Naive Diversification with Fewer Assets
-
A Risk Reduction Approach Using Clustering Methods

Anna-Katharina Thös

Gutachter:	Prof. Dr. Jörn Saß
	Technische Universität Kaiserslautern
Gutachterin:	Prof. Dr. Nicole Branger
	Westfälische Wilhelms-Universität Münster

Datum der Disputation: 10. Januar 2020

Vom Fachbereich Mathematik der Technischen Universität Kaiserslautern zur Verleihung des akademischen Grades Doktor der Naturwissenschaften (Doctor rerum naturalium, Dr. rer. nat.) genehmigte Dissertation.

D386

Abstract

Diversification is one of the main pillars of investment strategies. The prominent $\frac{1}{N}$ portfolio, which puts equal weight on each asset is, apart from its simplicity, a method which is hard to outperform in realistic settings, as many studies have shown. However, depending on the number of considered assets, this method can lead to very large portfolios. On the other hand, optimization methods like the mean-variance portfolio suffer from estimation errors, which often destroy the theoretical benefits. We investigate the performance of the equal weight portfolio when using fewer assets. For this we explore different naive portfolios, from selecting the best Sharpe ratio assets to exploiting knowledge about correlation structures using clustering methods. The clustering techniques separate the possible assets into non-overlapping clusters and the assets within a cluster are ordered by their Sharpe ratio. Then the best asset of each portfolio is chosen to be a member of the new portfolio with equal weights, the cluster portfolio. We show that this portfolio inherits the advantages of the $\frac{1}{N}$ portfolio and can even outperform it empirically. For this we use real data and several simulation models. We prove these findings from a statistical point of view using the framework by DeMiguel, Garlappi, and Uppal (2009). Moreover, we show the superiority regarding the Sharpe ratio in a setting, where in each cluster the assets are comonotonic. In addition, we recommend the consideration of a diversification-risk ratio to evaluate the performance of different portfolios.

Acknowledgements

First and foremost, I sincerely thank my advisor Prof. Dr. Jörn Saß, for his continuous support and guidance. I am very grateful that he offered me the opportunity and freedom to develop and pursue my own research interests. His door was always open and he always took the time for fruitful discussions.

I owe special thanks to Prof. Dr. Nicole Branger for agreeing to act as a referee for this thesis.

I am truly grateful for the opportunity to work and research at the Department of Mathematics at the TU Kaiserslautern. I particularly thank Dr. habil. Christoph Lossen for creating the administrative working environment which was an important contribution to make this thesis possible.

Furthermore, I want to thank the whole Financial Mathematics group for the excellent working atmosphere and their valuable support. Especially I want to thank my colleagues Dr. Elisabeth Leoff, Barbara Ermisch and Dorothee Westphal for their continuous support and friendship.

Special thanks go to my loving husband Bruno who accompanied me through everything, and to my daughters Marie-Luise and Klara-Sophie for their love and sympathy.

Last but not least, I am grateful to my family and friends, in particular to my father who early aroused my interest in mathematics and supported me with valuable comments and discussions during my PhD studies.

Contents

1 Introduction

The success of a portfolio depends strongly on the selection of assets and the choice of their weights in the portfolio. Special attention is required if risky assets are included. It is crucial to get as good an assessment as possible of the magnitude of the risk, that one may not jeopardize the performance of the portfolio. In order to achieve that, it is necessary to specify the weight of each asset. This determines the impact of the individual asset on the portfolio. Investment in risky assets means to deal with randomness and unknowns.

One way to solve the problem is to assume a model including the parameters for the assets and to theoretically find an optimal solution. This was done by Markowitz (1952) who started the by now classical framework of mean-variance optimization. However, when it comes to real settings, we must first estimate the parameters to apply the optimization results. Thus, we suffer from estimation errors. Even in the mean-variance world, where only a few parameters have to be estimated, this is a huge problem, which gets worse with the number of assets. Especially the estimation of the mean is a task which hardly leads to useful results, as was already pointed out by Merton (1980). One way to overcome this is by improving the estimation procedures. So far, this has not led to more satisfactory results regarding the mean.

Another approach is to exclude estimation at all and find other ways to create portfolios. This is mainly inspired by the idea of diversification. It goes back to the ancient rule for diversification, first orded in the Babylonian Talmud in the fourth century in late antiquity. It recommends diversifying wealth, a third in land, a third in merchandise and keeping a third, which were all investment possibilities at that time. Diversification in terms of tradable assets results in allocating equal amounts of money in several branches or more general in each available asset. The latter is called $\frac{1}{N}$ or naive portfolio. And indeed it performs well when it comes to real data settings, as many studies have shown.

DeMiguel, Garlappi, and Uppal (2009) show that a lot of optimization-based portfolios are not consistently better than the naive portfolio, meaning that they do not outperform it for each performance criterion at the same time. Further, they examine, how much data is needed to overcome the losses from estimation errors, so that the classical mean-variance portfolio outperforms the naive strategy. In fact, thousands of data are needed even in a market model with fewer than 100 assets.

Pflug, Pichler, and Wozabal (2012) approach the problem from the opposite side. They start with a model, where everything is known. Then they increase the uncertainty about the model. And as uncertainty rises the optimal portfolio converges to the $\frac{1}{N}$ portfolio.

Even in a continuous-time model the $\frac{1}{N}$ portfolio can be beneficial. In a jump-diffusion framework Alp and Korn (2011) show that the optimal strategy solving the continuous-time mean-variance problem can be interpreted as a $\frac{1}{N}$ strategy.

On the other hand, the naive portfolio does not necessarily lead to good performance regarding the reward for taking risks. Therefore, it is a common approach to mix optimization procedures with naive investment rules. For example, Tu and Zhou (2011) and Branger, Lucivjanska, and Weissensteiner (2019) suggest a combination of mean-variance procedures and the naive method.

However, there is also a different effect which improves the performance of portfolios, namely sparseness. The mean-variance strategy invests in all available assets thus also allowing short-selling. This often leads to extreme positions, which itself can be risky. If we impose a no-shortselling constraint, the portfolios become more stable and we observe that some assets are excluded from the portfolio. Brodie, Daubechies, De Mol, Giannone, and Loris (2009) also suggest portfolios with fewer assets.

Another issue is related to the branching idea. Bjerring, Ross, and Weissensteiner (2017) use hierarchical clustering methods to reduce the asset universe. Moreover, Branger, Lucivjanska, and Weissensteiner (2019) apply grouping of assets to combine naive diversification with mean-variance optimization.

This thesis aims at analyzing improvements of the performance of the equal weight portfolio strategy by using fewer assets. As in the references given above, we also recommend clustering of the assets, but then applying the naive strategy on the set with one representative from each cluster. For this, we use a hierarchical cluster algorithm using the Pearson correlation to seperate the assets into groups. We use a rolling sample approach like in Brodie, Daubechies, De Mol, Giannone, and Loris (2009) to collect several performance criteria on real and simulated data. Stepwise, we compare naive portfolios starting with the pure naive portfolio over a naive portfolio with fewer assets to two versions of a cluster portfolio. As benchmark, we additionally compare the mean-variance portfolio.

An important point is how to measure performance. Therefore, we introduce and analyze several performance measures. Usually, a reward-risk ratio is used to evaluate the performance of investments. The reward, which is mainly the mean of a portfolio, is always hard to measure. Thus, it is important to estimate the risk of a portfolio, and there is a huge discussion on how to measure risk. First, Artzner, Delbaen, Eber, and Heath (1999) came up with desired properties of a risk measure. Then Rachev, Ortobelli, Stoyanov, Fabozzi, and Biglova (2008) collect and further categorize risk measures and their properties.

The resulting good risk measures stay in contrast to the in practice applied performance measures such as the Sharpe ratio, where the standard deviation is used as risk measure. Many of the applied risk measures do not possess the theoretical best properties. Following the considerations by Cont, Deguest, and Scandolo (2010), the problems of estimation errors must be taken into account for risk measures, as well. According to Cont, Deguest, and Scandolo (2010), there is a conflict between the robustness of a risk measure and the subadditivity property. Thus, the effects of both theoretical and statistical properties must be considered together. As it is generally hard to estimate the reward of a portfolio, we recommend substituting the reward measure by a diversification measure. For this, Carmichael, Koumou, and Moran (2015) suggest a general approach to diversification measures, which also includes formerly applied measures.

We analyze the performance measures we use with respect to good risk measures. And we recommend the usage of a diversification-risk ratio, which relates a diversification measure with a risk measure. We can show that in some cases the $\frac{1}{N}$ strategy can be improved, only by taking fewer assets. In terms of Cover (1991), we show that the growth rate of the naive portfolio with fewer assets equals the growth rate of the full naive portfolio. Following the proof by DeMiguel, Garlappi, and Uppal (2009) about the number of data needed to outperform the mean-variance strategy, we can prove a similar result for the naive portfolio with fewer assets. We can show that the cluster portfolio performance ranking usually is between the naive method and the mean-variance optimization. Thus, it is on the second place in cases which favor the mean-variance strategy and in cases, which favor the full naive portfolio. Therefore, we see it as a robust strategy when facing model uncertainty. We use the comonotonicity framework (see e.g. Deelstra, Dhaene, and Vanmaele (2011) and Dhaene, Denuit, Goovaerts, Kaas, and Vyncke (2002)) to explain the results of the data experiments towards the Sharpe ratio and the diversification-risk ratio in an extreme case setting.

This thesis is organized as follows: In Chapter 2 we are going to introduce the portfolios of interest. There we recapitulate the naive portfolio and the mean-variance portfolio. The main part will be the description of the cluster portfolio. Chapter 3 discusses desirable properties of risk and reward measures in terms of theoretical and statistical issues. Then, a general framework for diversification measures is introduced. The last section introduces and analyzes the considered performance measures including the new diversification-risk ratio.

Chapter 4 contains the first approach to take fewer than all assets in a naive portfolio. Therefore, we take an excursion to the Cover (1991) framework. In Theorem 4.3 we give results on statistical robustness regarding estimation errors in the framework of DeMiguel, Garlappi, and Uppal (2009) examining the performance of the naive portfolio with fewer assets. We include a short comparison of three mean-variance strategies. In Chapter 5 we examine the performance of two cluster portfolios, one with a fixed number of clusters and one with free detection of clusters. Then, we introduce a setting with comonotonic behavior in the single clusters. In that setting we are able to explain by theoretical results some properties of the cluster portfolio we observe on real data. A conclusion can be found in Chapter 6 together with an outlook on further research possibilities.

Appendix A presents the real data set we use for the comparison of the portfolio methods. In Appendix B we explain the simulation models we work with. We shortly introduce the Higham (2002) algorithm for the nearest correlation matrix, which is used for the simulations to transform a pseudo-correlation matrix with the desired correlation coefficients to a true correlation matrix being also positive definite in Appendix C. Finally, Appendix D shows some more detailed diagrams corresponding to Chapters 4 and 5.

2 Portfolio Strategies

In this chapter we introduce the portfolios we consider throughout the thesis. As it is the aim of our work to improve the performance of the naive portfolio strategy, which puts equal weight on each asset, we start with its description in Section 2.1. The classical framework about portfolio optimization was popularized by Harry Markowitz using the mean vector and the covariance matrix of the assets as the only parameters for portfolio selection. We show several versions of the mean-variance optimization portfolios in Section 2.2. Moreover, we use them as benchmark portfolios because they are often referred to in the literature and because they are applied by practitioners. The main part of this chapter is Section 2.3, where we introduce the cluster portfolio, which we use to explore if and how the naive strategy can be improved.

2.1 Equal Weight (Naive) Portfolio

The first evidence of naive diversification can be found in the Babylonian Talmud, tractate Baba Mezi'a, folio 42a from the fourth century. There, Rabbi Isaac bar Aha recommends to divide the wealth into three equal parts: "One should always divide his wealth into three parts: [investing] a third in land, a third in merchandise, and [keeping] a third ready to hand." This general rule of investment is now part of our common sense. Thus, it is not only influencing investment strategies but also all considerations when facing risks.

So this portfolio allocates the equal amount of money in each available asset. That is, if we are given N different assets we invest the fixed amount of money c in every asset. Usually one works with (relative) weight vectors $w \in \mathbb{R}^N$, where each component is the fraction of wealth assigned to an asset. In this case this is just $w_i = \frac{c}{Nc} = \frac{1}{N}$ for all $i = 1, 2, \ldots N$. This is the reason why it is often also called $\frac{1}{N}$ portfolio.

Although it is an easy method and often regarded as naive in the sense of not well thought about, the $\frac{1}{N}$ portfolio possesses many advantages. In the first place it diversifies the wealth across many investment opportunities and thus diversifies away risk by balancing gains and losses.

