0.0091	0.0095		0.0100		0.0155	0.0155	0.0156	0.0089	0.0089	0.0089
Std dev Alpha return		0.0229	0.0227	0.0230	0.0229	0.0226	0.0228		0.0235		0.0282	0.0282	0.0282	0.0287	0.0287	0.0287
Info ratio		0.3349	0.4348	0.4407	0.3452	0.4026	0.4177		0.4231		0.5494	0.5499	0.5532	0.3083	0.3083	0.3083
p-value Info ratio		0.1798	0.2451	0.2391	0.1729	0.2262	0.2297		0.2131		0.2497	0.2502	0.2537	0.1012	0.1012	0.1012
CEQ		0.0664	0.0676	0.0671	0.0655	0.0675	0.0672		0.0658		0.0628	0.0629	0.0629	0.0592	0.0592	0.0592
p-value CEQ		0.9753	0.8978	0.9339	0.9703	0.9060	0.9240		0.9877		0.8117	0.8121	0.8151	0.6368	0.6368	0.6368
Turnover		0.0356	0.0191	0.0182	0.0190	0.0204	0.0188		0.0164		0.0148	0.0148	0.0148	0.0206	0.0206	0.0206
								M = 750								
Excess return	0.1438	0.1556	0.1565	0.1570	0.1558	0.1562	0.1566	0.1561	0.1568	0.1572	0.1639	0.1640	0.1640	0.1353	0.1600	0.1600
Std dev Excess return	0.1617	0.1684	0.1689	0.1695	0.1690	0.1685	0.1689	0.1673	0.1683	0.1709	0.1808	0.1808	0.1808	0.1634	0.1832	0.1832
Sharpe ratio	0.8895	0.9243	0.9262	0.9262	0.9220	0.9267	0.9275	0.9333	0.9314	0.9202	0.9068	0.9068	0.9071	0.8284	0.8733	0.8733
p-value Sharpe ratio		0.4557	0.4257	0.4220	0.4823	0.4234	0.4133	0.3468	0.3648	0.5021	0.7032	0.7028	0.7002	0.0232	0.7984	0.7984
Alpha return		0.0118	0.0127	0.0132	0.0120	0.0124	0.0128	0.0123	0.0130	0.0134	0.0201	0.0202	0.0202	-0.0085	0.0162	0.0162
Std dev Alpha return		0.0222	0.0222	0.0224	0.0222	0.0221	0.0222	0.0215	0.0218	0.0230	0.0293	0.0293	0.0293	0.0122	0.0355	0.0355
Info ratio		0.5336	0.5714	0.5917	0.5419	0.5591	0.5774	0.5740	0.5961	0.5849	0.6876	0.6878	0.6892	-0.6982	0.4562	0.4562
p-value Info ratio		0.4795	0.5203	0.5392	0.4836	0.5105	0.5298	0.5418	0.5594	0.5169	0.5746	0.5749	0.5773	0.0018	0.2687	0.2687
CEQ		0.0607	0.0609	0.0608	0.0601	0.0610	0.0611	0.0624	0.0619	0.0594	0.0545	0.0545	0.0545	0.0459	0.0476	0.0476
p-value CEQ		0.7525	0.7429	0.7505	0.7796	0.7364	0.7354	0.6782	0.7003	0.8192	0.8959	0.8960	0.8970	0.5563	0.6074	0.6074
Turnover		0.0194	0.0177	0.0173	0.0174	0.0181	0.0175	0.0172	0.0168	0.0161	0.0152	0.0152	0.0152	0.0157	0.0187	0.0187

Next, I evaluate the impact of lowering the hedging coefficient ν from 0.8 to 0.4. Overall, the results in table 3 and 7 do not show large differences, indicating changing the hedging coefficient only slightly impacts the performance of the best hedge models. Considering the Information ratio, I see that the portfolios with the lower hedging coefficient outperform the portfolios with the higher hedging coefficient. This result is intuitive, as a milder hedging restriction gives space to invest more risky in order to obtain higher returns. Despite the dynamic-factor models, I obtain higher volatilities for the lower hedging coefficient $\nu = 0.4$. This underlines the idea that an investor with a lower hedging coefficient starts to invest more risky. I further see that the CEQ is positively dependent on the hedging coefficient, thus CEQs for $\nu = 0.8$ are higher than for $\nu = 0.4$. Finally, I observe slightly higher turnover ratios for $\nu = 0.8$ than for $\nu = 0.4$. Both results are in line with our hypothesis, as hedging more heavily for factor risk is likely to lead to a higher risk-free return, and a higher trading volume to preserve the hedge.

The best hedge models with dynamic-factor covariance estimators account twice for factor exposure to the 3 Fama-French factors making it a special case. Table 3 shows that the best hedge model with sample estimates strictly outperforms the best hedge model with the dynamic-factor covariance estimates. Therefore, I find that accounting for factor exposure twice does not improve portfolio performance. Moreover, the results in table 7 indicate the dynamic-factor models for lower hedging coefficients perform better than for higher hedging coefficients relative to alternative optimization techniques. This result is intuitive as a lower hedging coefficient means the optimizations assign less weight to hedge the factor exposure. Therefore, the models entail a lower factor exposure in the best hedge function, resulting in relatively better performance of the dynamic-factor estimators. To summarize, accounting for factor exposure twice in the dynamic-factor best hedge models is not beneficial, though the negative impact erodes lower hedging coefficients.

Furthermore, tables 3 and 7 show that the results of the dynamic-factor models are almost identical. Only the DCC-3FF deviates from the DCC-3FF-LS and DCC-3FF-NLS models for $\nu = 0.8$ and $M = 750$. I presume this is caused by how I define the dynamic-factor models in this study. I allow the factor covariance matrix to be time-variant but assume the factor-exposure and residual covariance matrix to be time-invariant. Not allowing for time variance in the residual covariance matrix reduces the dynamic element of the covariance estimation, and likely results in identical output over different estimation windows. Moreover, Ledoit and Wolf (2022b) stated that allowing for time-variance in the residual covariance matrix improves the dynamic-factor estimation.

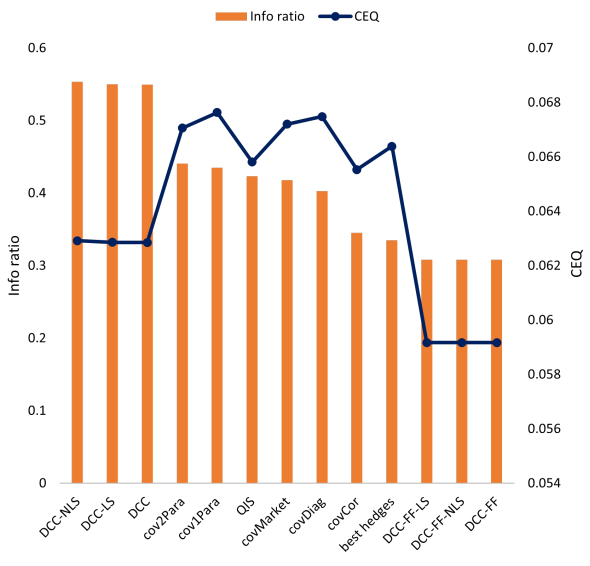
Comparing the alternative optimization techniques in the best hedge model, I see in table 3 that linearly or non-linearly shrinking the covariance matrix improves the sample covariance estimator. This is in line with our hypothesis, as previous papers by Ledoit and Wolf (2004b, 2022b) state shrinking the covariance matrix in high-dimensional settings lead to better-performing portfolios. Table 7 shows that the linear and non-linear shrinkage estimators even further improve the best hedge estimator for $\nu = 0.4$. This result is intuitive as reducing the weight to the factor exposure reduces the 'shrinking' in the best hedge function. As a result, shrinking the covariance matrix results in larger gains for the best hedge models. Therefore, I conclude that shrinking the covariance matrix with a large weight to the factor exposure results in better-performing portfolios. The turnover ratio stresses this conclusion as the best hedge model with the sample covariance matrix attains higher trading volumes than the linear and non-linear shrinkage estimators.

Figure 2 shows the Information ratios of the dynamic covariance estimators outperform the alternative optimizations. Also, they obtain the lowest turnover ratios. I see in table 7 and figure 4 that changing the hedging coefficient does not affect this result. The results indicate different optimal portfolios for the Sharpe and Information ratios. An investor that intends to attain the highest Sharpe ratio benefits from shrinking the covariance matrix, while one that intends to attain the highest Information ratio benefits from accounting for time-variance in the covariance estimation.

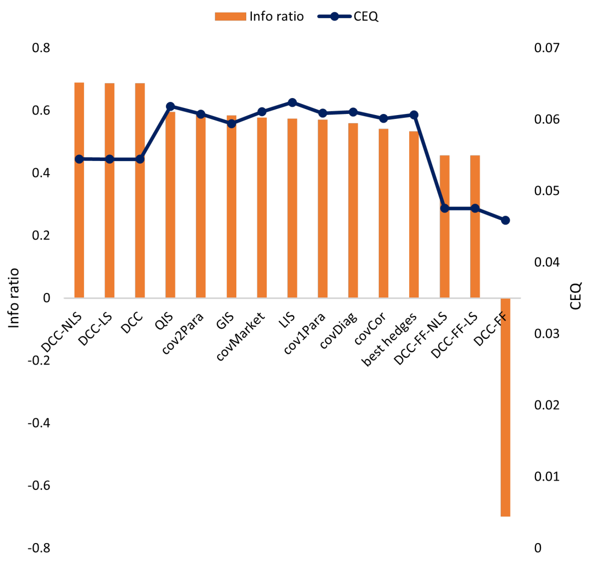
Despite the turnover ratios of the best hedge model with the sample covariance matrix, the results show that the estimation window does not largely impact the trading volume as the turnover ratios remain roughly consistent between $M = 250$ and $M = 750$. I obtain larger differences in trading volume between the different optimizations. The dynamic covariance estimators obtain the lowest turnover ratios. This result is intuitive as updating the covariance estimation on a daily basis results in more naturally updating the covariance estimator on market conditions (Engle 2002). Therefore, holding a portfolio with a time-variant covariance estimator results in lower trading volumes.

In order to assess whether the same results for the best hedge models are observed in a more simple setting, I test similar extensions in a minimum-variance model. I already found that switching to a minimum-variance model erodes the portfolio performance. Furthermore, I see in Table 4 that the dynamic-covariance estimators outperform the alternative optimizations for both the Sharpe and Information ratio, similar to the result I obtained for Information ratios of the best hedge models. Moreover, the dynamic-factor covariance estimators obtain higher Information ratios than the linear and non-linear shrinkage estimators and the minimum-variance portfolio. Therefore, I conclude that accounting for time-variance in the covariance estimator is beneficial in the individual-stock setting ($N = 564$, $T = 2509$), though incorporating factors in the dynamic-covariance estimator is not.

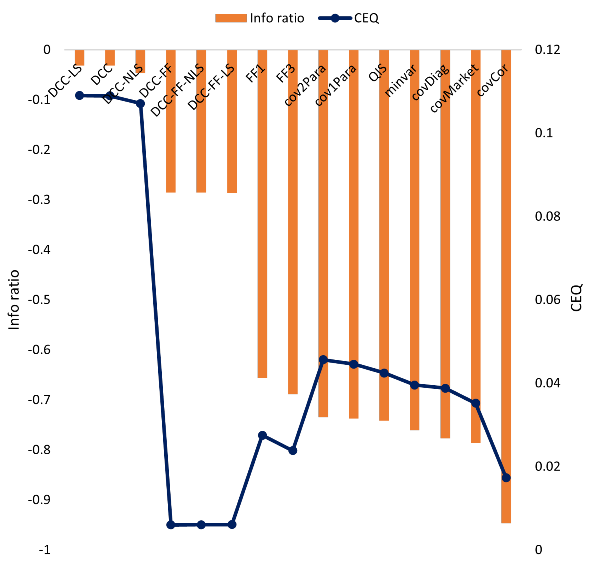
Notably, table 4 shows estimating the covariance matrix with factor models or linearly shrinkage does not strictly outperform the sample estimator, except for the LS cov1Para and LScov2Para estimators. The cov1Para uses a matrix with constant variances on the diagonal and cov2Para constant variances on the diagonal and constant correlations on the off-diagonal. Moreover, I find that non-linearly shrinking the covariance results in a portfolio that outperforms the sample covariance estimator. This result underlines the conclusion of Ledoit and Wolf (2012) that non-linearly shrinking the covariance matrix leads to better-performing portfolios than linearly shrinking the covariance matrix in high-dimensional settings.