Other portfolio methods try to optimize the investment regarding reward and risk. However, this leads to many pitfalls, which can be avoided by the naive method. The most important problem arises from estimation errors, which only can be overcome by large data sets. Then, there is the general problem of forecasting using historical returns as there is no guarantee that the assets behave the same way in the future. In particular, large data sets may not be sufficient for good estimates, since model parameters change over time. In addition, there is model uncertainty. Moreover, there is the problem of defining and measuring the risk appropriately. Usually the variance is used as measure of risk, but only the downward deviation is the dangerous part of risk we are facing. When applying the naive methods in models, which are driven by certain parameters, e.g. mean and covariance, and the length of estimation periods is no longer a problem, then other optimizing strategies easily outperform the naive portfolio. However, turning to real data in realistic settings, it is hard to outperform the naive portfolio consistently, i.e. in each performance criterion, see e.g. DeMiguel, Garlappi, and Uppal (2009).

Still, there are drawbacks with this method. If we simply invest in all available assets globally, we are facing huge portfolios with only small fractions of wealth invested in each asset. It would be expensive in time and money to invest and rebalance the portfolio from time to time. The portfolio turnover would always be quite high as the gains of some very good performing assets must be distributed over all bad performing assets. Certainly investing in all available assets is also an unrealistic setting. Usually, the portfolio contains only assets from one or two areas, such as Asia, USA, Europe or parts of it. As there are still too many assets available, the market is usually divided into branches. This is done merely from the economic point of view. Companies working in the same branch are in general dependent on the same resources, environments and consumers or clients. If something is affecting their common basis, they all will either suffer or profit from it in a similar way. Thus, one makes sure that the assets in a portfolio are from different branches. However, by correlation analysis, we see that there are dependencies between assets across branch borders. Another problem is that this procedure leads to low volatility but is usually bad in returns.

In the following we will not only see the naive portfolio on all assets, but also compare the naive portfolio with half the number of assets, where the assets can e.g. be chosen by a performance criterion like the Sharpe ratio.

2.2 Mean-Variance Optimization

Harry Markowitz proposed and derived an optimized portfolio in the mean-variance model (see Markowitz (1952)). In this model only the static parameters, the means, the variances and covariances of the assets explain the market. Since then, many variants of the theoretical optimal solution were created to improve the method for applications in real-world markets.

We consider the constrained, i.e. shortselling is not allowed, and unconstrained sample-based mean-variance optimization as well as the sample-based minimum-variance portfolio. 'Sample-based' means, that instead of the true parameters, which are only known in a model setting, the estimates of the moments are used. This is also referred to as plug-in strategy. In contrast to the unconstrained version the constrained portfolio strategy always leads to reasonable solutions, as portfolio weights do not become negative and thus also extreme long positions are avoided. It also leads to sparse portfolios. Moreover, in contrast to the minimum-variance portfolio, it consistently performs better, as the mean of assets is not ignored.

In Section 4.5 we provide a comparison of these three mean-variance strategies. To this end, we mainly use the constrained mean-variance strategy as benchmark. However, all methods are only optimal in a mean-variance setting where the true parameters are known. If we are facing estimation errors, then the results are no longer optimal. The challenge is to estimate the mean, which is known to need very long time series. Even the variance is hard to estimate and also needs a lot of data. A rule of thumb says: The volatility of a stock can be estimated within 30 days, for the drift one needs more than 90 years.

There is a discussion on how much data is needed for the mean-variance portfolio to outperform the naive portfolio in DeMiguel, Garlappi, and Uppal (2009), see also Section 4.3 for an adaptation to naive portfolios with fewer assets. Apart from estimation errors these methods do not adjust for parameter changes or switching in the underlying data. And last but not least, historical data can only be a hint to the future evolution rather than a reliable forecast.

2.3 Cluster Portfolio

The purpose of our cluster portfolio is to generate a naive portfolio with fewer assets. Therefore, all available assets are grouped into clusters, where one representative from each cluster is chosen for the portfolio. This method consists of two main steps, namely the clustering (Section 2.3.1) and the selection of assets (Section 2.3.3). The clustering is always performed by a hierarchical agglomerative clustering algorithm with complete linkage. The distance function (Section 2.3.2), however, could be chosen in different ways. Also the selection method could be changed, but we use selection of the representatives by the best Sharpe ratio. The following subsections will give a detailed description of the cluster portfolio strategy.

2.3.1 Clustering Method

The hierarchical agglomerative cluster algorithm is a bottom-up procedure which merges clusters from the bottom, where only one asset is in one cluster, upwards until all assets are in one cluster. Therefore, a pairwise distance function $\text{dist}(o_i, o_j)$ for all objects $o_i, i \in I$, is used to define the distance of objects, where I is the index set of all objects. Moreover, we need a distance function for clusters based on the object distance function. The advantage of this procedure is that it is possible to choose the number of clusters or specify a critical distance between the clusters. Then the distance between clusters is at least at this level.

The algorithm produces a hierarchical tree which can be visualized in a dendrogram, see Figure 2.1. The x-axis contains the numbers, which are assigned to the objects, and the distance between the clusters is displayed on the y-axis. The pruning of this tree defines the clusters. If a fixed number of clusters is desired, the horizontal line crossing exactly the specified number of vertical lines produces the desired clusters. The height of the tree gives the distance which is needed to put all assets in one cluster. Thus, the higher the tree the more variety is in the data. If we want that the internal distance in a cluster is not higher than a specified level, then the line going through this level provides the desired clusters. In Figure 2.1 the red line displays the cutting of the tree, this would be either setting 11 clusters or a level of 0.5. We are going to use both methods in the sequel. Usually we will choose the number of clusters to half the number of available assets for the fixed cluster portfolio. Moreover, we found that a level of 35% of the maximal distance in the tree often leads to the desired clusters, which is why we use it for the free cluster portfolio.

A further advantage of this method is that we can observe the clusters by sight and we can change the number of clusters or the cutting level afterwards without having to start the algorithm all over again.

Here we use complete linkage, which is also called furthest neighbor method, to merge the clusters. It uses the maximal distance for comparing two clusters. Thus, only if the maximal distance between the assets, one of each cluster, is near enough two clusters can be merged. The distance of two clusters C and D using complete linkage is then:

$$\text{dist}(C, D) = \max\{\text{dist}(o_s, o_t) \mid s \in I_C,\ t \in I_D\},$$

where I_C and I_D are the index sets of the clusters C and D, respectively. This ensures that the internal distance of a cluster is not higher than the maximum distance of the last merging step.

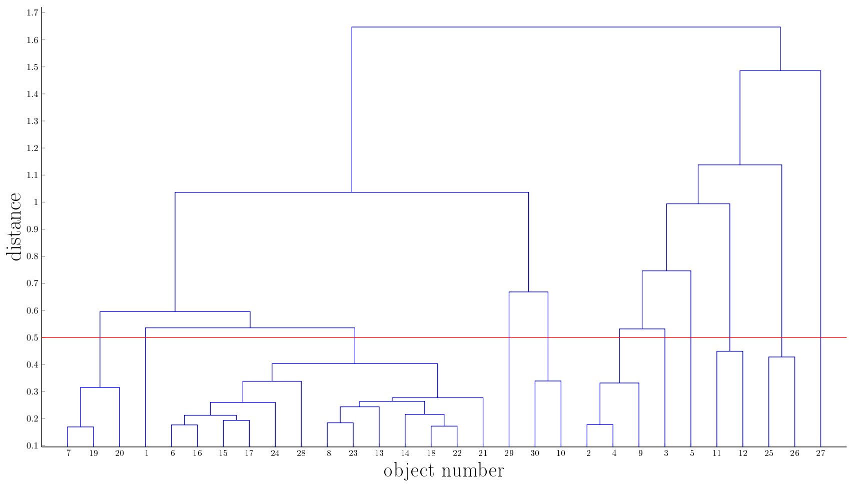


Figure 2.1: Dendrogram of a hierarchical agglomerative cluster tree.

For a general introduction to clustering algorithms see for example *'Clustering methods'* by Rokach and Maimon (2005) or *'The Elements of Statistical Learning'* by Hastie, Tibshirani, and Friedman (2017). Next, we need to specify the distance function for two objects.

2.3.2 Distance Functions

In general, we measure distances with a metric, which is defined as follows:

Definition 2.1. *A metric on a set X is a function* $\text{dist} : X \times X \to [0, \infty)$*, such that for all $x, y, z \in X$:*

1. $\text{dist}(x, y) = 0 \iff x = y$ *(identity of indiscernibles),*
2. $\text{dist}(x, y) = \text{dist}(y, x)$ *(symmetry),*
3. $\text{dist}(x, z) \leq \text{dist}(x, y) + \text{dist}(y, z)$ *(triangle inequality).*

For clustering a distance function which does not satisfy all properties of a metric still leads to reasonable results. Quite often the triangle inequality is violated. Although this property has computational advantages, the clustering works even if it is not fulfilled. If the attribute, which we want to use to compare objects, is a measure of similarity like the correlation, we need to transfer it into a dissimilarity measure with a suitable monotone decreasing function.

Pearson Correlation

The Pearson correlation is the standard linear correlation coefficient $\rho \in [-1; 1]$ for square integrable random variables X and Y:

$$\rho = \text{corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}[X]\text{Var}[Y]}}.$$

We use the sample version for paired sample points (x_i, y_i) for $i = 1, \ldots, M$, where the covariance, the variance and the expectation are replaced by their estimates:

$$\hat{\rho} = \frac{\sum_{i=1}^{M}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{M}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{M}(y_i - \bar{y})^2}},$$

where $\bar{x} = \frac{1}{M}\sum_{i=1}^{M} x_i$ denotes the sample mean of x and analogously $\bar{y}$ the sample mean of y. To convert this similarity measure into a dissimilarity measure, we chose the function $f_1(x) = 1-x$. In the literature also different functions are used, e.g. $f_2(x) = 1 - x^2$ or $f_3(x) = \sqrt{2(1-x)}$. We prefer the first one, because it assigns to each correlation coefficient a different value and changes in correlation have the same impact over the whole range. In Figure 2.2 we easily see, that these properties only hold for the first function.

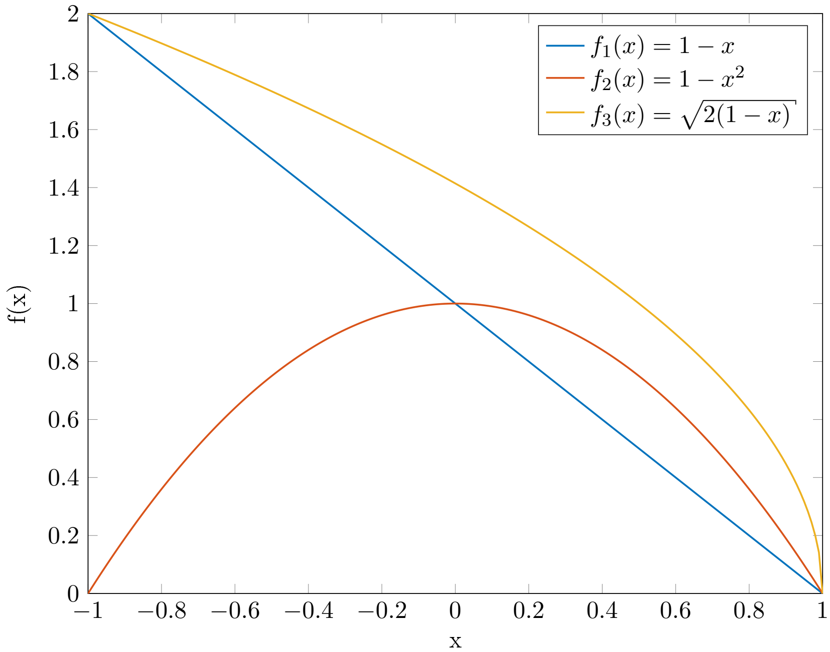


Figure 2.2: Distance Functions

Then, the following formula defines the distance function with the Pearson correlation coefficient we are going to use in this thesis:

$$\text{dist}_{Pearson}(X, Y) = 1 - \text{corr}(X, Y) \tag{2.1}$$

for square integrable X, Y. This function takes values in $[0; 2]$, where 0 means perfect positive correlation, 2 stands for perfect negative correlation and 1 indicates uncorrelated X and Y. As we do not know the true correlation between the assets, it has to be estimated and unfortunately, the sample correlation is sensitive to noise and outliers. The triangle inequality for this distance function is violated. As we could not find a good reference, we include the proof for completeness.

Proposition 2.2. *The Pearson correlation distance function defined in* (2.1) *does not satisfy all properties of a metric in* L^2. *It satisfies the symmetry property but violates the identity of indiscernibles and triangle inequality.*

Proof. We need to check the properties given in Definition 2.1: The Pearson correlation coefficient ranges from -1 to 1 thus the distance function ranges from 0 to 2: $\rho \in [-1; 1] \to 1-\rho \in [0; 2]$. Thus $\text{dist}(x, y) \geq 0$ for $X, Y \in L^2$.