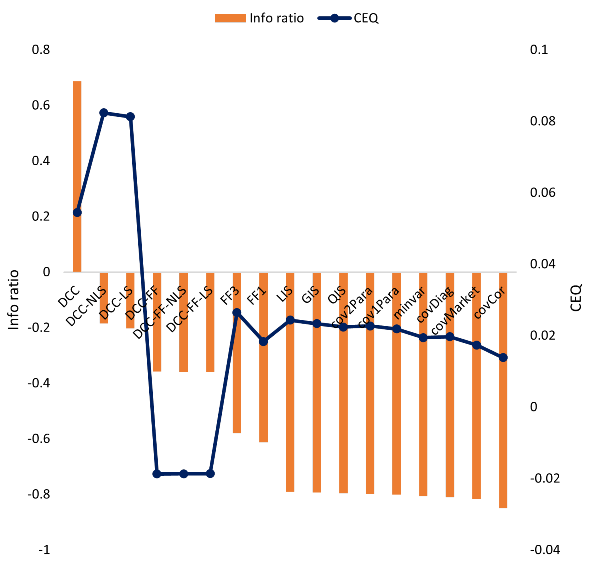
(a) Best hedges models with $M = 250$



(b) Best hedges models with $M = 750$



(c) Minimum-variance models with $M = 250$



(d) Minimum-variance models with $M = 750$

Figure 2: Empirically obtained Information ratios and CEQs from tables 3 and 4 of the one-period out-of-sample daily returns over the time period October 2012 - October 2022. The portfolios are sorted on the x-axis on descending order of the Information ratio. The Information ratios are on the left y-axis, and the CEQ on the right y-axis.

Table 4: Annualized empirical excess and alpha returns of the minimum-variance model (equation 17) from Oct 2012 - Oct 2022 such that $T = 2509$ days. The number of assets equals 564. The estimation windows differ between $M = 250$ and $M = 750$ days. The returns are the one-period out-of-sample returns, and the optimal weights are updated once every 21 trading days. For the excess returns I subtract the risk-free rate, and for alpha returns the bmk weight returns. All the models are different estimations of the covariance matrix in the minimum-variance model, where 'min-var' uses the sample covariance matrix. If CEQ equals 0.01, it indicates an investor values the portfolio at a 1% risk-free return. A turnover of 0.01 indicates 1% of the wealth is traded on a daily basis.

	Benchmark		*Factor model*		*Linear shrinkage*					*Non-linear shrinkage*			*DCC*					
	bmk	minvar	1FF	3FF	LS cov1Para	LS cov2Para	LS covCor	LS covDiag	LS covMkt	NLS LIS	NLS QIS	NLS GIS	DCC	DCC-LS	DCC-NLS	DCC-3FF	DCC-3FF-LS	DCC-3FF-NLS
									M = 250									
Excess return	0.1446	0.0668	0.0650	0.0603	0.0713	0.0723	0.0460	0.0658	0.0617		0.0690		0.1476	0.1477	0.1460	0.0913	0.0912	0.0912
Std dev Excess return	0.1532	0.0901	0.1058	0.1044	0.0893	0.0893	0.0926	0.0897	0.0889		0.0890		0.1074	0.1074	0.1077	0.1595	0.1593	0.1594
Sharpe ratio	0.9441	0.7412	0.6146	0.5781	0.7986	0.8103	0.4973	0.7333	0.6940		0.7752		1.3742	1.3751	1.3558	0.5721	0.5722	0.5722
p-value Sharpe ratio		0.3296	0.1098	0.0740	0.4361	0.4589	0.0441	0.3062	0.2525		0.3956		0.1108	0.1096	0.1421	0.3279	0.3281	0.3280
Alpha return		-0.0842	-0.0859	-0.0906	-0.0797	-0.0786	-0.1049	-0.0852	-0.0893		-0.0820		-0.0034	-0.0033	-0.0050	-0.0597	-0.0598	-0.0598
Std dev Alpha return		0.1106	0.1311	0.1317	0.1081	0.1071	0.1109	0.1097	0.1136		0.1106		0.1069	0.1068	0.1089	0.2093	0.2091	0.2092
Info ratio		-0.7608	-0.6555	-0.6886	-0.7370	-0.7341	-0.9466	-0.7765	-0.7861		-0.7416		-0.0313	-0.0311	-0.0460	-0.2854	-0.2860	-0.2856
p-value Info ratio		0.0049	0.0027	0.0021	0.0053	0.0052	0.0019	0.0045	0.0043		0.0051		0.0826	0.0828	0.0803	0.0341	0.0340	0.0341
CEQ		0.0396	0.0275	0.0238	0.0446	0.0456	0.0173	0.0388	0.0352		0.0425		0.1090	0.1090	0.1071	0.0060	0.0061	0.0061
p-value CEQ		0.5055	0.5024	0.5023	0.5055	0.5056	0.5056	0.5055	0.5057		0.5057		0.4978	0.4977	0.4981	0.4985	0.4988	0.4986
Turnover		0.0349	0.0250	0.0257	0.0325	0.0317	0.0306	0.0333	0.0309		0.0294		0.0454	0.0456	0.0468	0.0239	0.0239	0.0239
									M = 750									
Excess return	0.1438	0.0462	0.0569	0.0682	0.0482	0.0490	0.0407	0.0463	0.0436	0.0505	0.0483	0.0494	0.1639	0.1215	0.1231	0.0671	0.0669	0.0670
Std dev Excess return	0.1617	0.0893	0.1072	0.1117	0.0887	0.0886	0.0896	0.0891	0.0885	0.0884	0.0879	0.0882	0.1808	0.1095	0.1102	0.1600	0.1598	0.1599
Sharpe ratio	0.8895	0.5172	0.5303	0.6111	0.5434	0.5529	0.4546	0.5196	0.4925	0.5716	0.5496	0.5600	0.9068	1.1090	1.1164	0.4192	0.4188	0.4190
p-value Sharpe ratio		0.2001	0.1747	0.2207	0.2167	0.2237	0.1432	0.1974	0.1752	0.2380	0.2210	0.2285	0.7032	0.4110	0.4119	0.3125	0.3124	0.3124
Alpha return		-0.0976	-0.0869	-0.0756	-0.0956	-0.0948	-0.1031	-0.0975	-0.1002	-0.0932	-0.0955	-0.0944	0.0201	-0.0223	-0.0207	-0.0767	-0.0769	-0.0768
Std dev Alpha return		0.1210	0.1419	0.1303	0.1194	0.1187	0.1214	0.1204	0.1227	0.1179	0.1201	0.1190	0.0293	0.1105	0.1123	0.2142	0.2141	0.2141
Info ratio		-0.8062	-0.6127	-0.5796	-0.8008	-0.7985	-0.8489	-0.8099	-0.8165	-0.7906	-0.7950	-0.7933	0.6876	-0.2021	-0.1847	-0.3583	-0.3591	-0.3586
p-value Info ratio		0.0163	0.0155	0.0145	0.0163	0.0163	0.0145	0.0160	0.0156	0.0162	0.0161	0.0161	0.5746	0.0965	0.1038	0.0626	0.0626	0.0626
CEQ		0.0195	0.0184	0.0265	0.0218	0.0227	0.0138	0.0197	0.0173	0.0243	0.0224	0.0233	0.0545	0.0813	0.0824	-0.0187	-0.0186	-0.0187
p-value CEQ		0.5068	0.5023	0.5014	0.5067	0.5067	0.5067	0.5066	0.5067	0.5068	0.5071	0.5072	0.8959	0.4975	0.4973	0.4988	0.4985	0.4987
Turnover		0.0215	0.0190	0.0259	0.0212	0.0210	0.0206	0.0213	0.0209	0.0205	0.0203	0.0204	0.0152	0.0455	0.0467	0.0240	0.0239	0.0239

Figures 5c and 2d show that the dynamic-factor covariance estimators obtain higher Information ratios than the factor, linear, and non-linear shrinkage estimators. Nevertheless, the dynamic-factor models also obtain the lowest CEQ compared to the alternative optimizations. This indicates an investor demands the lowest risk-free return for holding the specific portfolio. The CEQ values for the dynamic-factor models even obtain negative values. When the CEQ equals -0.0187 I state that an investor expects a negative risk-free return of -1.87% on the portfolio. As a result, it is unlikely that an agent would invest in such a portfolio.

Altogether, I find that the minimum-variance model with sample estimates is not easily outperformed by alternative optimization techniques, despite the dynamic-covariance estimators. Moreover, the minimum-variance model does not strictly outperform the benchmark weights resulting in negative alpha returns and Information ratios. Although the factor, linear, and non-linear shrinkage models obtain lower turnover ratios than the sample minimum-variance model, the dynamic and dynamic-factor covariance estimators do not. This is in contrast with our hypothesis as accounting for time-variance generally leads to lower turnover ratios (Engle 2002).

5.2 Multi-manager setting

Tables 5, 6, and 8 show the portfolios' performance for $\gamma = 2, 7$, and the implied level of risk-aversion $\gamma \simeq 22$. The benchmark (bmk) weight portfolio equals APG's benchmark weights over time, and the benchmark model equals Markowitz (1952) mean-variance optimization. For all risk aversion levels, I find that the mean-variance portfolio only outperforms the benchmark portfolio for a short estimation window of $M = 12$. Therefore, I find that the performance of the mean-variance portfolio worsens for larger estimation windows. The decreasing performance of the mean-variance portfolio is likely caused by the presence of estimation errors in the mean and covariance.

Further evaluating the effect of the estimation errors, I compare the mean-variance model to the minimum-variance model both with sample estimates. The mean-variance model deals with mean and covariance estimation errors, while the minimum-variance model only deals with covariance estimation errors. Tables 5, 6, and8 show that for $M = 12$ the mean-variance portfolio outperforms the minimum-variance portfolio. The opposite is observed for $M = 36$ and $M = 72$. This implies that the mean estimation errors enlarge when more observations are considered in the estimation window. This result underlines the conclusion of Merton (1980) which states the covariance estimation tends to improve for more observations, while the mean estimation does not. Moreover, the result shows that estimation errors in the mean weaken the portfolio performance.

Table 5: Annualized empirical excess and alpha returns of the mean-variance model (equation 10) with $\gamma = 2$ from Oct 2012 - Oct 2022 such that $T = 121$ months. The number of assets $N = 8$. The estimation windows differ between $M = 12, 36, 72$ months. The returns are the one-period, i.e., one-month, out-of-sample alpha returns. APG subtracts the respective geographical MSCI Index to obtain the alpha returns. All the models are different estimations of the mean or covariance in the mean-variance model. If CEQ equals 0.01, it indicates an investor values the portfolio at a 1% risk-free return. A turnover of 0.01 indicates 1% of the wealth is traded on a daily basis.