1. (identity of indiscernibles may not hold)
   If $X = Y$ then $\text{dist}_{Pearson}(X, Y) = 0$. However, the correlation of two linear dependent

random variables is also zero. Therefore $\text{dist}_{Pearson}(X,Y) = 0$ does not imply the equality of the random variables.

2. (symmetry)
The symmetry property of the Pearson correlation passes on to the Pearson distance function: $\text{dist}_{Pearson}(X,Y) = 1 - \text{corr}(X,Y) = 1 - \text{corr}(Y,X) = \text{dist}_{Pearson}(Y,X)$.

3. (triangle inequality may not hold)
We can show the violation of the triangle inequality with the following counterexample: Let X, Y, Z be random variables, where X and Z are independent but they have the same variance ($\text{Var}(X) = \text{Var}(Z)$) and $Y = X + Z$. Then the triangle inequality for the Pearson correlation distance,

$$1 - \text{corr}(X,Z) \leq 1 - \text{corr}(X,Y) + 1 - \text{corr}(Y,Z),$$

would be equivalent to

$$\text{corr}(X,Y) + \text{corr}(Y,Z) \leq 1 + \text{corr}(X,Z).$$

As X and Z are independent, $\text{corr}(X,Z) = \text{Cov}(X,Z) = 0$. In addition we have $Y = X+Z$, thus the triangle inequality in this example is equivalent to

$$\text{corr}(X, X+Z) + \text{corr}(X+Z, Z) \leq 1.$$

We now show that the left hand side is greater one (leading to a contradiction) using the general properties of the variance and the covariance together with $\text{Var}(X) = \text{Var}(Z)$ and $\text{Var}(X+Z) = \text{Var}(X) + \text{Var}(Z) = 2\text{Var}(X)$.

$$\begin{aligned}
&\text{corr}(X, X+Z) + \text{corr}(X+Z, Z) \\
&= \frac{\text{Cov}(X, X+Z)}{\sqrt{\text{Var}(X)\text{Var}(X+Z)}} + \frac{\text{Cov}(X+Z, Z)}{\sqrt{\text{Var}(X+Z)\text{Var}(Z)}} \\
&= \frac{\text{Var}(X) + \text{Cov}(X,Z)}{\sqrt{2\ \text{Var}(X)^2}} \\
&\quad + \frac{\text{Var}(Z) + \text{Cov}(X,Z)}{\sqrt{2\ \text{Var}(X)^2}} \\
&= \frac{\text{Var}(X)}{\sqrt{2}\ \text{Var}(X)} + \frac{\text{Var}(X)}{\sqrt{2}\ \text{Var}(X)} \\
&= \frac{2}{\sqrt{2}} > 1.
\end{aligned}$$

□

To check how good the cluster method works with the Pearson correlation, we examine the performance on different underlying correlation matrices. In Figure 2.3, we show examples for a setting with 30 assets. The correlation coefficients of assets within a group are always around 0.9 and the external correlation coefficients are chosen randomly from the intervals $[-0.3; 0.3]$, $[0; 0.3]$ and $[0.3; 0.6]$. In Figure 2.3, we see the graphs of the estimation error for the correlation matrix in the left column in terms of the 2-norm, i.e. $\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|$ for the difference A of estimated and true correlation matrix. The plots in the right column show the range for the cutting level to produce the right clusters, where the blue line is the minimum level and the red line the maximum level. The free cut of 35% of the maximal distance of the whole tree is given as yellow line. We see that the estimation error decreases very fast in the first two cases, where the difference between the internal and external correlation is high. In the third cases 100 more data points are needed for the estimation error to fall below 2.5 compared to the first cases. The range for the cut to produce the correct cluster becomes narrower, when the difference between internal and external correlation decreases. The free cutting level is nearly always in the range of optimal cut. Only in the first case it goes below the optimal range between 45 and 63 data points for the estimation period. One important observation is, that although the correlation matrix is quite well estimated, the resulting cluster tree is never exactly the same. Therefore, the estimation errors indeed have an impact on the clustering.

Spearman Correlation

The Spearman's rank correlation or simply Spearman's rho is a copula-based method to express dependence. It is a statistically robust alternative to the Pearson correlation, which includes other than only linear dependence structures. It can be used for clustering in the same way as the Pearson correlation. This procedure shows better results, when the underlying data is heavy-tailed, see e.g. Marti, Andler, Nielsen, and Donnat (2016). For the data examined in this thesis the difference in results were only marginal, therefore we focused on Pearson correlation only.

2.3.3 Selection

In terms of diversification, it is not important, which asset we choose from each cluster. Nevertheless, regarding performance, we should make a good choice. The easiest way to achieve that, is to choose the asset with best historical Sharpe ratio from each cluster, which we will do in this thesis. If the estimation error is low and there are only moderate changes in the parameters from the past to the future, this will allow an improvement of the Sharpe ratio for the whole portfolio. Further selection methods could be the following: If we are interested in other performance measures, we might use them as selection criteria. Apart from that, if we want to achieve the smallest possible turnover, we might stick to an asset which was present in the former portfolio although its performance now is not the best under its fellow assets in the cluster. An alternative way is to find an optimal portfolio under all possible combinations. However, this needs high computational effort and cannot be performed in every setting. As the mean-variance optimization without shortselling also leads to good performances one could optimize over the best representatives. In case of a small amount of clusters this is not an appropriate procedure because we lose too many assets to maintain a well diversified portfolio. However, for a very large number of clusters this could be a profitable strategy.

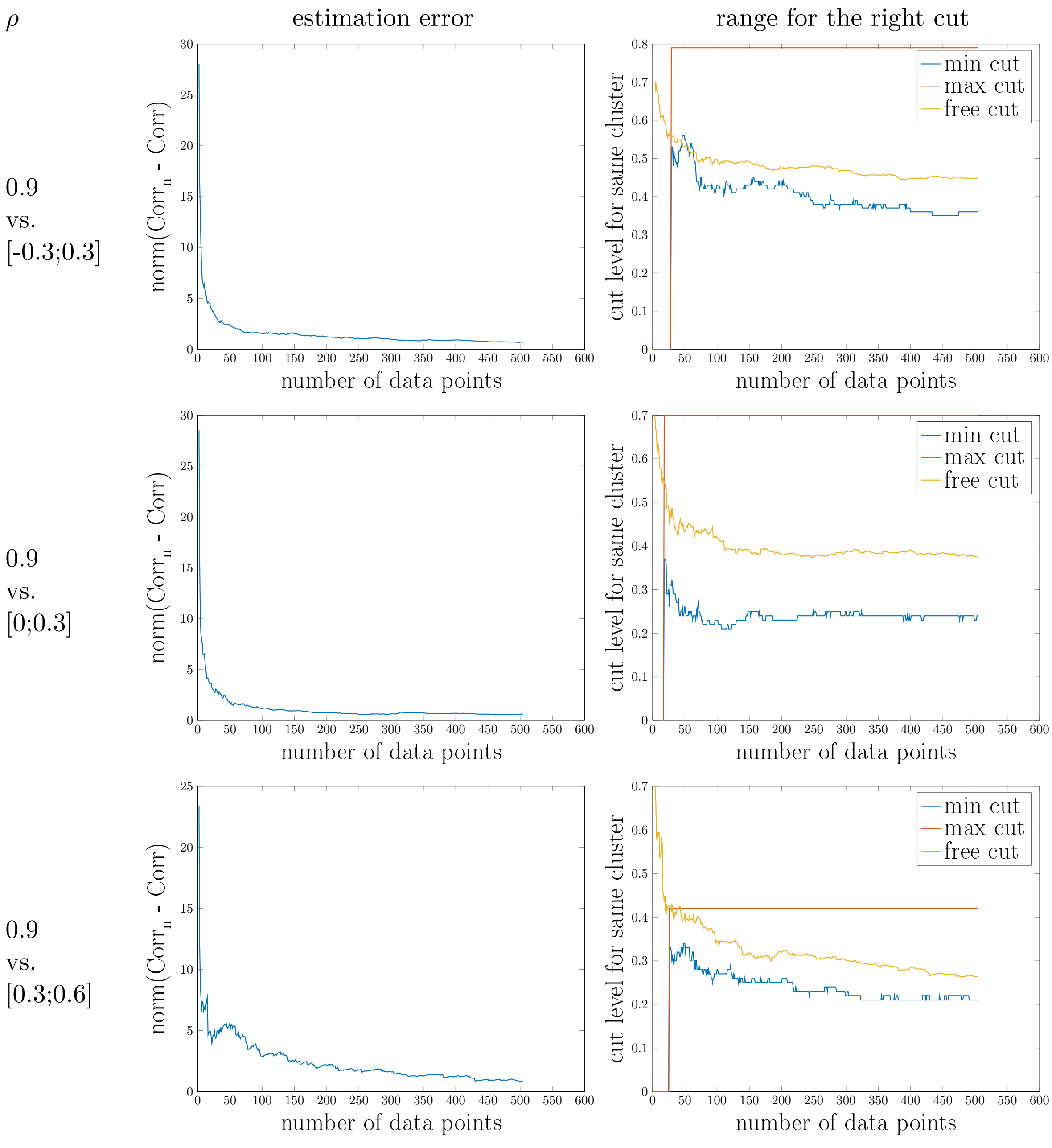


Figure 2.3: Cluster robustness test: Estimation error for the correlation matrix and ranges for the optimal cut producing the correct clusters including the level for the free cut using daily simulated data for 30 assets and correlation matrices with internal correlation of 0.9 and external correlation from ranges $[-0.3; 0.3]$, $[0; 0.3]$ and $[0.3; 0.6]$.

3 Performance

To evaluate the performance of a portfolio, usually a reward-risk ratio is used. The reward part is difficult to handle since the estimation of the mean is often not reliable. Thus it is necessary to focus on the risk for an optimal portfolio. The cluster portfolio realizes this by grouping assets according to their correlation. We need to check to which extent this carries over to the out-of-sample results. For this purpose, we must find a good way to measure risk. Unfortunately, there is no best risk measure to complete the task.

There are a lot of measures that are commonly used, but each of them has drawbacks. Therefore, it is important to know about the strengths and weaknesses of each measure. We need to find the appropriate ones under the particular consideration of a good risk measure, to evaluate the performance of a portfolio. In addition, there is always the question: What actually is a good risk measure? Another question is, how to define risk and how to measure it? And finally, can we estimate the risk properly?

In the first part of this chapter (Section 3.1) we will recapitulate the properties of good risk measures, starting with the definitions by Artzner, Delbaen, Eber, and Heath (1999). Then we will present the differences between risk measures, uncertainty measures and deviation measures as defined in Rachev, Ortobelli, Stoyanov, Fabozzi, and Biglova (2008). This terminology is extended by the definition of a reward measure as proposed in Rachev, Ortobelli, Stoyanov, Fabozzi, and Biglova (2008). In the second part (Section 3.2) we will introduce statistical robustness of risk measures following the paper of Cont, Deguest, and Scandolo (2010). The third part (Section 3.3) provides the diversification measure defined in a general framework by Carmichael, Koumou, and Moran (2015). The last part (Section 3.4) deals with the collection of performance measures, which are used in this thesis to compare the portfolios from Chapter 2. We explain how we record the performance for a given data set and describe the performance measures we choose. We also introduce a new performance ratio, which relates risk to diversification.

3.1 Good Risk Measures

As there is no general definition of risk in the area of portfolio optimization, it is naturally hard to define a good risk measure. Every existing risk measure is limited in its abilities. Therefore, it is necessary to define desirable properties to get appropriate risk measures.

A systematic attempt to do this was the definition of a coherent risk measure by Artzner, Delbaen, Eber, and Heath (1999). They use the future wealth approach, which just considers the future values of an asset or portfolio. Then they divide the future outcomes into two sets, one including outcomes of acceptable risk and the other containing the outcomes of unacceptable risk. That is, if we invest in a portfolio at time $t = 0$ the future outcome at time $t = T$ falls either in the set, where we accept the risk of the portfolio, or not. If the outcome does not fall into the acceptance region, one can calculate the amount of money, which invested at time $t = 0$ in a riskless asset and added to the portfolio, would be sufficient to bring the outcome into the region of acceptance. Then this amount of money can be seen as a measure of risk. The more money we need to invest in addition, the riskier the portfolio will be (bad risk). However, if this amount is even negative we are facing a portfolio with acceptable (good) risk. Therefore, Artzner, Delbaen, Eber, and Heath (1999) define risk as the future net value $\sum_{1 \leq i \leq I} e_i \cdot A_i(T)$ of a portfolio. In this setting $A_i(T)$ is the value of asset i at the end of the holding period at time $t = T$, discounted by the factor e_i (in the paper they consider various currencies in which assets are traded: "e_i denotes the random number of units of currency 1 which one unit currency i buys at date T") to the time of investment at time $t = 0$. The acceptance sets will then lead to a risk measure, which states how far a future net value is from the set of acceptable risks. Therefore, a risk measure is just a mapping from the set of all possible random outcomes $\mathcal{G}$ into $\mathbb{R}$. To formally create this correspondence between the acceptance sets and the risk measures they provide two definitions:

Definition 3.1. (Risk measure defined by an acceptance set)
Given the total rate of return r on a reference instrument, the risk measure defined by an acceptance set $\mathcal{A}$ is the mapping from $\mathcal{G}$ to $\mathbb{R}$ denoted by $m_{\mathcal{A},r}$ and defined by $m_{\mathcal{A},r}(X) = \inf\{c \mid c \cdot r + X \in \mathcal{A}\}$.