	Benchmark		*Mean shrinkage*			*Factor models*		*Linear shrinkage*					*Non-linear shrinkage*			*DCC*					
	bmk	mv	minvar	Pastor	James-Stein	1FF	3FF	LS cov1Para	LS cov2Para	LS covCor	LS covDiag	LS covMkt	NLS LIS	NLS QIS	NLS GIS	DCC	DCC-LS	DCC-NLS	DCC-3FF	DCC-3FF-LS	DCC-3FF-NLS
M = 12																					
Alpha return	0.0192	0.0454	0.0117	0.0120	0.0431	0.0453	0.0453	0.0458	0.0457	0.0454	0.0454	0.0455	0.0469	0.0455	0.0459	0.0123	0.0127	0.0123	0.0243	0.0249	0.0249
Std dev Alpha return	0.0302	0.0454	0.0276	0.0276	0.0455	0.0454	0.0454	0.0455	0.0454	0.0454	0.0454	0.0454	0.0456	0.0454	0.0455	0.0240	0.0240	0.0241	0.0293	0.0294	0.0294
Info ratio	0.6382	1.0003	0.4241	0.4350	0.9473	0.9975	0.9976	1.0073	1.0065	0.9986	0.9988	1.0005	1.0277	1.0006	1.0086	0.5117	0.5291	0.5103	0.8274	0.8453	0.8453
p-value Info ratio		0.4673	0.4598	0.4822	0.5304	0.4709	0.4709	0.4583	0.4592	0.4695	0.4692	0.4670	0.4334	0.4672	0.4568	0.7246	0.7382	0.7061	0.5501	0.5270	0.5270
CEQ		0.0434	0.0109	0.0112	0.0411	0.0432	0.0432	0.0437	0.0437	0.0433	0.0433	0.0434	0.0448	0.0434	0.0438	0.0117	0.0121	0.0117	0.0234	0.0240	0.0240
p-value CEQ		0.5020	0.4989	0.4988	0.5021	0.5019	0.5019	0.5020	0.5020	0.5019	0.5019	0.5020	0.5020	0.5020	0.5019	0.5005	0.5005	0.5004	0.5011	0.5012	0.5012
Turnover		0.4918	0.3469	0.3482	0.5428	0.4999	0.4997	0.5004	0.4940	0.4953	0.4953	0.4917	0.4907	0.4911	0.4897	0.0124	0.0135	0.0124	0.3111	0.3144	0.3144
M = 36																					
Alpha return	0.0202	0.0006	0.0111	0.0116	0.0079	-0.0004	0.0001	-0.0006	-0.0008	0.0002	0.0004	0.0005	0.0001	0.0004	0.0004	0.0129	0.0131	0.0130	0.0158	0.0161	0.0161
Std dev Alpha return	0.0473	0.0473	0.0253	0.0249	0.0449	0.0476	0.0472	0.0478	0.0479	0.0474	0.0473	0.0473	0.0474	0.0472	0.0472	0.0253	0.0252	0.0253	0.0264	0.0263	0.0263
Info ratio	0.4276	0.0137	0.4384	0.4648	0.1755	-0.0079	0.0029	-0.0129	-0.0160	0.0053	0.0081	0.0105	0.0031	0.0084	0.0087	0.5119	0.5190	0.5145	0.5987	0.6100	0.6100
p-value Info ratio		0.0870	0.4716	0.5137	0.1613	0.0788	0.0824	0.0793	0.0785	0.0840	0.0849	0.0857	0.0836	0.0847	0.0851	0.6660	0.6459	0.6523	0.7755	0.8012	0.8012
CEQ		-0.0016	0.0104	0.0109	0.0059	-0.0026	-0.0021	-0.0029	-0.0031	-0.0020	-0.0019	-0.0017	-0.0021	-0.0018	-0.0018	0.0123	0.0125	0.0124	0.0151	0.0154	0.0154
p-value CEQ		0.4975	0.5002	0.5007	0.4978	0.4969	0.4974	0.4967	0.4965	0.4972	0.4973	0.4972	0.4972	0.4974	0.4974	0.5001	0.5002	0.5002	0.4987	0.4985	0.4985
Turnover		0.4476	0.1199	0.1195	0.4410	0.4562	0.4494	0.4453	0.4463	0.4493	0.4465	0.4498	0.4479	0.4449	0.4452	0.0142	0.0154	0.0142	0.1494	0.1523	0.1523
M = 72																					
Alpha return	0.0216	0.0061	0.0065	0.0063	0.0115	0.0055	0.0056	0.0051	0.0052	0.0058	0.0057	0.0060	0.0055	0.0056	0.0055	0.0154	0.0163	0.0156	0.0152	0.0153	0.0153
Std dev Alpha return	0.0245	0.0462	0.0291	0.0290	0.0288	0.0455	0.0456	0.0461	0.0461	0.0461	0.0461	0.0462	0.0461	0.0461	0.0461	0.0290	0.0290	0.0291	0.0299	0.0298	0.0298
Info ratio	0.8814	0.1316	0.2222	0.2178	0.3980	0.1209	0.1223	0.1118	0.1124	0.1259	0.1246	0.1307	0.1199	0.1203	0.1203	0.5332	0.5623	0.5364	0.5082	0.5132	0.5133
p-value Info ratio		0.0677	0.3213	0.3056	0.2736	0.0580	0.0583	0.0558	0.0562	0.0631	0.0625	0.0670	0.0599	0.0598	0.0599	0.7958	0.8148	0.7874	0.5535	0.5737	0.5738
CEQ		0.0039	0.0056	0.0055	0.0106	0.0034	0.0035	0.0030	0.0031	0.0037	0.0036	0.0039	0.0034	0.0034	0.0034	0.0146	0.0154	0.0148	0.0143	0.0144	0.0144
p-value CEQ		0.4981	0.4987	0.4986	0.4987	0.4984	0.4982	0.4978	0.4979	0.4978	0.4978	0.4980	0.4979	0.4979	0.4979	0.4986	0.4986	0.4986	0.4993	0.4994	0.4994
Turnover		0.3905	0.0681	0.0680	0.2408	0.3797	0.3790	0.3937	0.3942	0.3879	0.3884	0.3900	0.3899	0.3900	0.3899	0.0147	0.0155	0.0146	0.1403	0.1367	0.1367

Due to the presence of estimation errors in the mean and covariance in the mean-variance portfolio, I expect this optimization to be outperformed by all the alternative optimizations. Nevertheless, figures 3 and 5 show that the mean-variance portfolio performs average in terms of Information ratio and CEQ. Moreover, I see that the performance of the mean-variance portfolio further excels compared to the alternative optimizations for smaller estimation windows. This result shows that: 1) indeed estimation errors erode the performance of the mean-variance portfolio for the largest estimation windows as concluded above, and 2) optimizations that intend to improve these estimation errors do not seem to successfully do so in this empirical study. The latter is in contrast with the results of the Ledoit and Wolf (2004b, 2012), Pástor (2000), Sharpe (1963), and Stein (1956) as they introduce these estimators to improve the sample estimator.

I see that the dynamic and dynamic-factor covariance estimators outperform the alternative strategies for the estimation windows $M = 36$ and $M = 72$. Tables 6 and 8 show that for higher risk-aversion levels, the dynamic-factor covariance estimators obtain higher Information ratios and CEQs than the dynamic-covariance estimators. This means that it is beneficial for more risk-averse investors to incorporate factor exposure in the dynamic-covariance estimation, which is not surprising as previous results in the individual-stock setting indicate that accounting for factor exposure functions as a form of hedging. Moreover, I deduce that in general, it is beneficial to account for time-variance by estimating the covariance matrix. Although this result underlines the conclusion by Engle (2002), they state it is likely to be beneficial for daily or weekly return data. Despite the lower frequency of the monthly return data in the multi-manager setting, I still find the out-performance of the dynamic covariance estimators.

Furthermore, I see in figure 3 and 5 that the non-linear shrinkage estimators result in slight performance gains for the shorter estimation window $M = 12$ by obtaining the highest Information ratios and CEQs. Nonetheless, the difference between linear shrinkage, mean shrinkage and factor models is small. The overall performance of linear shrinkage estimators slightly improves compared to the non-linear shrinkage estimators. Ledoit and Wolf (2012) state the non-linear shrinkage is especially beneficial in high-dimensional settings, thus N close to M. In the multi-manager setting, I deal with $N = 8$ assets and an estimation window of $M = 12$. This means the number of assets is slightly larger than the moments in time in the estimation window. This might cause the non-linear shrinkage estimators to perform similarly or better than the linear shrinkage estimators.

Table 6: Annualized empirical excess and alpha returns of the mean-variance model (equation 10) with $\gamma = 7$ from Oct 2012 - Oct 2022 such that $T = 121$ months. The number of assets $N = 8$. The estimation windows differ between $M = 12, 36, 72$ months. The returns are the one-period, i.e., one-month, out-of-sample alpha returns. APG subtracts the respective geographical MSCI Index to obtain the alpha returns. All the models are different estimations of the mean or covariance in the mean-variance model. If CEQ equals 0.01, it indicates an investor values the portfolio at a 1% risk-free return. A turnover of 0.01 indicates 1% of the wealth is traded on a daily basis.