Definition 3.2. (Acceptance set associated with a risk measure)
The acceptance set associated with a risk measure m is the set denoted by $\mathcal{A}_m$ and defined by $\mathcal{A}_m = \{X \in \mathcal{G} \mid m(X) \leq 0\}$.

After that, they define some properties of risk measures that arise from a monetary interpretation and from investors' preferences.

Axiom 3.3. Translation invariance.
For all $X \in \mathcal{G}$ and all real numbers α, we have $m(X + \alpha \cdot r) = m(X) - \alpha$.
Adding a riskless investment to the portfolio should lower the risk of the new portfolio in comparison with the old one.

Axiom 3.4. Subadditivity.
For all X_1 and $X_2 \in \mathcal{G}$, $m(X_1 + X_2) \leq m(X_1) + m(X_2)$.
If the investor adds another asset to his portfolio the new risk should not be higher than the sum of the individual risks.

Axiom 3.5. Positive homogeneity.
For all $\lambda \geq 0$ and all $X \in \mathcal{A}$, $m(\lambda X) = \lambda\, m(X)$.
The risk for an investment should be independent of the amount of the investment.

Axiom 3.6. Monotonicity.
For all X and $Y \in \mathcal{G}$ with $X \leq Y$, we have $m(Y) \leq m(X)$.
A higher future net value is associated with a smaller risk.

Then they propose that every good risk measure should have the properties above. This leads to the famous class of coherent risk measures.

Definition 3.7. (Coherence)
A risk measure satisfying the four axioms of translation invariance, subadditivity, positive homogeneity, and monotonicity is called coherent.

A prominent example of coherent risk measures is the 'average value at risk' (AV@R), which is also called 'conditional average value at risk' or 'expected shortfall'. Also the spectral risk measures which are a generalization of the AV@R are coherent. Another important class of axiomatic risk measures was introduced by Föllmer and Schied (2002), the so-called convex risk measures. Many utility-based risk measures only satisfy the properties of this class.

Definition 3.8. (Convex Risk Measure)
A risk measure that satisfies the axioms of translation invariance and monotonicity together with the following convexity property

$$m(\lambda X + (1-\lambda)Y \leq \lambda m(X) + (1-\lambda)m(Y)), \text{ for all } X, Y \text{ and } 0 \leq \lambda \leq 1,$$

is called a convex risk measure*.*

This axiomatic approach can even be extended. Rachev, Ortobelli, Stoyanov, Fabozzi, and Biglova (2008) present an overview of several desirable properties and enlighten the difficulties of finding an ideal risk measure. For that, they compare the concepts of uncertainty measure, risk measure and risk metric and show their abilities to evaluate risk in a portfolio optimization context. The standard deviation is often used to describe the risk of an investment. It measures not only the downside risk but also the deviation above the mean. In a normally distributed world this would not be a problem, but when there are more heavy tailed distributions in question, the standard deviation is just a measure of uncertainty.

Definition 3.9. (Uncertainty Measure)
Any increasing function of a positive functional d defined on the space of random variables is called an uncertainty measure *if it satisfies the following properties:*

1. $d(X+\alpha) \leq d(X)$ *for all* X *and constants* $\alpha \geq 0$.
2. $d(0) = 0$, *and* $d(\lambda X) = \lambda d(X)$ *for all* X *and* $\lambda > 0$.
3. $d(X) \geq 0$ *for all* X*, with* $d(X) > 0$ *for non-constant* X.

In fact the standard deviation belongs to a subclass of uncertainty measures, namely the deviation measures:

Definition 3.10. (Deviation Measure)
An uncertainty measure that satisfies the first property as equality ($d(X+\alpha) = d(X)$ for all X and constants $\alpha \geq 0$) and the following triangle inequality

4. $d(X+Y) \leq d(X) + d(Y)$ *for all* X *and* Y*,*

is called a deviation measure*.*

It is important that, although the standard deviation is a deviation measure, this class does not exclude asymmetric representatives as $d(X) = d(-X)$ is not a consequence of the definition. If the return distribution functions are only defined by the mean and a risk measure, then most of the representatives of the convex and the deviation measure are equivalent. However, in the case of portfolio optimization, both classes cannot be interchanged. The risk measure definitions above try to relate the risk to a stochastic order which can express the investors preference as in maximizing expected utility. As it is also important to take the reward of an investment into account and not only the risk, Rachev, Ortobelli, Stoyanov, Fabozzi, and Biglova (2008) formulate suitable conditions for a reward measure.

Definition 3.11. *(*Coherent Reward Measure*)*
A functional v is isotonic with respect the order of preference of the market if

$$X \geq Y \Longrightarrow v(X) \geq v(Y). \tag{3.1}$$

It is called a coherent reward measure if it additionally satisfies the following conditions:

1. $v(X + \alpha) = v(X) + \alpha$, *for all X and constants α,*
2. $v(0) = 0$, *and* $v(\lambda X) = \lambda v(X)$, *for all X and $\lambda > 0$,*
3. $v(X + Y) \geq v(X) + v(Y)$, *for all X and Y.*

In Section 3.4 we will show how the performance measures we use can be categorized in this context.

3.2 Robust Risk Measures

In the previous section many desirable aspects of risk measures were developed. From a theoretical point of view this would be enough to find a suitable risk measure. However, in practice it is not possible to obtain the distribution for the returns of an asset exactly. One has to estimate the parameters of the return distribution as good as possible. This opens a new field of considerations. All statistical procedures must be analyzed for their abilities to determine a good estimate. This includes robustness and sensitivity analysis. The first approach to these problems regarding risk measures was provided by Cont, Deguest, and Scandolo (2010). They define a two step procedure of first estimating the distribution of the profit and loss of a portfolio and second applying the (distribution based) risk measure and call it the risk measurement procedure. They observe a general problem concerning the subadditivity property, which contradicts statistical robustness.

This procedure leads to an effective risk measure, which is closely related to the empirical distribution function and therefore represents the measure that is actually computed. However, it is not generally true that this effective measure has the same properties as the originally chosen risk measure.

For example they show,

$$m_p(X) = -\int_0^1 q_u^-(X) p(\mathrm{d}u), \tag{3.2}$$

where p is a probability measure on $(0, 1)$, and $q_u^-(X) = q_u^-(F_X) \triangleq \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}$ is the lower quantile of $F \in \mathcal{G}$ of order $\alpha \in (0, 1)$, then m_p has a certain robustness property (C-robustness). For details see Cont, Deguest, and Scandolo (2010). This shows that, at least for the risk part of the reward-risk measures, robustness could be an additional feature one should take into account.

3.3 Diversification Measures

In the literature, there is a big discussion on how to measure diversification in a portfolio. This discussion is also about how to define diversification at all. Several attempts to do that range from the mere number of assets in a portfolio, preferring more to less, over variance analysis up to higher moments analysis. Carmichael, Koumou, and Moran (2015) set up a general framework for diversification measures, using Rao's quadratic entropy (RQE). By translating it to the portfolio setting they derive a formula which includes several existing diversification measures. By the notion of entropy, which is a measure of chaos or nonconformity, they can examine how different the ingredients of a portfolio are. They use the discrete version of the RQE, which is the average distance between two randomly drawn individuals X_i, X_j from a set of individuals Ω:

$$H(P) = \sum_{i,j}^{N} d_{i,j} p_i p_j,$$

where P is the probability distribution function of X, $p_i = P(X = x_i)$, for $i = 1, \dots, N$, and $d_{i,j} = d(X_i, X_j)$ is a non-negative, symmetric dissimilarity function, see below.
In terms of portfolio selection, Carmichael, Koumou, and Moran (2015) transfer the RQE to the new setting: Take a family (universe) $U = (A_i)_{i=1}^N$ of N different assets, and denote a specific portfolio (without shortselling) by $w = (w_i)_{i=1}^N$, where w_i is the relative weight of asset i in w. Thus each w can be seen as a population of individuals, the former Ω. Then the random variable X can take values $1, \dots, N$, the number of assets in portfolio w with the probability distribution $P(X = i) = w_i$, the probability that asset i is in portfolio w. This is associated to a random experiment, where the assets are chosen randomly with replacement from portfolio w. The dissimilarity function is then a matrix $D = (d_{i,j})_{i,j=1}^N$, where

$$\begin{aligned} &d_{i,j} \geq 0, \quad \text{for } i, j = 1, \dots, N, \\ &d_{i,j} = d_{j,i}, \quad \text{for } i, j = 1, \dots, N \text{ and} \\ &d_{i,i} = 0, \quad \text{for } i = 1, \dots, N. \end{aligned}$$

Thus, the REQ of a portfolio is defined as half of the average difference between two randomly drawn (with replacement) assets:

$$H_D(w) = \frac{1}{2} \sum_{i,j=1}^{N} d_{i,j} w_i w_j.$$

This can be interpreted as the average pairwise dissimilarity between assets in a portfolio w and the impact on the portfolio by the product of the relative weights. Further properties are discussed in Rao and Nayak (1985). In terms of $H_D(w)$ being the average degree of information concentration, the dissimilarity measures the amount of unshared information. If the assets in a portfolio are completely similar, i.e. $d_{i,j} = 0$, then the dissimilarity equals zero and there is no diversification effect.

This framework includes the Gini-Simpson Index with $d_{i,j} = d(1 - \delta_{i,j})$, where $\delta_{i,j}$ is the Kronecker's delta and d a strictly positive constant. And it equals the diversification return or excess growth rate with $d_{i,j} = \sigma_i^2 + \sigma_j^2 - 2\sigma_{i,j}$, where σ_i^2 is the variance of asset i and $\sigma_{i,j}$ the covariance between asset i and j. Assuming that the variances of all assets are equal and normalized to 1, this leads to an RQE which uses correlation $\rho_{i,j}$ as an indicator for diversification:

$$H_D(w) = \sum_{i,j=1}^{N} (1 - \rho_{i,j}) w_i w_j. \tag{3.3}$$

As we are examining portfolios under correlation aspects, this is the right measure to compare the portfolios in our settings. If we want to optimize the portfolio in terms of this diversification measure and we assume that the correlation between all assets is the same, then the optimal weights are the weights of the naive portfolio.

Proposition 3.12. *Let $0 \leq \rho < 1$ be the correlation coefficient for all asset pairs (X_i, X_j) for $i \neq j$ for $i, j = 1, \ldots, N$. Then the optimal no shortselling portfolio w^* maximizing diversification with respect to 3.3 is the naive portfolio.*

Proof. We have to solve the following nonlinear optimization problem:

$$\max_w \sum_{i,j=1}^{N} (1 - \rho_{i,j}) w_i w_j$$

$$s.t. \sum_{i=1}^{N} w_i = 1, \tag{3.4}$$

$$0 \leq w_i \leq 1, \text{ for } i = 1, \ldots, N. \tag{3.5}$$

Using the Lagrange multiplier approach leads to the Lagrange function:

$$L(w, \lambda) = \sum_{i,j=1}^{N} (1 - \rho_{i,j}) w_i w_j - \lambda(\sum_{i=1}^{N} w_i - 1)$$

Then the first order condition leads to the following system of equations:

$$\begin{cases} \sum_{j=1}^{N} (1 - \rho_{1,j}) w_j - \lambda = 0, \\ \vdots \\ \sum_{j=1}^{N} (1 - \rho_{N,j}) w_j - \lambda = 0, \\ \sum_{i=1}^{N} w_i - 1 = 0. \end{cases}$$

Taking into account that $\rho_{i,i} = 1$, which implies that the term with w_i in equation i of the first N equations vanishes, and $\rho_{i,j} = \rho$ for $i \neq j$ lead to

$$\begin{cases} (1 - \rho)(w_2 + \ldots + w_N) - \lambda = 0, \\ \vdots \\ (1 - \rho)(w_1 + \ldots + w_{N-1}) - \lambda = 0, \\ w_1 + \ldots + w_N - 1 = 0. \end{cases}$$

From the last equation we get that each sum of weights without weight i must be equal to $1 - w_i$:

$$w_1 + \ldots + w_{i-1} + w_{i+1} + \ldots + w_N = 1 - w_i.$$

Inserting this into the first N equation leaves to solve:

$$\begin{cases} (1 - \rho)(1 - w_1) - \lambda = 0, \\ \vdots \\ (1 - \rho)(1 - w_N) - \lambda = 0. \end{cases}$$

Rearranging terms yields:

$$\begin{cases} (1-\rho)(1-w_1) = \lambda, \\ \vdots \\ (1-\rho)(1-w_N) = \lambda. \end{cases}$$

Thus the optimal solution has to fulfill:

$$(1-\rho)(1-w_1) = \cdots = (1-\rho)(1-w_N),$$

for $\rho < 1$. This only holds if all weights are equal. The condition that the weights sum up to 1 in (3.4) leads to

$$1 = \sum_{j=1}^{N} w_j = N w_i,$$

and thus $w_i^* = \frac{1}{N}$ for $i = 1, \ldots, N$, for the optimal weights, which is the naive portfolio. □

3.4 Performance Criteria

To compare the different portfolio strategies we use different performance criteria. Apart from the mean (drift) and the standard deviation (volatility) we also use the Sharpe ratio (Section 3.4.1), the Sortino ratio (Section 3.4.2), the omega ratio (Section 3.4.3), the portfolio turnover (Section 3.4.4), the certainty equivalent return (Section 3.4.5) and the new diversification-risk ratio (Section 3.4.6). The riskless asset or target return is always set to zero, and the risk aversion to one.

For the collection of the performance measures we use a rolling sample approach like Brodie, Daubechies, De Mol, Giannone, and Loris (2009). First, we fix an estimation period between 60 and 360 months to estimate parameters. Given this information, the portfolios are fixed for the following testing period of 12 months. In the testing period, all necessary characteristics of the portfolios are collected. Then, the portfolios are rebalanced according to the information from the preceding new estimation period. Again, the performance of the portfolios is recorded in the following testing period. This is executed up to the last testing period available. All together we then have collected the performance of the portfolios for the whole data set excluding the first estimation period. To make it easier to compare the data, we will always provide an average value over the whole period and the average values of the last three five-year-periods, which is the only out-of-sample data used in all settings. In the following we always have 45 years of daily and monthly data available.

3.4.1 Sharpe Ratio

The most commonly used measure is the Sharpe ratio, introduced by William Sharpe (1994). It is a relative risk measures, which indicates that the outcome does not give an absolute specification of the portfolio. Only by comparing portfolios we can state that the higher the Sharpe ratio the better the performance of the portfolio will be. The ratio only works well, if we believe that the risk in a portfolio can be expressed by the standard deviation. However, this can be questioned. It is true that the so-called downside risk is what investors fear. This is the risk that the actual return of the portfolio will fall below the mean. The possibility that the true return will be above the mean is exactly what investors are waiting for. Nevertheless, the ratio combines two

important features of a portfolio and can also be adapted to a riskless investment opportunity. This makes the comparison of different portfolios very easy, especially since the calculation is pretty easy, as well.
There are two versions of the Sharpe ratio, which can be seen as the risk adjusted return. The ex ante Sharpe ratio is the expected return per unit of expected risk:

$$SharpeRatio(P) = \frac{\mathrm{E}[R - t_r]}{\sqrt{\mathrm{Var}(R - t_r)}}.$$

In addition, there is the ex post or historic Sharpe ratio, which uses historical data and is applied to approximate the future Sharpe ratio:

$$SharpeRatio(P) = \frac{\frac{1}{T}\sum_{i=1}^{T}(R_i - t_r)}{\sqrt{\widehat{\mathrm{Var}}(R_i - t_r)}}.$$

where P is the considered portfolio with return R in the ex ante case or returns R_i for $i = 1, \ldots, T$ in the ex post case, t_r is the return of the riskless asset and $\widehat{\mathrm{Var}}(R_i - t_r) = \frac{1}{T-1}\sum_{i=1}^{T}((R_i - t_r) - \frac{1}{T}\sum_{i=1}^{T}(R_i - t_r))^2$ is the estimator of the variance. These formulas tell us that the Sharpe ratio delivers the risk-adapted return of the portfolio. It specifies the return of the portfolio per unit of risk if not investing in the riskless asset.

We already stated the advantages of the Sharpe ratio, its easy calculation, and that one value suffices to consider two features of the portfolio together with the information about the riskless asset. However, there are several drawbacks. At first, the Sharpe ratio is very dependent on the time period on which it is calculated. Periods with high performance lead to high ratios, but it often happens that these periods are followed by poor performance. Thus we have to be cautious when taking the historic Sharpe ratio as a forecast. More importantly, the inclusion of assets with high positive returns may make the Sharpe ratio smaller in contrast to the portfolio where this asset is excluded. This is due to the rise of the standard deviation of the whole portfolio caused by this well-performing asset. This is due to the problem with up- and downside risk, as pointed out above. In general, portfolios with skew in their return distributions are not properly measured. In the end, the Sharpe ratio does not consider correlation between the assets. However, correlation is crucial for a well diversified portfolio. The Sharpe ratio is only appropriate for the comparison of investments of the same asset class, which assumes equal distribution families, but, if the distributions differ a lot as in Figure 3.3, then it is not possible to properly find a preference with changing threshold (riskless asset return). Thus, the Sharpe ratio cannot really tell the two distributions apart and has no change in preference, see Figure 3.4. Following the definitions from Section 3.1, the Sharpe ratio is a reward-risk ratio, where the risk measure is a deviation measure, see Rachev, Ortobelli, Stoyanov, Fabozzi, and Biglova (2008).

3.4.2 Sortino Ratio

One way to improve the disadvantage of the Sharpe ratio not recognizing the difference between downside risk and upside deviation, is using the Sortino ratio, see e.g. Rollinger and Hoffman (2013). In the ex ante version it can be described as follows:

$$SortinoRatio(P) = \frac{\mathrm{E}[R - t_r]}{\sqrt{\mathrm{E}[(\min\{0; R - t_r\})^2]}}.$$

To calculate the ratio for historical data one uses this formula:

$$SortinoRatio(P) = \frac{\frac{1}{T}\sum_{i=1}^{T}(R_i - t_r)}{\mathrm{TDD}(R)},$$

where $\mathrm{TDD}(R) = \sqrt{\frac{1}{T}\sum_{i=1}^{T}(\min\{0, R_i - t_r\})^2}$ is the target downside deviation over all asset returns replacing the standard deviation in Sharpe's formula. The further interpretation is quite similar to the Sharpe ratio. The change to downside deviation makes the Sortino ratio interpret only the deviation on the left of the desired target return as risk, see Figure 3.1. If there are mass changes in the asset return distributions, see Figure 3.3, this changes the preference with the threshold, see Figure 3.4. There, the red dots mark places where the ratio curves intersect. Again, the higher the ratio the better the portfolio performance will be. As the term target

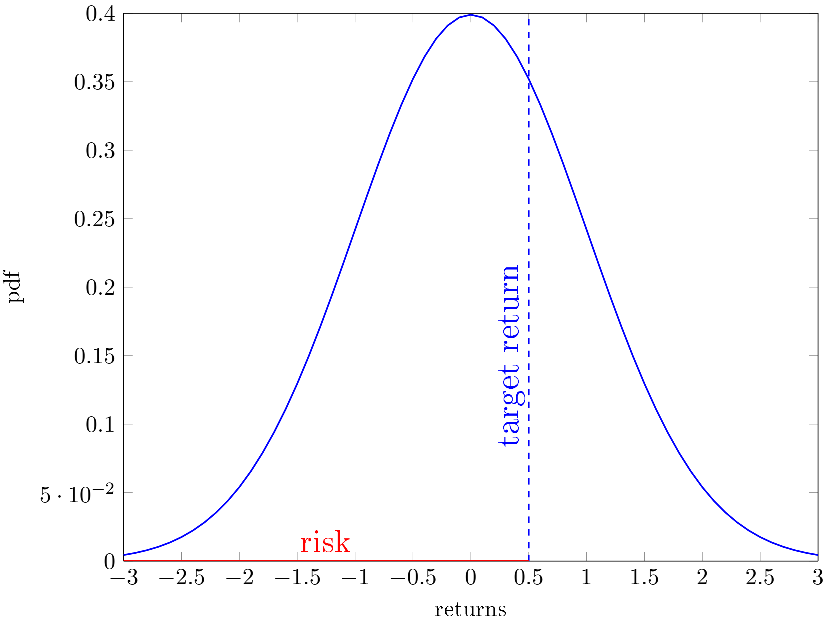


Figure 3.1: Interpretation of Risk by the Sortino Ratio.

return suggests, every return can be taken to fill the position of the former riskless asset return. This makes the Sortino ratio more flexible for comparison.

The Sortino ratio belongs to the reward-risk ratios. The reward is, as for the Sharpe ratio, the excess return. The TDD as measure of risk fulfills only partly the properties of an uncertainty measure as defined in Definition 3.9. Opposed to the standard deviation, the threshold does change the properties of the measure. If the threshold equals zero, then the first two properties can be fulfilled. For other thresholds only the first property is fulfilled. As we could not find a reference on the properties of the target downside deviation regarding uncertainty measures (Definition 3.9), we include a proof.

Proposition 3.13. *The target downside deviation measure for risk, which is used in the Sortino ratio*

$$\mathrm{TDD}(X) = \sqrt{\mathrm{E}[(\min\{0; X - t_r\})^2]}$$

satisfies the following conditions. In terms of the definition of an uncertainty measure as in Definition 3.9, the TDD fulfills the second and third property only under additional conditions.

1. $\text{TDD}(X+\alpha) \leq \text{TDD}(X)$ *for all* X *and constants* $\alpha \geq 0$.
2. $\text{TDD}(0) = 0$, *if* $t_r \leq 0$; *and* $\text{TDD}(\lambda X) = \lambda\,\text{TDD}(X)$ *for all* X *and* $\lambda > 0$, *if* $t_r = 0$.
3. $\text{TDD}(X) \geq 0$ *for all* X, *with* $\text{TDD}(X) > 0$, *if* $P(X < t_r) > 0$.

Proof. 1. $\text{TDD}(X+\alpha) \leq \text{TDD}(X)$ for all X and constants $\alpha \geq 0$:
As we only consider the positive square root, the results do not change if we square the TDD.

$$(\text{TDD}(X+\alpha))^2 = \text{E}[(\min\{0; X - t_r + \alpha\})^2].$$

Case 1: $X - t_r \leq -\alpha$

$$\begin{aligned}(\text{TDD}(X+\alpha))^2 &= \text{E}[(\min\{0; \underbrace{X - t_r + \alpha}_{\leq 0}\})^2] \\ &= \text{E}[(X - t_r + \alpha)^2] \\ &= \text{E}[(X-t_r)^2 + 2\alpha(X-t_r) + \alpha^2] \\ &= \text{E}[(X-t_r)^2] + 2\alpha\text{E}[X-t_r] + \alpha^2.\end{aligned}$$

Additionally, as $X - t_r \leq 0$, the first term can be expressed as $(\text{TDD}(X))^2$,

$$\begin{aligned}(\text{TDD}(X+\alpha))^2 &= \text{E}[(\min\{0; X - t_r\})^2] + 2\alpha\text{E}[X-t_r] + \alpha^2 \\ &= (\text{TDD}(X))^2 + 2\alpha(\text{E}[X] - t_r) + \alpha^2 \\ &= (\text{TDD}(X))^2 + 2\alpha\text{E}[X] - 2\alpha t_r + \alpha^2.\end{aligned}$$

From the condition in this case we know that $X \leq t_r - \alpha$, thus this also holds for the expectation: $\text{E}[X] \leq t_r - \alpha$. Thus,

$$\begin{aligned}(\text{TDD}(X+\alpha))^2 &\leq (\text{TDD}(X))^2 + 2\alpha(t_r - \alpha) - 2\alpha t_r + \alpha^2 \\ &= (\text{TDD}(X))^2 + 2\alpha t_r - 2\alpha^2 - 2\alpha t_r + \alpha^2 \\ &= (\text{TDD}(X))^2 - \alpha^2 \\ &\leq (\text{TDD}(X))^2.\end{aligned}$$

Taking again the square root we have:

$$\text{TDD}(X+\alpha) \leq \text{TDD}(X).$$

Case 2: $-\alpha \leq X - t_r \leq 0$

$$\begin{aligned}(\text{TDD}(X+\alpha))^2 &= \text{E}[(\min\{0; \underbrace{X - t_r + \alpha}_{\geq 0}\})^2] \\ &= \text{E}[0] \\ &= 0 \\ &\leq \text{E}[(X-t_r)^2] \\ &= \text{E}[(\min\{0; \underbrace{X - t_r}_{\leq 0}\})^2] \\ &= (\text{TDD}(X))^2.\end{aligned}$$

Taking the square root we have:

$$\mathrm{TDD}(X+\alpha) \leq \mathrm{TDD}(X).$$

Case 3: $0 \leq X - t_r$

$$\begin{aligned}(\mathrm{TDD}(X+\alpha))^2 &= \mathrm{E}[(\min\{0; \underbrace{X - t_r + \alpha}_{\geq 0}\})^2] \\ &= \mathrm{E}[(\min\{0; \underbrace{X - t_r}_{\geq 0}\})^2] \\ &= (\mathrm{TDD}(X))^2.\end{aligned}$$