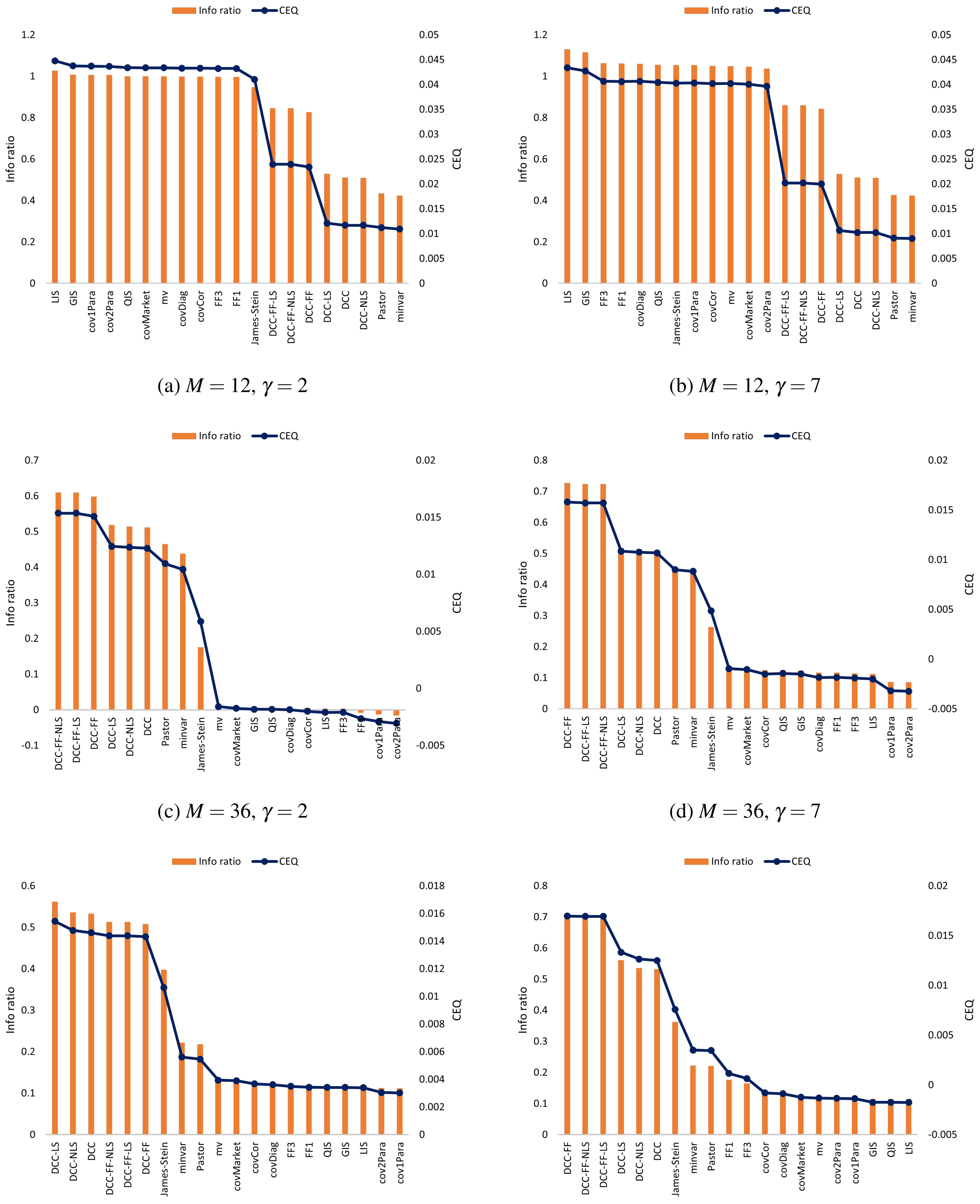
	Benchmark		*Mean shrinkage*			*Factor models*		*Linear shrinkage*					*Non-linear shrinkage*			*DCC*					
	bmk	mv	minvar	Pastor	James-Stein	1FF	3FF	LS cov1Para	LS cov2Para	LS covCor	LS covDiag	LS covMkt	NLS LIS	NLS QIS	NLS GIS	DCC	DCC-LS	DCC-NLS	DCC-3FF	DCC-3FF-LS	DCC-3FF-NLS
M = 12																					
Alpha return	0.0192	0.0474	0.0117	0.0118	0.0473	0.0477	0.0477	0.0474	0.0468	0.0473	0.0477	0.0472	0.0503	0.0476	0.0497	0.0123	0.0127	0.0123	0.0225	0.0226	0.0226
Std dev Alpha return	0.0302	0.0452	0.0276	0.0276	0.0449	0.0449	0.0449	0.0450	0.0451	0.0450	0.0450	0.0451	0.0445	0.0451	0.0446	0.0240	0.0240	0.0241	0.0267	0.0263	0.0263
Info ratio	0.6382	1.0490	0.4241	0.4271	1.0547	1.0622	1.0633	1.0540	1.0376	1.0506	1.0614	1.0465	1.1308	1.0554	1.1160	0.5110	0.5284	0.5095	0.8435	0.8609	0.8609
p-value Info ratio		0.3967	0.4598	0.4659	0.3778	0.3836	0.3826	0.3967	0.4161	0.3956	0.3832	0.4001	0.3147	0.3939	0.3284	0.7231	0.7367	0.7046	0.4871	0.4548	0.4547
CEQ		0.0402	0.0090	0.0091	0.0403	0.0406	0.0407	0.0403	0.0397	0.0402	0.0407	0.0401	0.0434	0.0405	0.0428	0.0102	0.0107	0.0102	0.0200	0.0202	0.0202
p-value CEQ		0.4997	0.4986	0.4986	0.4997	0.4997	0.4996	0.4996	0.4995	0.4997	0.4997	0.4995	0.4996	0.4995	0.4997	0.4991	0.4991	0.4988	0.5010	0.5016	0.5016
Turnover		0.4970	0.3469	0.3472	0.5330	0.5029	0.5054	0.5024	0.5045	0.5105	0.5074	0.4948	0.4690	0.4869	0.4814	0.0124	0.0135	0.0124	0.2489	0.2403	0.2403
M = 36																					
Alpha return	0.0202	0.0061	0.0111	0.0112	0.0113	0.0052	0.0051	0.0039	0.0039	0.0056	0.0053	0.0060	0.0050	0.0055	0.0055	0.0129	0.0131	0.0130	0.0179	0.0179	0.0179
Std dev Alpha return	0.0473	0.0449	0.0253	0.0252	0.0427	0.0450	0.0448	0.0449	0.0450	0.0450	0.0450	0.0449	0.0448	0.0446	0.0446	0.0253	0.0252	0.0253	0.0247	0.0247	0.0247
Info ratio	0.4276	0.1361	0.4384	0.4460	0.2634	0.1167	0.1145	0.0866	0.0858	0.1243	0.1167	0.1342	0.1124	0.1240	0.1229	0.5115	0.5185	0.5140	0.7274	0.7237	0.7237
p-value Info ratio		0.1343	0.4716	0.4835	0.2228	0.1146	0.1174	0.1031	0.1041	0.1236	0.1184	0.1319	0.1188	0.1240	0.1237	0.6656	0.6454	0.6518	0.9444	0.9527	0.9527
CEQ		-0.0009	0.0088	0.0090	0.0049	-0.0018	-0.0019	-0.0032	-0.0032	-0.0015	-0.0018	-0.0010	-0.0020	-0.0014	-0.0015	0.0107	0.0109	0.0108	0.0158	0.0157	0.0157
p-value CEQ		0.5000	0.4992	0.4992	0.5002	0.4999	0.5000	0.4999	0.5000	0.5000	0.5001	0.4999	0.5000	0.5002	0.4999	0.4992	0.4993	0.4993	0.5394	0.5304	0.5304
Turnover		0.3660	0.1199	0.1196	0.3914	0.3934	0.3869	0.3693	0.3714	0.3820	0.3784	0.3727	0.3670	0.3626	0.3632	0.0142	0.0154	0.0142	0.0981	0.0976	0.0976
M = 72																					
Alpha return	0.0216	0.0054	0.0065	0.0064	0.0105	0.0076	0.0071	0.0054	0.0054	0.0059	0.0058	0.0055	0.0050	0.0050	0.0050	0.0154	0.0163	0.0156	0.0197	0.0197	0.0197
Std dev Alpha return	0.0245	0.0439	0.0291	0.0290	0.0290	0.0428	0.0429	0.0439	0.0440	0.0438	0.0438	0.0439	0.0440	0.0440	0.0440	0.0290	0.0290	0.0292	0.0281	0.0281	0.0281
Info ratio	0.8814	0.1238	0.2222	0.2209	0.3626	0.1769	0.1652	0.1224	0.1234	0.1351	0.1333	0.1254	0.1142	0.1142	0.1143	0.5327	0.5618	0.5358	0.7015	0.7013	0.7013
p-value Info ratio		0.0662	0.3213	0.3168	0.2459	0.0825	0.0757	0.0541	0.0559	0.0694	0.0667	0.0670	0.0569	0.0561	0.0564	0.7953	0.8142	0.7868	0.9692	0.9699	0.9699
CEQ		-0.0013	0.0035	0.0035	0.0076	0.0012	0.0006	-0.0014	-0.0013	-0.0008	-0.0009	-0.0012	-0.0018	-0.0017	-0.0017	0.0125	0.0133	0.0126	0.0170	0.0169	0.0169
p-value CEQ		0.4994	0.4989	0.4988	0.4987	0.4996	0.4994	0.4994	0.4998	0.4994	0.4994	0.4994	0.4998	0.4999	0.4999	0.4987	0.4987	0.4989	0.6157	0.6093	0.6094
Turnover		0.3391	0.0681	0.0680	0.1796	0.2774	0.2882	0.3342	0.3365	0.3205	0.3191	0.3363	0.3404	0.3350	0.3360	0.0147	0.0155	0.0146	0.1286	0.1276	0.1276

Figures 3b, 3d, and 3f demonstrate that similar to figures 3a, 3c, and 3e, the linear, non-linear shrinkage, and factor models generate the highest information ratios and CEQ for $M = 12$, whereas the dynamic and dynamic-factor models achieve the highest information ratios and CEQ for longer estimation windows. This is due to the decreasing performance of the linear, non-linear, mean shrinkage, and factor models for larger estimation windows, while the performance of the dynamic and dynamic-factor models remains consistent. The consistent performance of the dynamic covariance estimators is explained by how I estimate the model. I estimate the time-variant $8 \times 8 \times 121$ for $N = 8$ assets and $T = 121$ months, where-after I use the covariance matrix at time t in the optimization for $t = 1, \ldots, T$. As a result, the dynamic covariance estimators might be less influenced by varying the estimation window, while the linear, non-linear shrinkage, and factor models are affected.

As I increase the level of risk-aversion from $\gamma = 2$ to $\gamma \simeq 22$ I expect higher alpha returns and volatilities, as more risk-seeking investors tend to invest more risky in order to obtain a higher return. Tables 5 and 8 show higher Information ratios for more risk-seeking investors with $\gamma \simeq 22$. This difference enlarges for larger estimation windows and seems to be mainly caused by the higher expected alpha returns a more risk-seeking investor attains as the volatilities are similar for both risk-aversion levels.

In summary, I see for all the different estimation windows and risk-aversion levels that the dynamic covariance estimators obtain the lowest turnover ratios. This means that accounting for time-variance results in lower trading volumes. The same result is obtained for the individual-stock setting in this study and in previous papers such as Engle (2002). They explain that accounting for time variance leads to estimates that more closely follow the market trends, thus requiring less trading.

Finally, Tables 9, 10, 11, 12, 13, and 14 present the weight statistics of all distinct models in both the individual-stock and multi-manager settings. Within the individual-stock setting, the benchmark weights exhibit the greatest disparity between the average minimum and maximum weights but at the same time yield the lowest residual sum of squares compared to the alternative models. Conversely, in the multi-manager setting, the dynamic models exhibit the smallest discrepancy between the average minimum and maximum weights and achieve the lowest residual sum of squares. Therefore, the weight analysis suggests that small discrepancies in the weights result in a lower residual sum of squares in the multi-manager setting, while the opposite is observed in the individual-stock setting. Furthermore, the dynamic-factor models in the individual-stock minimum-variance setting display a significantly larger residual sum of squares when compared to the alternative optimization approaches.



(a) $M = 12, \gamma = 2$

(b) $M = 12, \gamma = 7$

(c) $M = 36, \gamma = 2$

(d) $M = 36, \gamma = 7$

(e) $M = 72, \gamma = 2$

(f) $M = 72, \gamma = 7$

Figure 3: Empirically obtained Information ratios and CEQs from tables 5, 6 of the one-period out-of-sample monthly returns over the time period October 2012 - October 2022. The portfolios are sorted on the x-axis on descending order of the Information ratio. The Information ratios are on the left y-axis, and the CEQ on the right y-axis.

6 Conclusion

The aim of this research is to evaluate the consequences of sample estimation errors on the robustness of portfolios and to identify techniques that can minimize these errors. This is done by conducting an empirical study on daily returns in an individual-stock setting and monthly returns in a multi-manager setting. I find in both settings that estimation errors erode the performance of the portfolio. Moreover, the portfolios that incorporate dynamic covariance estimators tend to be the best-performing portfolios, although none of the alternative optimizations strictly outperforms the benchmark models at a 5% significance level. The main findings of this paper are divided into two parts: 1) main findings for the individual-stock setting that entails $N = 564$ assets over a time period of $T = 2509$ days and 2) for the multi-manager setting that entails $N = 8$ assets over a time period of $T = 121$ months.

In the individual-stock setting, the dynamic covariance estimators outperform the alternative strategies for the best hedges and minimum-variance models. In the best hedges model, the linear and non-linear covariance shrinkage estimators improve the performance of the best hedges optimization with sample estimates. On the other hand, the dynamic-factor estimators perform worse than the best hedges model with the sample estimates. In the minimum-variance model, I find that dynamic-factor model does outperform the minimum-variance model with sample estimates. This phenomena is explained by the how the best hedge function is defined. It minimizes the tracking error by accounting for factor exposures. These factors in the best hedges function are the 3 Fama-French factors thus similar to the factors in the dynamic-factor estimator. Therefore, implementing a dynamic-factor covariance estimator in the best hedges model is similar to accounting for factor exposure twice. Table 7 underlines this reasoning as it shows that lowering the hedging coefficient results in better performing dynamic-factor estimators.

Further comparing the performance in the best hedge models for both $\nu = 0.8$ and $\nu = 0.4$, I see that the best hedge models strictly outperform the minimum-variance models. The minimum-variance model with the sample covariance estimator and some more advanced covariance estimators is even outperformed by the benchmark weight portfolio. This indicates that accounting for factor risk exposure, while minimizing the tracking error results in better performing portfolios than only minimizing the volatility of a portfolio, and that an investor is better off by investing in a benchmark index than in a minimum-variance model. Furthermore, I find that increasing the hedging coefficient ν results in a higher trading volume and CEQ, which is intuitive as hedging for factor risk reduces riskiness of the portfolio but also requires additional trading.

Remarkably, the Sharpe and Information ratios of the best hedge models attain the highest values for different optimization techniques. The Sharpe ratios attain the highest values for the linear and non-linear shrinkage covariance estimators, while the Information ratios for the dynamic-covariance estimators. Moreover, as previously stated, I find that the gain in performance of the dynamic-factor covariance estimators erodes when the hedging coefficient increases. Therefore, I conclude that an investor that aims to construct the most beneficial portfolio should determine whether he values the Information ratio over the Sharpe ratio. If the latter is the case, the dynamic covariance estimators are the best-performing portfolios, while for low hedging coefficients the performance of the dynamic-factor models improve relative to the dynamic models. Also, the dynamic and dynamic-factor models require the least asset trading.

In the multi-manager setting, I find that the linear, non-linear shrinkage, and factor models perform best for the estimation window $M = 12$, while the dynamic-factor covariance estimators perform best for longer estimation windows. I find that the performance of the dynamic-factor models remains consistent over the different estimation windows, while the performance of the linear, non-linear shrinkage, and factor models erodes for longer estimation windows. Therefore, I conclude that shrinking the covariance matrix is mainly beneficial for smaller estimation windows, while accounting for dynamic factor covariance matrices more consistently outperforms the sample covariance matrices and alternative estimations for different estimation windows. This makes intuitively sense as for small estimation windows M, the ratio N/M increases with N the number of assets such that the non-linear shrinking estimators improves compared to linear shrinkage (Ledoit and Wolf 2012).

The benchmark portfolio is set to the mean-variance portfolio. I expect that mean and covariance shrinkage estimators outperform the mean-variance portfolio due to the occurrence of estimation errors in the mean and variance estimation. Moreover, I expect these estimation errors to decrease for larger estimation windows resulting in a better performing mean-variance portfolio. Compared to alternative portfolios, the mean-variance portfolio performs average, but still outperforms some mean and covariance shrinkage estimators which contradicts with our hypothesis. For shorter estimation window, the DCC, DCC-Factor, factor, and mean-shrinkage models are outperformed by the mean-variance portfolio. For larger estimation windows, the factor, linear and non-linear shrinkage models are outperformed by the mean-variance portfolio. Therefore, I conclude the mean-variance portfolio performs better than expected compared to several extensions that aim to reduce the estimation errors in the mean and covariance.