Taking again the square root we have:

$$\mathrm{TDD}(X+\alpha) \leq \mathrm{TDD}(X).$$

2. - $\mathrm{TDD}(0) = 0$ if $t_r \leq 0$:

   $$\mathrm{TDD}(0) = \sqrt{\mathrm{E}[(\min\{0; 0 - t_r\})^2]} = 0$$

   Thus, this condition only holds if $t_r \leq 0$.
   - $\mathrm{TDD}(\lambda X) = \lambda\,\mathrm{TDD}(X)$ for all X, $\lambda > 0$ and $t_r = 0$:
   In terms of squared TDD we need to show the following:

   $$(\mathrm{TDD}(\lambda X))^2 = \mathrm{E}[(\min\{0; \lambda X - t_r\})^2] \stackrel{!}{=} \lambda^2 (\mathrm{TDD}(X))^2$$

   $\mathrm{TDD}(\lambda X) = \lambda\,\mathrm{TDD}(X)$ for all X if $t_r = 0$
   Case 1: $\lambda X \leq t_r$

   $$\begin{aligned}(\mathrm{TDD}(\lambda X))^2 &= \mathrm{E}[(\min\{0; \underbrace{\lambda X - t_r}_{\leq 0}\})^2] \\ &= \mathrm{E}[(\lambda X - t_r)^2] \\ &= \lambda^2 \mathrm{E}[(X - \frac{1}{\lambda} t_r)^2]\end{aligned}$$

   This only leads to the desired terms if either $\lambda = 1$, or $t_r = 0$.
   Case 2: $\lambda X \geq t_r$

   $$\begin{aligned}(\mathrm{TDD}(\lambda X))^2 &= \mathrm{E}[(\min\{0; \underbrace{\lambda X - t_r}_{\geq 0}\})^2] \\ &= \mathrm{E}[0] \\ &= \lambda^2 \mathrm{E}[(\min\{0; \lambda X - t_r\})^2]\end{aligned}$$

   Thus we either need $\lambda = 1$, or $t_r = 0$ to get to the desired result as the second case holds trivially.

3. $\mathrm{TDD}(X) \geq 0$ for all X, with $\mathrm{TDD}(X) > 0$ if $P(X < t_r) > 0$:
   Here, we want to examine if the TDD is always greater or equal zero and only zero, when X is constant. It can easily be shown that the first assumption holds, as the TDD is constructed such that it can only be positive or zero. In particular, $\mathrm{TDD}(X) > 0$ if $P(X < t_r) > 0$. Note that, if $X \geq t_r$, unfortunately, the TDD is also zero for non-constant X.

□

Note also that, if $t_r > 0$ the target downside deviation of zero equals the threshold:

$$\mathrm{TDD}(0) = \sqrt{\mathrm{E}[(\min\{0; 0 - t_r\})^2]} = t_r.$$

This is reasonable since the threshold defines the zero risk level. However, in this case the second property of an uncertainty measure is violated as $\mathrm{TDD}(0) \neq 0$. The third property requires that the uncertainty measure only equals zero if the random variable is constant. However, as we shift the zero risk level with the threshold this property can no longer be fulfilled. For constant random variables, there is no uncertainty about the outcome, but with the threshold we already rate this outcome. Thus, we can say that the target downside deviation measures the dangerous part of uncertainty.

3.4.3 Omega Ratio

Another way to separate the downside risk from upside deviation is the omega ratio by Keating and Shadwick (2002). They derive a ratio of the returns above and below a threshold. In terms of the distribution function this leads to the following formula:

$$OmegaRatio(P) = \frac{\int_{t_r}^{\infty}(1 - F_R(x))\mathrm{d}x}{\int_{-\infty}^{t_r} F_R(x)\mathrm{d}x} = \frac{\mathrm{E}[(R - t_r)^+]}{\mathrm{E}[(R - t_r)^-]},$$

where F_R is the cumulative distribution function of the returns R of the portfolio P and t_r the threshold, which can be the return of a riskless asset, but as in the Sortino ratio it can be chosen freely. Moreover, $\mathrm{E}[(\ .\ )^+]$ indicates the expectation of the positive part of the term in brackets and $\mathrm{E}[(\ .\ )^-]$ the expectation of the negative part. That this last term equals the first formula can be shown by using integration by parts or applying double integrals.

To approximate the ratio for historical data, either the empirical distribution function or the resulting mean of the positive and negative part can be used. If the threshold equals the expected return or mean return, the omega ratio equals 1, it is greater 1, if the upside deviation is higher than the downside deviation, and less than 1, if it is the other way around. So, the omega ratio "is the probability weighted ratio of gains to losses, relative to the return", see Keating and Shadwick (2002). In fact, it is just the ratio of the blue area above to the red area below the exemplary distribution function in Figure 3.2, if the threshold is set to 4. Therefore, it is reasonable to examine the denominator of the omega ratio as a measure of risk. The advantage of this performance measure is, that it is not reduced to some parameters of the return distribution. It uses the whole empirical distribution function for the data to evaluate the performance. A better interpretation of the risk and the reward of the asset is the consequence. Only very rare events are unlikely to occur in a historical data set. However, if the distribution truly shows skew or another probability mass shifting structure, the omega ratio will take this

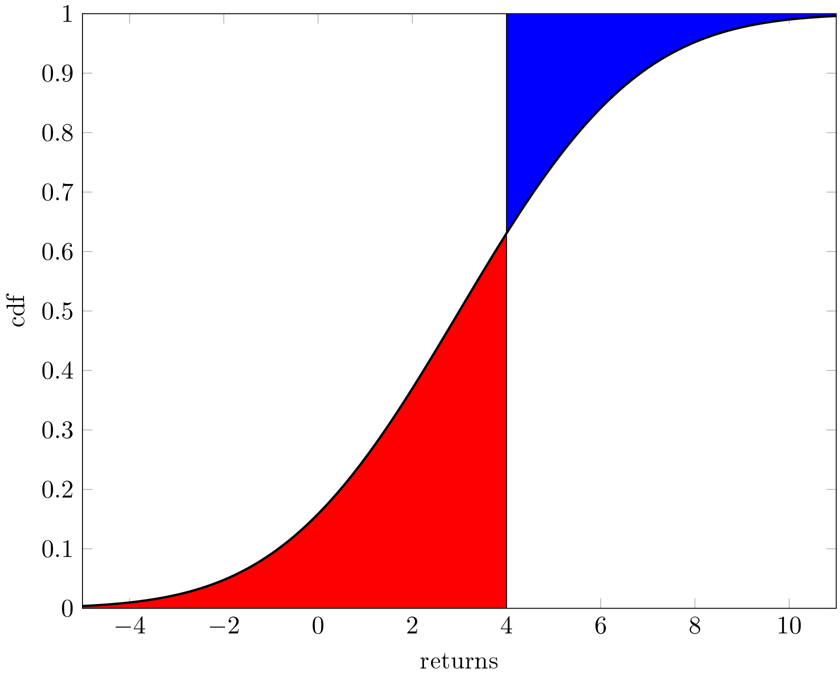


Figure 3.2: Omega Ratio

into account. An additional advantage of the omega ratio is that it shows how performance can change with the changing threshold.

Again, we include the proofs for the properties of the omega ratio in terms of Section 3.1, because we could not find a reference. We show that the omega ratio is a reward-risk ratio, where the risk measure only satisfies the first property of an uncertainty measure (3.9). And the reward measure only satisfies the general property, (3.1).

Proposition 3.14. *Denote the measure for risk, which is used in the omega ratio*

$$\Omega\text{-risk}(P) = \int_{-\infty}^{t_r} F_R(x)\mathrm{d}x$$

by Ω*-risk. It satisfies the following conditions. In terms of the definition of an uncertainty measure as in Definition 3.9, the* Ω*-risk only fulfills the first part of the second property and the third property under conditions.*

1. $\Omega\text{-risk}(X+\alpha) \leq \Omega\text{-risk}(X)$ *for all* X *and constants* $\alpha \geq 0$.
2. $\Omega\text{-risk}(0) = 0$, *if* $t_r \leq 0$.
3. $\Omega\text{-risk}(X) \geq 0$ *for all* X, *with* $\Omega\text{-risk}(X) = 0$ *if* $X = c$ *constant or* X *discrete (minimal possible value* c*) and* $t_r \leq c$.

Proof. 1. $\Omega\text{-risk}(X+\alpha) \leq \Omega\text{-risk}(X)$ for all X and constants $\alpha \geq 0$:

$$\Omega\text{-risk}(X+\alpha) = \int_{-\infty}^{t_r} F_{X+\alpha}(x)\mathrm{d}x.$$

In terms of probability functions, we get:

$$F_{X+\alpha}(x) = P(X + \alpha \leq x) = P(X \leq x - \alpha) = F_X(x - \alpha).$$

Since the cumulative distribution function is non-decreasing, we can deduce that:

$$\int_{-\infty}^{t_r} F_X(x - \alpha)\mathrm{d}x \leq \int_{-\infty}^{t_r} F_X(x)\mathrm{d}x = \Omega\text{-risk}(X),$$

which shows the desired property.

2. Ω-risk$(0) = 0$, if $t_r \leq 0$:
If $X = 0$ we get

$$\Omega\text{-risk}(0) = \int_{-\infty}^{t_r} F_0(x)\mathrm{d}x,$$

and the integral only equals zero if $t_r \leq 0$.

3. Ω-risk$(X) \geq 0$ for all X, with Ω-risk$(X) = 0$ if $X = c$ constant or X discrete (minimal possible value c) and $t_r \leq c$:
The first assumption holds naturally, as the cumulative distribution function takes values in [0;1] and therefore the integral is always greater or equal zero. In the case of constant random variables ($X = c$) the distribution function is a step function which equals zero for all values smaller c and equals 1 for all other values. Thus the value of the Ω-risk depends on the threshold, and we get:

$$\Omega\text{-risk}(c) = \int_{-\infty}^{t_r} F_c(x)\mathrm{d}x = 0 \iff t_r < c.$$

The case of discrete random variables is quite similar. The cumulative distribution function is again a step function, but this time with more than one step. Thus, the value of the Ω-risk again depends on the threshold and the smallest possible value for X, c:

$$\Omega\text{-risk}(X) = \int_{-\infty}^{t_r} F_X(x)\mathrm{d}x = 0 \iff t_r < c.$$

□

Note that the second part of property 2 can only be fulfilled in the trivial case, Ω-risk$(\lambda X) = \lambda\,\Omega$-risk$(X)$ for all X only if $\lambda = 1$:
Here we would need the following:

$$\Omega\text{-risk}(\lambda X) = \int_{-\infty}^{t_r} F_{\lambda X}(x)\mathrm{d}x \stackrel{?}{=} \int_{-\infty}^{t_r} \lambda F_X(x)\mathrm{d}x.$$

This would require that

$$P(\lambda X \leq x) = \lambda P(X \leq x),$$

which only holds for all x in the trivial case, i.e. $\lambda = 1$.
Again, there are problems with the uncertainty measure properties concerning the threshold, and the interpretation of multiples of assets is not unique.

Proposition 3.15. *The* Ω-reward *only satisfies the general property of a reward measure,* $X \geq Y \Longrightarrow v(X) \geq v(Y)$.

Proof. If $X \geq Y$, i.e. the investor prefers portfolio X, then the cumulative distribution function of X is pointwise smaller or equal than the cumulative distribution function of Y because

$$P(X \leq x) \leq P(Y \leq x)$$

for all $x \in \mathbb{R}$. This tells us that for each x the probability for X taking smaller values than x is smaller than for the Y, which indicates, that the probability for higher values is greater for X than for Y. Thus, we get

$$F_X(x) \leq F_Y(x)$$

pointwise. Then taking $1 - F(x)$ turns around the unequal sign:

$$1 - F_X(x) \geq 1 - F_Y(x).$$

Then this relation also holds for the integral:

$$\int_{t_r}^{\infty} 1 - F_X(x)\mathrm{d}x \geq \int_{t_r}^{\infty} 1 - F_Y(x)\mathrm{d}x,$$

i.e. Ω-reward$(X) \geq \Omega$-reward(Y). □

3.4.4 Portfolio Turnover

The portfolio turnover is a measure for the amount of rebalancing needed for the portfolio. It sums up how much the weights of the assets in the portfolio change in the holding period and have to be adjusted for the next rebalancing. The theoretical formula is as follows:

$$Turnover(P) = \mathrm{E}[\sum_{n=1}^{N} |w_{n,t+1} - w_{n,t^+}|].$$

Using historical data we apply this formula:

$$Turnover(P) = \frac{1}{T} \sum_{t=1}^{T} \sum_{n=1}^{N} |w_{n,t+1} - w_{n,t^+}|,$$

where $w_{n,t}$ is the weight of asset n at the beginning of period t, the one which is set by the portfolio strategy, and w_{n,t^+} is the weight of asset n at the end of period t produced by the asset price change in this period. Thus $|w_{n,t+1} - w_{n,t^+}|$ is the absolute value of the change in weights caused by rebalancing for the new period $t + 1$. Therefore, the portfolio turnover is a hint at potential transaction costs initiated by the portfolio strategy. So, low portfolio turnover is preferred.

3.4.5 Certainty Equivalent Return

The certainty equivalent (CEQ) return is another way to combine the mean and standard deviation of the asset returns to one value. However, in contrast to the Sharpe ratio the risk aversion of an investor can be included. In total, the following formula leads to the riskless return which an investor would rather choose than to invest in the risky portfolio. Thus, the higher the value the better the portfolio strategy.

$$CEQ(P) = \mathrm{E}[R - t_r] - \frac{\gamma}{2}\mathrm{Var}(R - t_r),$$

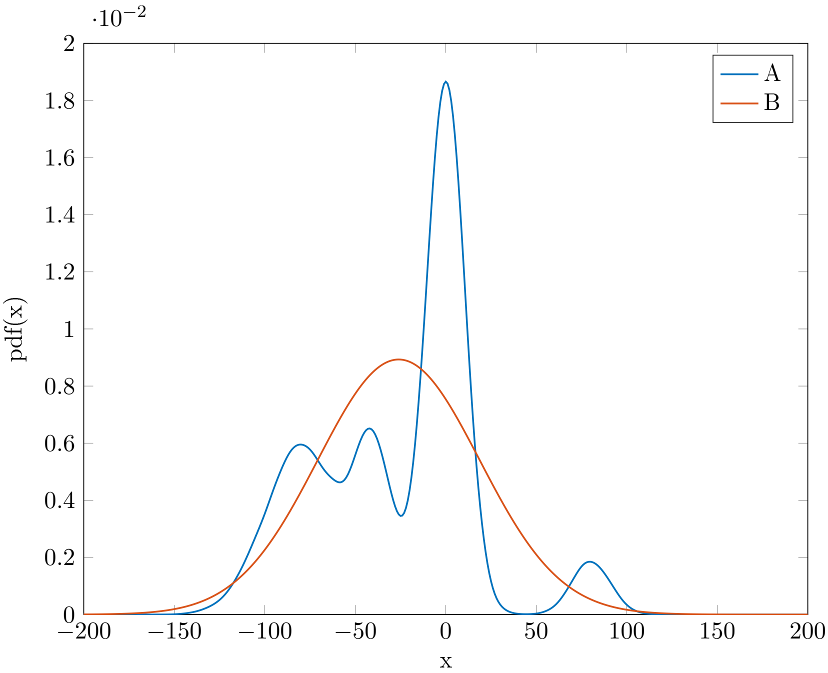


Figure 3.3: Comparison of two distributions with the same mean and variance. Curve B belongs to a normal distribution and curve A results from combining several normal distributions.

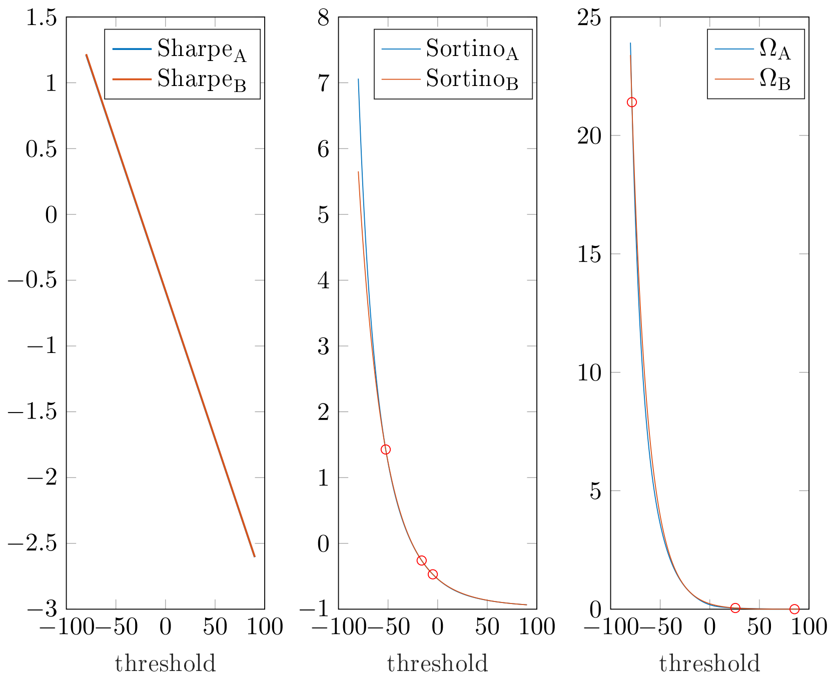


Figure 3.4: Comparison of the Sharpe, Sortino and Omega ratio. The red dots indicate points of intersection. Thus they show the threshold for which the preference order of the two distributions is turned around.

where the expectation and the variance are replaced by their corresponding estimators, when using historical data. The higher the risk aversion parameter γ, the more risk averse the investor. In our simulation studies we set $\gamma = 1$, for comparison with other studies in the literature, e.g. those in DeMiguel, Garlappi, and Uppal (2009).

3.4.6 Diversification-Risk Ratio

As we have seen, there is no best procedure to measure the risk of a portfolio. However, the procedures only rely on the given return series of the specific portfolio. Also diversification itself decreases the risk. Therefore, we suggest to involve a diversification per risk ratio as performance criterion into the selection process. When we know that two portfolios have the same risk but differ in terms of diversification, then we would choose the more diversified portfolio. In this case, the ratio of the more diversified portfolio would be higher than the one of the less diversified portfolio. If we know that two portfolios are at a common level of diversification but one leads to higher risk, then we would prefer the one with lower risk. In that case, the ratio of the portfolio with lower risk would lead to a larger ratio than the one with higher risk. Again, we would be able to judge two performance criteria from one value.

Now, the question is how to choose the risk and the diversification measure. From what we learned so far it is reasonable to choose the Ω-risk as risk measure and the diversification measure using Rao's quadratic entropy with correlation as in Carmichael, Koumou, and Moran (2015). The theoretical version is then given by:

$$DivRiskRatio(P) = \frac{H_D(w)}{\Omega\text{-risk}(P)} = \frac{\sum_{i,j=1}^{N}(1-\rho_{i,j})w_i w_j}{\int_{-\infty}^{t_r} F_R(x)\mathrm{d}x}.$$

The version for a given data sample is then:

$$DivRiskRatio(P) = \frac{\sum_{i,j=1}^{N}(1-\widehat{\rho}_{i,j})w_i w_j}{\frac{1}{T}\sum_{i=1}^{T}\min\{0, R_i - t_r\}}.$$

Here, the same notation is used as before: P is the portfolio with returns R and cumulative distribution function F_R, which results from the returns of N assets together with portfolio weights $w_i \geq 0$ for $i = 1\ldots, N$ and $\sum_{i=1}^{N} w_i = 1$. Then $\rho_{i,j}$ is the correlation coefficient of asset i and j with its estimate $\widehat{\rho}_{i,j}$, and t_r denotes the chosen threshold.

If test periods are too small, then the diversification-risk ratio may not lead to finite values. This is the case when the Ω-risk becomes zero. This has already happened for the test period of 12 months with monthly data. Thus, we have to evaluate the Ω-risk over longer periods. Doing so, several rebalancing steps might be present in the longer period. Thus, we get more than one value for the diversification measure. We take the mean over all these values to obtain one value for the whole period.

Let us take a look at an example where we are given 10 assets. They form 5 groups with 2 assets each, which are perfectly positive correlated. The groups are independent from each other. Then 3 naive portfolios are compared: the full naive portfolio, the portfolio which chooses the best asset from each group and the portfolio which chooses the worst asset from each group. As we have equally sized groups, the diversification measure is equal for each portfolio and only the risk changes, see Section 5.3. Thus, the portfolio selecting the best assets is superior to the naive portfolio, leaving out the bad ones. This results in a higher diversification risk ratio. The portfolio, which only selects the worst assets performs worse than the full naive portfolio. The following example illustrates this.

mean	0.0	0.1	0.0	0.1	0.0	0.2	0.0	0.25	0.0	0.45
volatility	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2

Table 3.1: annual mean and volatility pairs for this simulation

In Table 3.1 we see the annual mean and volatility pairs for the 10 assets used to derive daily log returns data. The asset groups are separated by vertical lines. We multiply the resulting returns with 100 such that the Ω-risk values become larger. Otherwise, the diversification risk ratios would be too large and we were not able to get a good fit for the densities, see Figure 3.5. Thus, we get the following diversification-risk ratios for the portfolios:

$$DivRiskRatio(P_{naive}) = 175.77$$
$$DivRiskRatio(P_{selectbest}) = 184.47$$
$$DivRiskRatio(P_{selectworst}) = 168.22$$

The density functions fitted to the data is given in Figure 3.5. There, the increasing risk is displayed by the left shift of the densities.

One drawback of this ratio is that, if the threshold is too low for the given data then the denominator becomes zero and the ratio is no longer defined. When we replace the ratio by infinity in this case, we result in a mathematically reasonable value (limit value, when the denominator goes to zero), but the diversification measure no longer plays a role. Thus, it is better to calculate the ratio either on more data, when values under the threshold are likely, or on a lower scale, e.g use monthly data instead of annual data and then annualize if necessary.

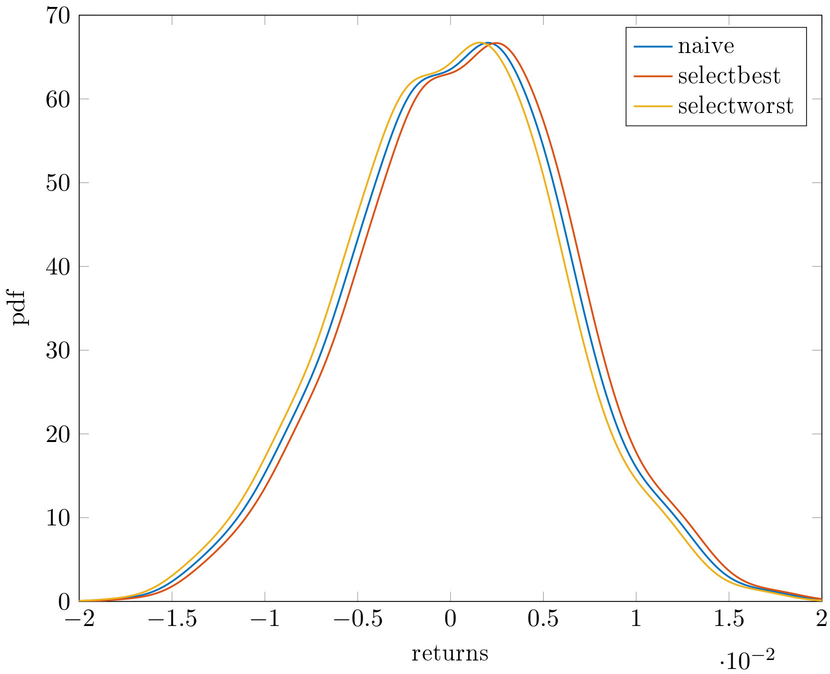


Figure 3.5: Densities fitted to the data of the full naive portfolio and the naive portfolios selecting assets from each group.

4 Naive Diversification

In this chapter we show how the advantages of the prominent $\frac{1}{N}$ rule pass on to a naive portfolio with fewer assets.

The first section (Section 4.1) recapitulates the advantages of the naive portfolio, when facing statistical or model uncertainty. Moreover, we show that the growth rate of the full naive portfolio equals the growth rate of the naive portfolio with fewer assets in a constant rebalancing setting following the approach by Cover (1991) (Section 4.2). Section 4.3 adds a discussion on how estimation errors influence the performance of the mean-variance portfolio compared to the naive portfolio following the guidelines of DeMiguel, Garlappi, and Uppal (2009) and provides an extension of their results to a naive portfolio with fewer assets. Results on real and simulated data complete this section. The next part (Section 4.5) compares the performance of three mean-variance optimization variants, the unconstrained, the constrained and the minimum variance version. This leads to the subsequent choice of the constrained mean-variance portfolio for the comparison of methods. A summary of the results of this chapter is given in Section 4.7.

4.1 Reduction of the Number of Assets

'Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?' by DeMiguel, Garlappi, and Uppal (2009) was the first paper about naive diversification we came across. The authors recommend the $\frac{1}{N}$ portfolio to be used as a benchmark for performance evaluation regarding the out-of-sample Sharpe ratio, certainty equivalent (CEQ) return and the turnover of the trading volume. In the end, they show the superiority of the naive diversification over several prominent investment strategies, including the outperformance of the sample mean-variance portfolio. This is caused by the simple form of the naive strategy, which needs no parameter estimation for its computation.