The minimum-variance portfolio performs worse than the mean-variance portfolio for the estimation window $M = 12$ months, but outperforms the mean-variance portfolio for $M = 36$ and $M = 72$ months. The worsening performance of the mean-variance portfolio relative to the minimum-variance portfolio is possibly caused by estimation errors in the mean, because the mean-variance portfolio uses mean estimation, while the minimum-variance does not. Furthermore, the turnover ratios are lower for the dynamic covariance estimators which is in line with what I observe in the individual-stock setting. Therefore, I conclude that accounting for time-variance on individual asset level reduces the trading volume. Finally, I see that the level of risk-aversion

γ does slightly impact the performance of the portfolios as I attain higher Information ratios for $\gamma \simeq 22$ than for $\gamma = 2$. This difference is mainly caused by higher alpha returns, as the volatility's remain constant.

In summary, I see that including time-variant covariance estimators improves the portfolio performance in the individual-stock and multi-manager setting. This means that estimation errors in the covariance cause sub-optimal performance of the benchmark portfolios. In the individual-stock setting the DCC-models outperform the alternative portfolios, though the DCC-Factor models in the best hedge function are outperformed. Therefore, I conclude that accounting for factor exposure twice erodes the gain from a time-variant covariance estimator. Moreover, the best hedge models outperform the minimum-variance models, thus accounting for factor exposure in the individual-stock setting is beneficial. In the multi-manager setting the DCC-Factor models outperform the alternative strategies for larger estimation windows. Finally, the estimation errors in the mean worsen the portfolio performance for larger estimation windows as the minimum-variance portfolio outperforms the mean-variance portfolio for larger estimation windows.

7 Discussion

This study focuses on optimizing mean and covariance estimation in both an individual-stock and multi-manager setting. I evaluate the impact of mean and covariance estimation errors on portfolios at APG by comparing their performance to portfolio optimization techniques aimed at reducing these errors. Portfolio optimization has been extensively researched, as theoretically optimal optimizations do not always outperform equally weighted or benchmark portfolios in practice (DeMiguel, Garlappi, and Uppal 2009; Kan and Zhou 2007). Therefore, I intend to test the portfolios in an empirical study that simulates closely a true investment setting, though this implies falling back on assumptions and restrictions.

Several optimizations incorporate factor exposure to the Fama-French factors in its optimization. The benchmark model in the individual-stock setting, the best hedge model, depends on its factor exposure determined by the varying hedging coefficient. Moreover, the 1FF, 3FF, DCC-FF, DCC-FF-LS, and DCC-FF-NLS models rely on a factor exposure. This study involves the 1 and 3 Fama-French factors. Alternatively, a study can include different factors to test whether similar results for these optimizations are obtained. This study concludes that heavily accounting for factor exposure in the best hedge function, while estimating the covariance matrix with a dynamic-factor covariance estimator is not beneficial. Further research could study whether similar results are obtained if the study accounts for distinct factors in both settings instead of accounting twice for the 3 Fama-French factors. Moreover, Chan, Karceski, and Lakonishok (1999) studies principal component analysis (PCA) in an empirical portfolio optimization study. They find that the first component already explains 75% of the variability in the assets. Therefore, further research might include PCA as a way to determine the factors instead of using the 3 Fama-French factors.

Another assumption I make in the dynamic-factor covariance estimation models is that the residual covariance matrix remains constant over time. Ledoit and Wolf (2022b) exhibit the setting with constant and time-varying covariance matrices. They even modeled the dynamic-factor models with time-varying factor exposure, though they conclude this does not improve the estimation. Capturing the time-varying component of the residual covariance matrix makes the dynamic-factor model computationally much heavier, therefore I assume it to be constant. However, this results in a residual covariance matrix that is a diagonal matrix and assumes an exact factor model (EFM) (Ledoit and Wolf 2022b). The latter assumption is often violated in practice, therefore likely impacts the performance. Further research could consider implying a sparse residual covariance matrix as a time-varying residual covariance matrix is computationally heavy, but a constant one seems to rely on strong assumptions. One could consider using the POET estimator introduced by Fan, Liao, and Mincheva (2013).

In the individual-stock setting that entails daily returns, I deal with many assets and moments in time resulting in longer computational time of the best hedge models. Therefore, I estimate the optimal weights once every month, assuming one month contains 21 trading days. This implies the proportion of the weights remains constant for these 21 trading days. However, I deal with daily returns indicating that keeping the same proportion implies a certain amount of trading. In further research, I could look at portfolio optimizations that optimize every trading day instead of once every 21 trading days. Furthermore, I could account for changes in the returns by updating the proportion of the optimized weights. Both methods might better reflect a true investment setting.

Furthermore, the results are obtained from an empirical study that entails daily and monthly returns from October 2012 - October 2022. A potential risk of conducting an empirical study for a certain time frame is a bias in the data. For example, the overall stock market did not entail any compelling financial crises. Though a crisis might influence how the portfolios perform compared to each other. For example, a minimum-variance portfolio that just minimizes a portfolio's volatility might perform better compared to a more risk-seeking portfolio. Moreover, I might see larger differences between the different parameter settings for the hedging coefficient ν and the risk-aversion level γ in the individual-stock and multi-manager settings, respectively. Another downside of an empirical study is that I was not able to test stylized settings. Therefore, further research could conduct a simulation study or an empirical study for different time periods to see whether similar results are obtained.

Looking at the results, I see that many of the performance measures are not significantly different from the APG benchmark weight portfolios at a 5% significance level. This is likely caused by the small differences between the benchmark returns and the several extension techniques. The high significance levels make it harder to draw strong conclusions on whether a certain extension truly outperforms the model in place at APG. Nonetheless, finding no extensions that strictly and significantly outperforms the benchmark models still gives valuable insights to APG. The results in this study give the first direction in searching for optimizations that do strictly and significantly outperform the best hedge and mean-variance model.

References

Andrews (1991). "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation". In: *Econometrica* 59.3, pp. 817–858. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/2938229 (visited on 02/14/2023).

Bauwens, Laurent, and Rombouts (2006). "Multivariate GARCH models: a survey". In: *Journal of applied econometrics* 21.1, pp. 79–109.

Bollerslev (1990). "Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model". In: *The review of economics and statistics*, pp. 498–505.

Brandt (2010). "Portfolio choice problems". In: *Handbook of financial econometrics: Tools and techniques*. Elsevier, pp. 269–336.

Callot, Laurent et al. (2021). "A nodewise regression approach to estimating large portfolios". In: *Journal of Business & Economic Statistics* 39.2, pp. 520–531.

Chan, Karceski, and Lakonishok (1999). "On portfolio optimization: Forecasting covariances and choosing the risk model". In: *The review of Financial studies* 12.5, pp. 937–974.

Connorand and Korajczyk (1988). "Risk and return in an equilibrium APT: Application of a new test methodology". In: *Journal of financial economics* 21.2, pp. 255–289.

DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal (2009). "Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?" In: *Review of Financial Studies* 22 (5), pp. 1915–1953. ISSN: 14657368. DOI: 10.1093/rfs/hhm075.

Engle (1982). "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation". In: *Econometrica: Journal of the econometric society*, pp. 987–1007.

— (2002). "Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models". In: *Journal of Business & Economic Statistics* 20.3, pp. 339–350.

Engle, Ledoit, and Wolf (2019). "Large dynamic covariance matrices". In: *Journal of Business & Economic Statistics* 37.2, pp. 363–375.

Fama and French (1993). "Common risk factors in the returns on stocks and bonds". In: *Journal of financial economics* 33.1, pp. 3–56.

Fan, Jianqing, Yuan Liao, and Martina Mincheva (2013). "Large covariance estimation by thresholding principal orthogonal complements". In: *Journal of the Royal Statistical Society. Series B, Statistical methodology* 75.4.

Gabor, Dennis (1946). "Theory of communication. Part 1: The analysis of information". In: *Journal of the Institution of Electrical Engineers-part III: radio and communication engineering* 93.26, pp. 429–441.

Jagannathan and Ma (2003). “Risk reduction in large portfolios: Why imposing the wrong constraints helps”. In: *The Journal of Finance* 58.4, pp. 1651–1683.

Jobson (1979). “Improved estimation for Markowitz portfolios using James-Stein type estimators”. In: 71, pp. 279–284.

Jobson and Korkie (1980). “Estimation for Markowitz efficient portfolios”. In: *Journal of the American Statistical Association* 75.371, pp. 544–554.

Jorion (1985). “International portfolio diversification with estimation risk”. In: *Journal of Business*, pp. 259–278.

— (1986). “Bayes-Stein estimation for portfolio analysis”. In: *Journal of Financial and Quantitative analysis* 21.3, pp. 279–292.

— (1991). “Bayesian and CAPM estimators of the means: Implications for portfolio selection”. In: *Journal of Banking and Finance* 15.3, pp. 717–727. ISSN: 0378-4266. DOI: https://doi.org/10.1016/0378-4266(91)90094-3. URL: https://www.sciencedirect.com/science/article/pii/0378426691900943.

Kan and Zhou (2007). “Optimal portfolio choice with parameter uncertainty”. In: *Journal of Financial and Quantitative Analysis* 42.3, pp. 621–656.

Kirby and Ostdiek (2010). “It’s All in the Timing: Simple Active Portfolio Strategies that Outperform Naive Diversification”. In: *Available at SSRN 1530022*.

Ledoit, Olivier and Sandrine Péché (2011). “Eigenvectors of some large sample covariance matrix ensembles”. In: *Probability Theory and Related Fields* 151.1-2, pp. 233–264.

Ledoit and Wolf (2003). “Improved estimation of the covariance matrix of stock returns with an application to portfolio selection”. In: *Journal of empirical finance* 10.5, pp. 603–621.

— (2004a). “A well-conditioned estimator for large-dimensional covariance matrices”. In: *Journal of multivariate analysis* 88.2, pp. 365–411.

— (2004b). “Honey, I shrunk the sample covariance matrix”. In: *The Journal of Portfolio Management* 30.4, pp. 110–119.

— (2008). “Robust performance hypothesis testing with the Sharpe ratio”. In: *Journal of Empirical Finance* 15.5, pp. 850–859.

— (2012). “Nonlinear shrinkage estimation of large-dimensional covariance matrices”. In.

— (2015). “Spectrum estimation: A unified framework for covariance matrix estimation and PCA in large dimensions”. In: *Journal of Multivariate Analysis* 139, pp. 360–384. ISSN: 0047-259X. DOI: https://doi.org/10.1016/j.jmva.2015.04.006. URL: https://www.sciencedirect.com/science/article/pii/S0047259X15000949.

Ledoit and Wolf (2020). “Analytical nonlinear shrinkage of large-dimensional covariance matrices”. In: *The Annals of Statistics* 48.5, pp. 3043–3065.

— (2022a). “Quadratic shrinkage for large covariance matrices”. In: *Bernoulli* 28.3, pp. 1519–1547.

— (2022b). “The power of (non-) linear shrinking: A review and guide to covariance matrix estimation”. In: *Journal of Financial Econometrics* 20.1, pp. 187–218.

Ledoit, Wolf, and Zhao (2019). “Efficient sorting: A more powerful test for cross-sectional anomalies”. In: *Journal of Financial Econometrics* 17.4, pp. 645–686.

Markowitz, Harry (1952). “Portfolio Selection”. In: *Source: The Journal of Finance* 7 (1), pp. 77–91.

Meinshausen, Nicolai and Peter Bühlmann (2006). “High-dimensional graphs and variable selection with the lasso”. In.

Merton (1980). “On estimating the expected return on the market: An exploratory investigation”. In: *Journal of financial economics* 8.4, pp. 323–361.

Michaud (1989). “The Markowitz optimization enigma: Is ‘optimized’ optimal?” In: *Financial analysts journal* 45.1, pp. 31–42.

Nard, De, Ledoit, and Wolf (2021). “Factor models for portfolio selection in large dimensions: The good, the better and the ugly”. In: *Journal of Financial Econometrics* 19.2, pp. 236–257.

Pástor (2000). “Portfolio selection and asset pricing models”. In: *The Journal of Finance* 55.1, pp. 179–223.

Sharpe (1963). “A simplified model for portfolio analysis”. In: *Management science* 9.2, pp. 277–293.

Stein (1956). “Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution”. In.

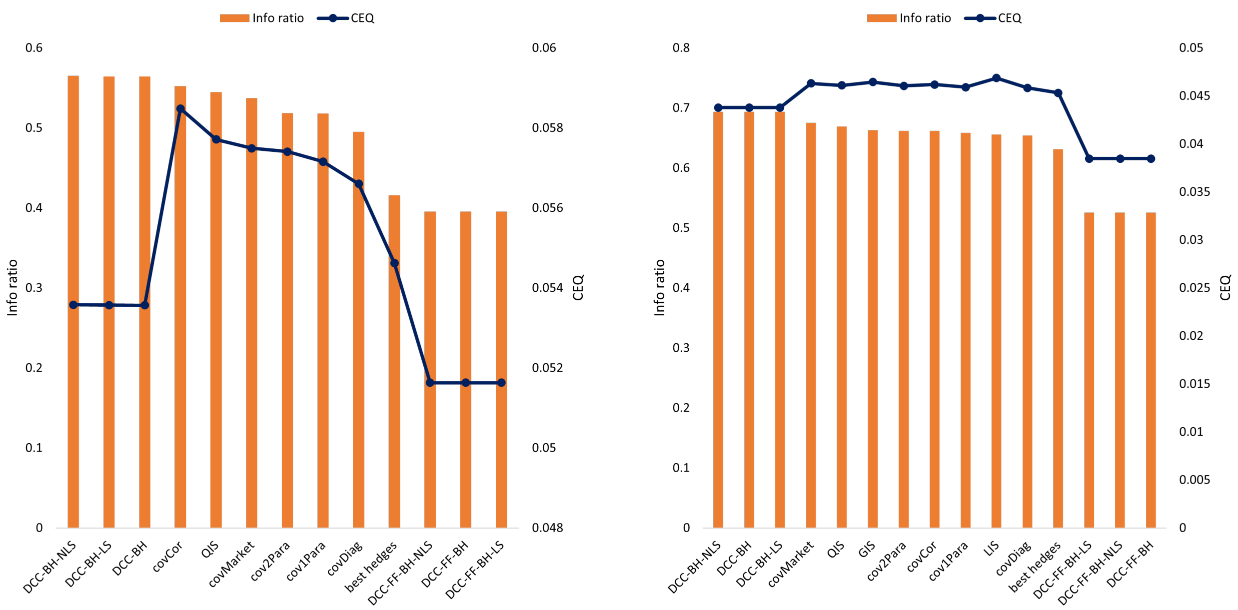
Stevens, Guy VG (1998). “On the inverse of the covariance matrix in portfolio analysis”. In: *The Journal of Finance* 53.5, pp. 1821–1827.

A Empirical result for best hedge model with hedging coefficient 0.4

Table 7: Annualized empirical excess and alpha returns of the best hedge model (equation 1) with $v = 0.4$ from Oct 2012 - Oct 2022 such that $T = 2509$ days. The number of assets equals 1197 for the benchmark portfolio and 564 for the portfolio. The estimation windows differ between $M = 250$ and $M = 750$ days. The returns are the one-period out-of-sample returns, and the optimal weights are updated once every 21 trading days. For the excess returns I subtract the risk-free rate, and for alpha returns the bmk weight returns. All the models are different estimations of the covariance matrix in the best hedge model, where 'best hedges' uses the sample covariance matrix. If CEQ equals 0.01, it indicates an investor values the portfolio at a 1% risk-free return. A turnover of 0.01 indicates 1% of the wealth is traded on a daily basis.

	Benchmark		*Linear shrinkage*					*Non-linear shrinkage*			*DCC*					
	bmk	best hedges	LS cov1Para	LS cov2Para	LS covCor	LS covDiag	LS covMkt	NLS LIS	NLS QIS	NLS GIS	DCC	DCC-LS	DCC-NLS	DCC-3FF	DCC-3FF-LS	DCC-3FF-NLS
								M = 250								
Excess return (ER)	0.1446	0.1545	0.1574	0.1573	0.1581	0.1569	0.1577		0.1580		0.1606	0.1606	0.1606	0.1557	0.1557	0.1557
Std dev ER	0.1532	0.1648	0.1651	0.1648	0.1646	0.1651	0.1651		0.1651		0.1706	0.1706	0.1706	0.1682	0.1682	0.1682
Sharpe ratio	0.9441	0.9375	0.9533	0.9544	0.9607	0.9499	0.9553		0.9567		0.9414	0.9414	0.9416	0.9256	0.9256	0.9256
p-value Sharpe ratio		0.8727	0.8317	0.8121	0.7050	0.8917	0.7932		0.7697		0.9521	0.9523	0.9538	0.7076	0.7076	0.7076
Alpha		0.0099	0.0127	0.0127	0.0135	0.0122	0.0131		0.0134		0.0160	0.0160	0.0160	0.0111	0.0111	0.0111
Std dev Alpha		0.0237	0.0246	0.0244	0.0245	0.0247	0.0244		0.0245		0.0284	0.0284	0.0284	0.0281	0.0281	0.0281
Info ratio		0.4158	0.5181	0.5185	0.5523	0.4952	0.5373		0.5449		0.5643	0.5645	0.5651	0.3954	0.3954	0.3954
p-value Info ratio		0.1717	0.2770	0.2807	0.3232	0.2531	0.2954		0.3055		0.2655	0.2656	0.2664	0.1448	0.1448	0.1448
CEQ		0.0546	0.0572	0.0574	0.0585	0.0566	0.0575		0.0577		0.0536	0.0536	0.0536	0.0516	0.0516	0.0516
p-value CEQ		0.7854	0.9274	0.9438	0.9898	0.8927	0.9492		0.9618		0.7393	0.7393	0.7398	0.6504	0.6504	0.6504
Turnover		0.0276	0.0175	0.0172	0.0177	0.0179	0.0168		0.0165		0.0147	0.0147	0.0147	0.0189	0.0189	0.0189
								M = 750								
Excess return (ER)	0.1438	0.1607	0.1615	0.1615	0.1615	0.1613	0.1619	0.1611	0.1617	0.1614	0.1642	0.1642	0.1642	0.1604	0.1604	0.1604
Std dev ER	0.1617	0.1772	0.1773	0.1772	0.1771	0.1772	0.1773	0.1762	0.1773	0.1768	0.1810	0.1810	0.1810	0.1821	0.1821	0.1821
Sharpe ratio	0.8895	0.9073	0.9109	0.9115	0.9120	0.9104	0.9133	0.9141	0.9122	0.9129	0.9074	0.9074	0.9074	0.8810	0.8810	0.8810
p-value Sharpe ratio		0.7052	0.6498	0.6400	0.6332	0.6569	0.6147	0.6116	0.6306	0.6220	0.6935	0.6935	0.6930	0.8733	0.8733	0.8733
Alpha		0.0169	0.0177	0.0177	0.0177	0.0175	0.0181	0.0173	0.0179	0.0176	0.0204	0.0204	0.0204	0.0166	0.0166	0.0166
Std dev Alpha		0.0268	0.0268	0.0268	0.0268	0.0268	0.0268	0.0264	0.0268	0.0266	0.0294	0.0294	0.0294	0.0317	0.0317	0.0317
Info ratio		0.6313	0.6584	0.6619	0.6617	0.6540	0.6755	0.6557	0.6690	0.6631	0.6933	0.6933	0.6935	0.5256	0.5256	0.5256
p-value Info ratio		0.5240	0.5676	0.5743	0.5757	0.5608	0.5967	0.5793	0.5849	0.5826	0.5840	0.5840	0.5844	0.3239	0.3239	0.3239
CEQ		0.0453	0.0459	0.0460	0.0462	0.0458	0.0463	0.0469	0.0461	0.0464	0.0438	0.0438	0.0438	0.0385	0.0385	0.0385
p-value CEQ		0.8663	0.9005	0.9072	0.9132	0.8968	0.9221	0.9535	0.9101	0.9292	0.7967	0.7968	0.7969	0.6030	0.6030	0.6030
Turnover		0.0183	0.0168	0.0166	0.0167	0.0171	0.0165	0.0164	0.0161	0.0162	0.0152	0.0152	0.0152	0.0172	0.0172	0.0172

B Information ratios and CEQs for best hedge model with hedging coefficient 0.4



(a) Best hedges model with $v = 0.4$ and $M = 250$

(b) Best hedges model with $v = 0.4$ and $M = 750$

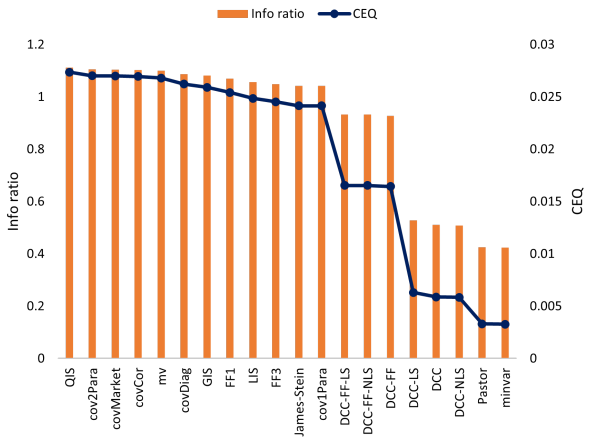
Figure 4: Empirically obtained Information ratios and CEQs from appendix A of the one-period out-of-sample daily returns of the best hedge model with $v = 0.4$ over the time period October 2012 - October 2022. The portfolios are sorted on the x-axis on descending order of the Information ratio. The Information ratios are on the left y-axis, and the CEQ on the right y-axis.

C Empirical result for mean-variance model with implied risk-aversion

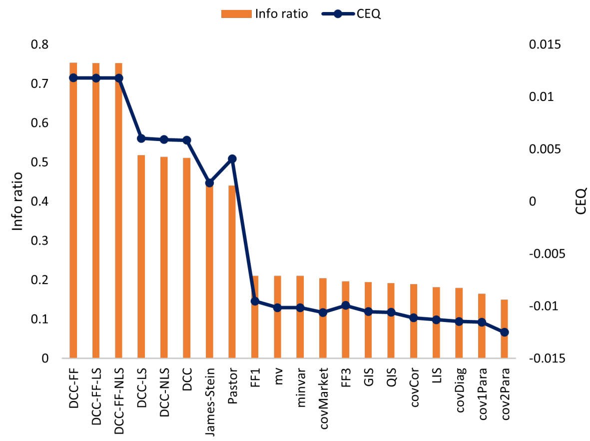
Table 8: Annualized empirical excess and alpha returns of the mean-variance model (equation 10) with implied level of $\gamma \simeq 22$ from Oct 2012 - Oct 2022 such that $T = 121$ months. The number of assets $N = 8$. The estimation windows differ between $M = 12, 36, 72$ months. The returns are the one-period, i.e., one-month, out-of-sample alpha returns. APG subtracts the respective geographical MSCI Index to obtain the alpha returns. All the models are different estimations of the mean or covariance in the mean-variance model. If CEQ equals 0.01, it indicates an investor values the portfolio at a 1% risk-free return. A turnover of 0.01 indicates 1% of the wealth is traded on a daily basis.

	Benchmark		*Mean shrinkage*			*Factor models*		*Linear shrinkage*					*Non-linear shrinkage*			*DCC*					
	bmk	mv	minvar	Pastor	James-Stein	1FF	3FF	LS cov1Para	LS cov2Para	LS covCor	LS covDiag	LS covMkt	NLS LIS	NLS QIS	NLS GIS	DCC	DCC-LS	DCC-NLS	DCC-3FF	DCC-3FF-LS	DCC-3FF-NLS
M = 12																					
Alpha return	0.0192	0.0468	0.0117	0.0117	0.0426	0.0449	0.0440	0.0426	0.0470	0.0473	0.0461	0.0472	0.0443	0.0477	0.0454	0.0123	0.0127	0.0123	0.0236	0.0237	0.0237
Std dev Alpha return	0.0302	0.0425	0.0276	0.0276	0.0408	0.0420	0.0420	0.0408	0.0426	0.0429	0.0424	0.0428	0.0420	0.0429	0.0420	0.0240	0.0240	0.0241	0.0254	0.0254	0.0254
Info ratio	0.6382	1.1003	0.4241	0.4250	1.0429	1.0694	1.0483	1.0429	1.1052	1.1027	1.0869	1.1039	1.0563	1.1117	1.0814	0.5107	0.5282	0.5092	0.9276	0.9326	0.9326
p-value Info ratio		0.3335	0.4598	0.4617	0.3926	0.3650	0.3848	0.3926	0.3361	0.3281	0.3454	0.3288	0.3820	0.3221	0.3564	0.7227	0.7363	0.7042	0.3061	0.3000	0.3000
CEQ		0.0268	0.0033	0.0033	0.0242	0.0254	0.0245	0.0242	0.0270	0.0270	0.0262	0.0270	0.0249	0.0274	0.0259	0.0059	0.0063	0.0059	0.0164	0.0165	0.0165
p-value CEQ		0.5013	0.5002	0.5002	0.5018	0.5012	0.5013	0.5018	0.5013	0.5021	0.5012	0.5020	0.5012	0.5022	0.5013	0.5012	0.5021	0.5015	0.5002	0.4999	0.4999
Turnover		0.5232	0.3469	0.3470	0.5372	0.5191	0.5374	0.5372	0.4989	0.5229	0.5228	0.5297	0.4793	0.4937	0.4735	0.0124	0.0135	0.0124	0.1482	0.1451	0.1451
M = 36																					
Alpha return	0.0202	0.0087	0.0087	0.0111	0.0161	0.0085	0.0079	0.0067	0.0062	0.0078	0.0075	0.0085	0.0075	0.0078	0.0080	0.0129	0.0131	0.0130	0.0184	0.0184	0.0184
Std dev Alpha return	0.0473	0.0413	0.0413	0.0252	0.0359	0.0404	0.0402	0.0406	0.0411	0.0414	0.0414	0.0416	0.0412	0.0408	0.0409	0.0253	0.0253	0.0253	0.0244	0.0245	0.0245
Info ratio	0.4276	0.2107	0.2107	0.4408	0.4468	0.2108	0.1966	0.1651	0.1501	0.1893	0.1804	0.2045	0.1814	0.1922	0.1949	0.5114	0.5184	0.5138	0.7541	0.7530	0.7530
p-value Info ratio		0.1521	0.1521	0.4753	0.4441	0.1207	0.1158	0.0986	0.0963	0.1273	0.1171	0.1470	0.1202	0.1236	0.1272	0.6654	0.6453	0.6516	0.8848	0.8876	0.8876
CEQ		-0.0101	-0.0101	0.0041	0.0018	-0.0095	-0.0099	-0.0115	-0.0125	-0.0111	-0.0114	-0.0106	-0.0113	-0.0106	-0.0105	0.0059	0.0060	0.0059	0.0118	0.0118	0.0118
p-value CEQ		0.4984	0.4984	0.4995	0.4983	0.4991	0.4989	0.4991	0.4985	0.4982	0.4983	0.4983	0.4986	0.4988	0.4987	0.4995	0.4995	0.4995	0.7016	0.7087	0.7087
Turnover		0.2777	0.2777	0.1198	0.2862	0.2791	0.2875	0.2728	0.2810	0.2889	0.2828	0.2871	0.2809	0.2739	0.2742	0.0142	0.0154	0.0142	0.0885	0.0885	0.0885
M = 72																					
Alpha return	0.0216	0.0037	0.0065	0.0064	0.0102	0.0092	0.0080	0.0041	0.0041	0.0045	0.0043	0.0036	0.0037	0.0038	0.0038	0.0154	0.0163	0.0156	0.0216	0.0216	0.0216
Std dev Alpha return	0.0245	0.0385	0.0291	0.0291	0.0276	0.0376	0.0375	0.0385	0.0388	0.0385	0.0383	0.0386	0.0385	0.0381	0.0382	0.0290	0.0290	0.0292	0.0266	0.0266	0.0266
Info ratio	0.8814	0.0962	0.2222	0.2218	0.3712	0.2457	0.2138	0.1063	0.1047	0.1182	0.1133	0.0942	0.0948	0.1010	0.0998	0.5326	0.5616	0.5357	0.8128	0.8126	0.8126
p-value Info ratio		0.0510	0.3213	0.3199	0.2685	0.1151	0.0935	0.0443	0.0444	0.0524	0.0494	0.0498	0.0433	0.0453	0.0448	0.7952	0.8141	0.7867	0.7260	0.7272	0.7272
CEQ		-0.0127	-0.0029	-0.0029	0.0018	-0.0064	-0.0075	-0.0123	-0.0126	-0.0118	-0.0119	-0.0128	-0.0128	-0.0122	-0.0123	0.0062	0.0070	0.0062	0.0138	0.0138	0.0138
p-value CEQ		0.4979	0.4997	0.4998	0.5003	0.4992	0.4990	0.4979	0.4986	0.4979	0.4985	0.4981	0.4980	0.4982	0.4981	0.4997	0.4999	0.4997	0.5011	0.5010	0.5010
Turnover		0.2635	0.0681	0.0680	0.1228	0.1781	0.1929	0.2498	0.2544	0.2543	0.2548	0.2618	0.2630	0.2605	0.2609	0.0147	0.0155	0.0146	0.1028	0.1027	0.1027

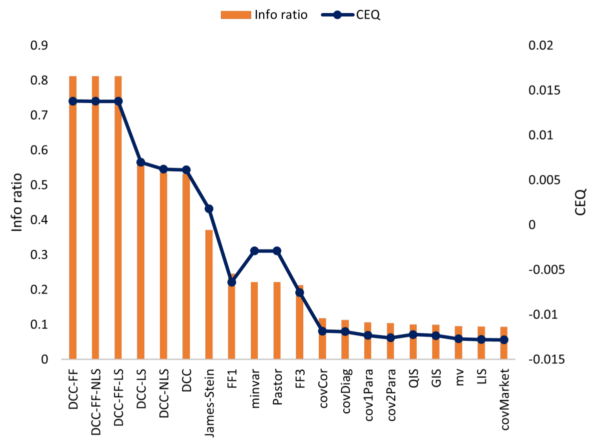
D Information ratios and CEQs for mean-variance model with implied risk-aversion



(a) $M = 12$, implied γ



(b) $M = 36$, implied γ



(c) $M = 72$, implied γ

Figure 5: Empirically obtained Information ratios and CEQs from appendix C of the one-period out-of-sample monthly returns of mean-variance models over the time period October 2012 - October 2022. The portfolios are sorted on the x-axis on descending order of the Information ratio. The Information ratios are on the left y-axis, and the CEQ on the right y-axis.

E Weight statistics of individual-stock best hedge model with hedging coefficient 0.8

Table 9: This table shows the statistics of the portfolio weights of the best hedge model (equation 1) with $v = 0.8$ from Oct 2012 - Oct 2022 such that $T = 2509$ days. The best hedge model optimally allocates over 564 assets and accounts for 1197 assets in the benchmark. The estimation windows differ between $M = 250$ and $M = 750$ days. The weights are updated once every 21 trading days. The average portfolio weights equals $\bar{w_i} \times 100 = 0.1773$. The table shows the average minimum and maximum portfolio weights, $\min w_i \times 100$ and $\max w_i \times 100$, and the residual sum of squares, $\sum(w_i - \bar{w_i})^2$.

	$\min w_i \times 100$	$\max w_i \times 100$	$\sum(w_i - \bar{w_i})^2$	$\min w_i \times 100$	$\max w_i \times 100$	$\sum(w_i - \bar{w_i})^2$
Models		**M = 250**			**M = 750**	
bmk	0.0022	9.7055	0.0190	0.0022	9.7055	0.0190
best hedges	0.0378	0.7346	42.0493	0.0835	0.3490	32.5147
LS cov1Para	0.0960	0.3315	41.8764	0.1052	0.2907	32.2800
LS cov2Para	0.1042	0.3062	41.8371	0.1124	0.2744	32.1750
LS covCor	0.0961	0.3459	42.0805	0.1081	0.2887	32.4090
LS covDiag	0.0852	0.3739	41.9636	0.0996	0.3063	32.3739
LS covMkt	0.0966	0.3256	41.9904	0.1079	0.2848	32.3683
NLS LIS				0.1165	0.2658	31.9647
NLS QIS	0.1273	0.2506	42.3791	0.1216	0.2558	31.9140
NLS GIS				0.1351	0.2319	32.7169
DCC	0.1700	0.1844	42.9691	0.1737	0.1806	33.4582
DCC-LS	0.1702	0.1841	42.9685	0.1738	0.1805	33.4578
DCC-NLS	0.1722	0.1821	42.9644	0.1748	0.1795	33.4552
DCC-3FF	0.0477	0.5153	43.2751	0.1376	0.2188	9.0307
DCC-3FF-LS	0.0477	0.5153	43.2752	0.0458	0.4873	34.9641
DCC-3FF-NLS	0.0477	0.5153	43.2751	0.0458	0.4873	34.9640

F Weight statistics of individual-stock best hedge model with hedging coefficient 0.4

Table 10: This table shows the statistics of the portfolio weights of the best hedge model (equation 1) with $\nu = 0.4$ from Oct 2012 - Oct 2022 such that $T = 2509$ days. The best hedge model optimally allocates over 564 assets and accounts for 1197 assets in the benchmark. The estimation windows differ between $M = 250$ and $M = 750$ days. The weights are updated once every 21 trading days. The average portfolio weights equals $\bar{w}_i \times 100 = 0.1773$. The table shows the average minimum and maximum portfolio weights, $\min w_i \times 100$ and $\max w_i \times 100$, and the residual sum of squares, $\sum(w_i - \bar{w}_i)^2$.

	$\min w_i \times 100$	$\max w_i \times 100$	$\sum(w_i - \bar{w}_i)^2$	$\min w_i \times 100$	$\max w_i \times 100$	$\sum(w_i - \bar{w}_i)^2$
Models		**M = 250**			**M = 750**	
bmk	0.0022	9.7055	0.0190	0.0022	9.7055	0.0190
best hedges	0.0562	0.4825	35.8673	0.1109	0.2817	34.3102
LS cov1Para	0.1218	0.2649	42.2856	0.1295	0.2425	34.0106
LS cov2Para	0.1244	0.2590	42.2710	0.1335	0.2355	33.9612
LS covCor	0.1169	0.2760	42.1784	0.1315	0.2394	33.9463
LS covDiag	0.1169	0.2817	42.3173	0.1259	0.2498	34.0639
LS covMkt	0.1277	0.2516	42.0777	0.1340	0.2347	33.9506
NLS LIS				0.1351	0.1773	33.9341
NLS QIS	0.1338	0.2385	42.0726	0.1415	0.2226	33.8388
NLS GIS				0.1387	0.2268	33.8794
DCC	0.1759	0.1786	42.9634	0.1767	0.1779	33.4540
DCC-LS	0.1760	0.1786	42.9633	0.1767	0.1778	33.4539
DCC-NLS	0.1764	0.1782	42.9627	0.1769	0.1777	33.4536
DCC-3FF	0.0615	0.4117	42.9419	0.0615	0.4117	42.9419
DCC-3FF-LS	0.0615	0.4117	42.9419	0.0749	0.3496	33.8851
DCC-3FF-NLS	0.0615	0.4117	42.9419	0.0749	0.3496	33.8851

G Weight statistics of the individual-stock minimum-variance model

Table 11: This table shows the statistics of the portfolio weights of the minimum-variance model (equation 17) from Oct 2012 - Oct 2022 such that $T = 2509$ days. The number of assets equals 564. The estimation windows differ between $M = 250$ and $M = 750$ days. The weights are updated once every 21 trading days. The average portfolio weights equals $\bar{w}_i \times 100 = 0.1773$. The table shows the average minimum and maximum portfolio weights, $\min w_i \times 100$ and $\max w_i \times 100$, and the residual sum of squares, $\sum(w_i - \bar{w}_i)^2$.

	$\min w_i \times 100$	$\max w_i \times 100$	$\sum(w_i - \bar{w}_i)^2$	$\min w_i \times 100$	$\max w_i \times 100$	$\sum(w_i - \bar{w}_i)^2$
Models		**M = 250**			**M = 750**	
bmk	0.0022	9.7055	0.0190	0.0022	9.7055	0.0190
minvar	0.0000	1.7390	97.2458	0.0124	0.6994	64.9317
FF1	0.0040	0.7540	28.0893	0.0311	0.4437	24.6311
FF3	0.0044	0.7746	29.2853	0.0046	0.7119	21.9592
LS cov1Para	0.0000	1.4255	61.5580	0.0142	0.6678	56.1427
LS cov2Para	0.0000	1.3381	53.4485	0.0151	0.6548	52.2004
LS covCor	0.0008	1.5294	109.3205	0.0151	0.6933	70.9866
LS covDiag	0.0000	1.5882	84.2461	0.0134	0.6853	62.0457
LS covMkt	0.0004	1.3919	77.4595	0.0159	0.6577	60.9057
NLS LIS				0.0193	0.5972	40.6853
NLS QIS	0.0006	1.1887	46.1687	0.0199	0.5994	45.9761
NLS GIS				0.0196	0.5982	43.2692
DCC	0.0177	1.2975	49.5167	0.1737	0.1806	33.4582
DCC-LS	0.0174	1.3064	49.1297	0.0146	1.1906	37.8671
DCC-NLS	0.0147	1.4031	49.4944	0.0123	1.2786	38.0706
DCC-3FF	0.0096	0.3480	1140.7420	0.0100	0.3462	907.5020
DCC-3FF-LS	0.0089	0.3484	1138.9148	0.0094	0.3466	906.2498
DCC-3FF-NLS	0.0094	0.3481	1139.9695	0.0099	0.3463	906.9614

H Weight statistics of multi-manager model with risk-aversion of 2

Table 12: This table shows the statistics of the portfolio weights of the mean-variance model (equation 10) with risk-aversion $\gamma = 2$ from Oct 2012 - Oct 2022 such that $T = 121$ months. The number of assets equals $N = 8$. The estimation windows differ between $M = 12, 36, 72$ months. The weights are updated every month. The average portfolio weights equals $\bar{w}_i \times 100 = 12.5$. The table shows the average minimum and maximum portfolio weights, $\min w_i \times 100$ and $\max w_i \times 100$, and the residual sum of squares, $\sum(w_i - \bar{w}_i)^2$.

	$\min w_i \times 100$	$\max w_i \times 100$	$\sum(w_i - \bar{w}_i)^2$	$\min w_i \times 100$	$\max w_i \times 100$	$\sum(w_i - \bar{w}_i)^2$	$\min w_i \times 100$	$\max w_i \times 100$	$\sum(w_i - \bar{w}_i)^2$
Models		**M = 12**			**M = 36**			**M = 72**	
bmk	9.1342	17.9727	13.4660	9.1342	17.9727	13.4660	9.1342	17.9727	13.4660
mv	0.0000	100.0000	94.3696	0.0000	75.0000	67.2251	0.0000	50.0000	39.6474
minvar	0.0000	46.0728	18.3264	0.4967	33.6536	10.2682	6.9449	20.2666	4.2750
Pastor	0.0000	45.5682	18.1029	0.3033	33.3824	10.0466	6.9458	20.1712	4.2440
James-Stein	0.0000	100.0000	91.3242	0.0000	73.3635	58.9995	1.5030	38.0468	12.0800
1FF	0.0000	100.0000	93.9067	0.0000	75.0000	67.1914	0.0000	50.0000	35.0524
3FF	0.0000	100.0000	94.0041	0.0000	75.1820	66.6924	0.0000	50.0000	35.5045
LS cov1Para	0.0000	100.0000	93.2662	0.0000	75.5777	67.2064	0.0000	50.0000	38.3514
LS cov2Para	0.0000	100.0000	93.8112	0.0000	75.0000	67.6310	0.0000	50.0000	38.7299
LS covCor	0.0000	100.0000	94.2185	0.0000	75.0000	67.5782	0.0000	50.0000	38.7935
LS covDiag	0.0000	100.0000	94.1556	0.0000	75.0000	67.1875	0.0000	50.0000	38.5772
LS covMkt	0.0000	100.0000	94.3561	0.0000	75.0000	67.5020	0.0000	50.0000	39.5074
NLS LIS	0.0000	100.0000	92.5842	0.0000	75.0000	67.5655	0.0000	50.0000	39.3660
NLS QIS	0.0000	100.0000	93.8798	0.0000	75.0000	67.1548	0.0000	50.0000	39.1177
NLS GIS	0.0000	100.0000	93.3256	0.0000	75.0000	67.1973	0.0000	50.0000	39.1709
DCC	12.4582	12.5505	9.3705	12.4596	12.5485	7.3059	12.4647	12.5345	4.2109
DCC-LS	12.0401	13.1149	5.2389	12.0535	13.1005	4.0842	12.0976	12.9449	2.3224
DCC-NLS	12.4518	12.5650	7.4968	12.4525	12.5626	5.8447	12.4557	12.5480	3.3653
DCC-3FF	0.0000	59.2419	31.9887	0.0000	37.7078	21.3249	4.5158	27.3003	13.1203
DCC-3FF-LS	0.0000	60.6694	31.3886	0.0000	37.6044	20.4036	4.9763	26.5479	12.8377
DCC-3FF-NLS	0.0000	60.6698	31.3885	0.0000	37.6050	20.4028	4.9765	26.5475	12.8377

I Weight statistics of multi-manager model with risk-aversion of 7

Table 13: This table shows the statistics of the portfolio weights of the mean-variance model (equation 10) with risk-aversion $\gamma = 7$ from Oct 2012 - Oct 2022 such that $T = 121$ months. The number of assets equals $N = 8$. The estimation windows differ between $M = 12, 36, 72$ months. The weights are updated every month. The average portfolio weights equals $\bar{w}_i \times 100 = 12.5$. The table shows the average minimum and maximum portfolio weights, $\min w_i \times 100$ and $\max w_i \times 100$, and the residual sum of squares, $\sum(w_i - \bar{w}_i)^2$.

	$\min w_i \times 100$	$\max w_i \times 100$	$\sum(w_i - \bar{w}_i)^2$	$\min w_i \times 100$	$\max w_i \times 100$	$\sum(w_i - \bar{w}_i)^2$	$\min w_i \times 100$	$\max w_i \times 100$	$\sum(w_i - \bar{w}_i)^2$
Models		**M = 12**			**M = 36**			**M = 72**	
bmk	9.1342	17.9727	13.4660	9.1342	17.9727	13.4660	9.1342	17.9727	13.4660
mv	0.0000	100.0000	88.0749	0.0000	73.5214	57.0047	0.0000	51.2748	28.1667
minvar	0.0000	46.0728	18.3264	0.4967	33.6536	10.2682	6.9449	20.2666	4.2750
Pastor	0.0000	45.9535	18.2635	0.4408	33.5513	10.1990	6.9451	20.2381	4.2648
James-Stein	0.0000	100.0000	80.3262	0.0000	67.7011	43.0017	1.5030	32.9426	9.6887
1FF	0.0000	100.0000	85.4449	0.0000	75.1209	53.1100	0.0000	43.3900	19.9663
3FF	0.0000	100.0000	86.1910	0.0000	75.1023	53.8166	0.0000	44.9550	20.7267
LS cov1Para	0.0000	100.0000	84.7040	0.0000	75.5905	54.7741	0.0000	49.4809	24.4652
LS cov2Para	0.0000	100.0000	86.3625	0.0000	75.9648	55.8718	0.0000	49.7795	25.2733
LS covCor	0.0000	100.0000	87.2407	0.0000	74.1995	57.0184	0.0000	49.0063	26.9081
LS covDiag	0.0000	100.0000	86.2061	0.0000	74.8641	56.0854	0.0000	49.0388	26.1747
LS covMkt	0.0000	100.0000	88.5094	0.0000	73.6630	57.5589	0.0000	50.7960	28.0696
NLS LIS	0.0000	100.0000	79.3456	0.0000	73.9476	57.0495	0.0000	50.4226	27.5343
NLS QIS	0.0000	100.0000	85.9220	0.0000	74.1016	56.2403	0.0000	50.1861	27.0049
NLS GIS	0.0000	100.0000	80.8946	0.0000	74.0920	56.2940	0.0000	50.2117	27.1386
DCC	12.4582	12.5505	9.3871	12.4596	12.5484	7.3188	12.4647	12.5345	4.2183
DCC-LS	12.0400	13.1148	5.2568	12.0533	13.1003	4.0982	12.0974	12.9448	2.3303
DCC-NLS	12.4518	12.5650	7.5181	12.4525	12.5626	5.8612	12.4557	12.5479	3.3748
DCC-3FF	0.0334	43.0867	27.2582	5.3364	26.8614	20.3110	7.1836	22.1520	14.0083
DCC-3FF-LS	0.1603	42.4525	26.1527	5.5337	26.5654	20.2373	7.2340	22.0211	13.9832
DCC-3FF-NLS	0.1602	42.4524	26.1527	5.5340	26.5651	20.2373	7.2341	22.0209	13.9832

J Weight statistics of multi-manager model with implied risk-aversion

Table 14: This table shows the statistics of the portfolio weights of the mean-variance model (equation 10) with implied level of risk-aversion $\gamma \simeq 22$ from Oct 2012 - Oct 2022 such that $T = 121$ months. The number of assets equals $N = 8$. The estimation windows differ between $M = 12, 36, 72$ months. The weights are updated every month. The average portfolio weights equals $\bar{w}_i \times 100 = 12.5$. The table shows the average minimum and maximum portfolio weights, $\min w_i \times 100$ and $\max w_i \times 100$, and the residual sum of squares, $\sum(w_i - \bar{w}_i)^2$.

	$\min w_i \times 100$	$\max w_i \times 100$	$\sum(w_i - \bar{w}_i)^2$	$\min w_i \times 100$	$\max w_i \times 100$	$\sum(w_i - \bar{w}_i)^2$	$\min w_i \times 100$	$\max w_i \times 100$	$\sum(w_i - \bar{w}_i)^2$
Models		**M = 12**			**M = 36**			**M = 72**	
bmk	9.1342	17.9727	13.4660	9.1342	17.9727	13.4660	9.1342	17.9727	13.4660
mv	0.0000	100.0000	68.5784	0.0000	61.8884	38.1169	0.8903	41.2679	16.0656
minvar	0.0000	46.0728	18.3264	0.0000	61.8884	38.1169	6.9449	20.2666	4.2750
Pastor	0.0000	46.0428	18.3071	0.4787	33.6212	10.2472	6.9450	20.2574	4.2717
James-Stein	0.0000	99.4282	60.0563	0.0000	56.8674	24.8213	3.5195	26.1646	6.2241
1FF	0.0000	99.1888	59.7689	0.0000	57.6286	29.1813	1.4759	30.9991	9.7252
3FF	0.0000	99.9840	62.7995	0.0000	59.1910	30.7177	0.7875	32.7044	10.0121
LS cov1Para	0.0000	99.4282	60.0563	0.0000	60.6734	32.1056	0.6589	37.6954	11.4962
LS cov2Para	0.0000	96.6995	61.9248	0.0000	61.7403	33.9040	0.5104	38.4415	12.1539
LS covCor	0.0000	99.3604	65.9673	0.0000	61.1597	36.8029	0.7600	38.7622	14.7660
LS covDiag	0.0000	99.0104	62.6680	0.0000	60.9851	34.9951	0.7697	38.4320	13.8583
LS covMkt	0.0000	100.0000	68.1164	0.0000	62.0923	38.5018	0.7674	41.0237	16.1244
NLS LIS	0.0000	95.6783	57.1839	0.0000	62.3534	37.7270	0.3750	40.2822	14.4973
NLS QIS	0.0000	98.4915	63.5369	0.0000	61.3350	36.6175	0.5496	39.4894	13.9654
NLS GIS	0.0000	95.1596	57.9113	0.0000	61.4177	36.6100	0.5210	39.6596	14.1098
DCC	12.4582	12.5505	9.3916	12.4596	12.5484	7.3223	12.4647	12.5345	4.2204
DCC-LS	12.0400	13.1148	5.2618	12.0533	13.1003	4.1021	12.0974	12.9448	2.3324
DCC-NLS	12.4518	12.5650	7.5239	12.4525	12.5626	5.8658	12.4557	12.5479	3.3774
DCC-3FF	4.7601	30.0239	26.2161	7.1623	23.5376	21.1161	7.4011	21.2781	12.1224
DCC-3FF-LS	4.9107	29.7150	26.3315	7.2858	23.5090	21.1921	7.4234	21.2513	12.1449
DCC-3FF-NLS	4.9107	29.7150	26.3315	7.2859	23.5090	21.1921	7.4234	21.2513	12.1449