A rigorous asymptotic view on the advantages of the $\frac{1}{N}$ rule is given in the paper *'The 1/N investment strategy is optimal under high model ambiguity'* by Pflug, Pichler, and Wozabal (2012). Here, the risk of model misspecification is considered. They show that with increasing uncertainty about the underlying model the $\frac{1}{N}$ rule becomes the best possible investment strategy.

However, as good as the naive diversification may seem, it has several drawbacks. First of all, this method leads to very large portfolios. Considering all available assets, naive diversification allocates very small amounts of the given capital to thousands of investment positions.

Now, the goal is to find a trading strategy which maintains the advantages of the naive diversification and at the same time avoids its disadvantages, mainly the large portfolios. Therefore, we reduce the number of assets and examine the performance with respect to the full naive portfolio, which includes all available assets. For this purpose we explore the portfolios from different point of views.

4.2 Constant Rebalanced Portfolios

First, we examine the expected growth rate of naive portfolios with fewer assets. In the context of constant rebalanced portfolios, a fundamental paper is *'Universal portfolios'* by Cover (1991). He proposes an average portfolio generated by every possible portfolio, using only the past stock prices. This so-called universal portfolio has many good properties and it gets near the best constant rebalanced portfolio knowing the future. The most important advantage is that this portfolio does not suffer from estimation errors, hence is motivated as the $\frac{1}{n}$ portfolio.

If we allow only for fewer assets in the naive portfolio, we need to choose from the available assets. In the constant rebalancing setting we can select the $n \leq N$ assets randomly in each rebalancing step. Following the approach by Cover (1991), we can show that the expected growth rate of the naive portfolio with fewer assets is the same as the growth rate of the full naive portfolio.

Therefore, we want to check the properties of the $\frac{1}{N}$ constant rebalanced portfolio and compare it to the $\frac{1}{n}$ constant rebalanced portfolio which has fewer assets.

The Cover (1991) setting is different from the settings throughout this thesis. However, we will adapt the notation to our setting. His setting is based on the so-called price relative $r_i = \frac{p_i^{close}}{p_i^{open}}$, for the open and close prices p_i^{open} and p_i^{close} of asset i, which is the factor of price change for one time step: $p_i^{open} r_{t+1,i} = p_i^{close}$. This is just the absolute return $r_{t+1,i} = \frac{p_{t+1}}{p_t}$ for the time step t to $t+1$, if there is no jump over night. Then $r_{t+1} = (r_1, \ldots, r_N)$ is the vector of price relatives for all considered assets, the stock market vector. A portfolio is specified by the relative weight vector $x = (x_1, \ldots, x_N)$, representing the proportion of wealth invested in the corresponding asset. In addition, we have a no-shortselling constraint as $x_i \geq 0$ and $\sum_{i=1}^N x_i = 1$. Let C be the initial capital invested in the portfolio, then the initial value of the portfolio can be calculated as follows:

$$V_0(x) = \sum_{i=1}^N \frac{x_i C}{p_i^0} p_i^0 = C \sum_{i=1}^N x_i = C.$$

Thus, the value of the portfolio at time $t+1$ is

$$V_{t+1}(x) = \sum_{i=1}^N \frac{x_i V_t}{p_i^t} p_i^{t+1} = V_t \sum_{i=1}^N x_i \frac{p_i^{t+1}}{p_i^t} = V_t \sum_{i=1}^N x_i r_{t+1,i}.$$

Then

$$Y_{t+1} = \sum_{i=1}^N x_i r_{t+1,i} = x' r_{t+1}$$

is the factor of wealth change: $V_t \cdot Y_{t+1} = V_{t+1}$, where r_{t+1} is the vector of price relatives for the time step t to $t+1$. Let $r_1, \ldots, r_T \in \mathbb{R}_+^N$ be a sequence of stock vectors, for days $t = 1, \ldots, T$. The final wealth of a constant rebalanced portfolio with initial capital $C = 1$ is then

$$V_T(x) = \prod_{t=0}^{T-1} Y_{t+1} = \prod_{t=1}^{T} x' r_t.$$

The optimal achievable wealth from a constant rebalanced portfolio is then:

$$V_T^* = \max_{x \in \mathbb{R}_+^N} V_T(x).$$

Now, we can show that the expected final wealth of the $\frac{1}{N}$ constant rebalanced portfolio equals the expected final wealth of the $\frac{1}{n}$ constant rebalanced portfolio with fewer assets.

Theorem 4.1. *Let x_{ew} and x_τ be the portfolio vectors for the $\frac{1}{N}$ constant rebalanced portfolio and the randomly chosen $\frac{1}{n}$ portfolio, respectively, consisting of the fractions of wealth invested in the different assets. The randomly chosen portfolio consists of exactly n assets. Further assume that the sequence of price relatives, $r_1, \dots, r_T$, is independent and also independent of the choices, $\tau_1, \dots, \tau_T$, of the assets. Then the expected growth-rate of the $\frac{1}{N}$ portfolio equals the expected growth rate of the $\frac{1}{n}$ portfolio,*

$$\mathrm{E}[V_T(x_{ew})] = \mathrm{E}[V_T(x_\tau)].$$

Proof. The growth rate of the (equal weight) constant rebalanced $\frac{1}{N}$ portfolio satisfies

$$V_T(x_{ew}) = \prod_{t=1}^{T} x'_{ew} r_t = \prod_{t=1}^{T} \frac{1}{N} \mathbf{1}'_N r_t = \frac{1}{N^T} \prod_{t=1}^{T} \mathbf{1}'_N r_t,$$

where $\mathbf{1}_N$ is the vector where all N entries equal 1. For the random portfolio we get

$$V_T(x_\tau) = \prod_{t=1}^{T} x'_{\tau t} r_t,$$

where $x_{\tau t}$ is a vector with n nonzero entries equal to $\frac{1}{n}$. The positions of the nonzero entries are randomly chosen by the random vector τ_t. Its entries are digital, 1 for this asset is chosen, 0 for this asset is not chosen. Thus the formula above can be altered to:

$$V_T(x_\tau) = \prod_{t=1}^{T} x'_{\tau t} r_t = \prod_{t=1}^{T} \frac{1}{n} \tau'_t r_t = \left(\frac{1}{n}\right)^T \prod_{t=1}^{T} \tau'_t r_t.$$

Taking expectation yields:

$$\mathrm{E}[V_T(x_\tau)] = \mathrm{E}[\left(\frac{1}{n}\right)^T \prod_{t=1}^{T} \tau'_i x_i] = \left(\frac{1}{n}\right)^T \mathrm{E}[\prod_{t=1}^{T} \tau'_t r_t].$$

As the τ_t and the r_t are assumed to be independent the expectation and the product can be interchanged,

$$\mathrm{E}[V_T(x_\tau)] = \left(\frac{1}{n}\right)^T \prod_{t=1}^{T} \mathrm{E}[\tau'_t r_t] = \left(\frac{1}{n}\right)^T \prod_{t=1}^{T} \sum_{i=1}^{N} \mathrm{E}[\tau_{t,i} r_{t,i}].$$

With the independence of the price relative and the choice of assets the expectation of the price relative for one asset leads to $\mathrm{E}[\tau_{t,i} r_{t,i}] = \mathrm{E}[r_{t,i}] P(\tau_{t,i} = 1)$, where $P(\tau_{t,i} = 1)$ denotes the probability that asset i is chosen for the portfolio $x_{\tau t}$. This probability can be derived from the hypergeometric distribution. Consider an urn that contains one white ball, which is the asset we are interested in, and $N-1$ black balls, which represent the other assets. Now, we draw exactly n balls and we want to know if the white ball is taken as well. Thus, we are in a setting where the probabilities for the hypergeometric distribution are nonzero and we get the following:

$$P(\tau_{ti} = 1) = \frac{\binom{1}{1}\binom{N-1}{n-1}}{\binom{N}{n}} = \frac{(N-1)!n!n!}{(n-1)!n!N!} = \frac{n}{N}.$$

Therefore, we conclude:

$$\mathrm{E}[V_T(x_\tau)] = \left(\frac{1}{n}\right)^T \prod_{t=1}^{T}\sum_{i=1}^{N}\frac{n}{N}\mathrm{E}[r_{ti}] = \left(\frac{n}{nN}\right)^T \prod_{t=1}^{T}\sum_{i=1}^{N}\mathrm{E}[r_{ti}] = \left(\frac{1}{N}\right)^T \prod_{t=1}^{T}\mathbf{1}'_N\mathrm{E}[r_t].$$

□

This includes the naive portfolio with half the number of assets, which we will further examine in this chapter. Moreover, it shows that we would not lose too much, if we choose a naive portfolio with fewer assets. Surely, this statement only holds if the rebalancing takes place very often so that the expected growth rate is reached in reality.

4.3 Statistical Robustness

In the same way as in the paper by DeMiguel, Garlappi, and Uppal (2009), we are going to calculate the critical length of the estimation period for the mean-variance portfolio to outperform the naive portfolio, but now with fewer than all assets. The setting is the same as in the paper extended by the possibility of choosing fewer than all available assets for the naive portfolio. Thus, the following theorem is a generalization of Proposition 1 from that paper.

Here, we want to show the relation of the expected performance of the sample mean-variance portfolio and of the naive portfolio with 3 up to N assets, where the last case is exactly the result by DeMiguel, Garlappi, and Uppal (2009). In the world where drift and volatility explain everything an investor wants to maximize the return and simultaneously minimize the variance, which represents the risk of the portfolio. The portfolio is represented by the weight vector x, which gives the amount of money invested in each risky asset. This leads to the following Markowitz utility function (4.1). (Note that the terms 'utility function' and 'expected utility' do not coincide with the terminology from decision theory. In some sense this is a second order Taylor expansion for the expected utility for a certain power utility function.):

$$U(x) = x'\mu - \frac{\gamma}{2}x'\Sigma x, \tag{4.1}$$

where γ is the relative risk aversion of the investor. The higher γ the higher is the risk aversion as it magnifies the variance in the utility function. Since $\gamma > 0$ and U is concave in x, there is an optimal solution to maximizing utility. We include the derivation of the results from the paper. For unconstrained $x \in \mathbb{R}^N$ we aim to solve

$$\max_{x\in\mathbb{R}^N}\left(x'\mu - \frac{\gamma}{2}x'\Sigma x\right).$$

Since the utility function is concave in x, we find the optimal solution with the first order condition,

$$\mu - \gamma\Sigma x = 0.$$

This yields $\frac{1}{\gamma}\mu = \Sigma x$ and this results in the optimal portfolio for this model,

$$x^* = \frac{1}{\gamma}\Sigma^{-1}\mu.$$

Then we can calculate the expected loss for using the portfolio rule $\hat{x}$ instead of the optimal solution x^*,

$$L(x^*, \hat{x}) = U(x^*) - \mathrm{E}[U(\hat{x})].$$

The first term is simply plugging in the optimal solution into the utility function,

$$\begin{aligned} U(x^*) &= \frac{1}{\gamma}(\Sigma^{-1}\mu)'\mu - \frac{\gamma}{2}\left(\frac{1}{\gamma}\Sigma^{-1}\mu\right)'\Sigma\left(\frac{1}{\gamma}\Sigma^{-1}\mu\right) \\ &= \frac{1}{\gamma}\mu'\Sigma^{-1}\mu - \frac{1}{2\gamma}\mu'\Sigma^{-1}\Sigma\Sigma^{-1}\mu \\ &= \frac{1}{\gamma}\mu'\Sigma^{-1}\mu - \frac{1}{2\gamma}\mu'\Sigma^{-1}\mu \\ &= \frac{1}{2\gamma}\mu'\Sigma^{-1}\mu. \end{aligned}$$

Together with the formula for the squared Sharpe ratio of the optimal portfolio:

$$S_*^2 = \frac{((x^*)'\mu)^2}{(x^*)'\Sigma x^*} = \frac{((\frac{1}{\gamma}\Sigma^{-1}\mu)'\mu)^2}{(\frac{1}{\gamma}\Sigma^{-1}\mu)'\Sigma\frac{1}{\gamma}\Sigma^{-1}\mu} = \frac{(\frac{1}{\gamma})^2(\mu'\Sigma^{-1}\mu)^2}{(\frac{1}{\gamma})^2\mu'\Sigma^{-1}\Sigma\Sigma^{-1}\mu} = \frac{(\mu'\Sigma^{-1}\mu)^2}{\mu'\Sigma^{-1}\mu} = \mu'\Sigma^{-1}\mu,$$

we get the optimal utility:

$$U(x^*) = \frac{1}{2\gamma}S_*^2.$$

In the following theorem we include the naive portfolios with fewer than N assets and compare their expected losses to the sample-based mean-variance portfolio. The proof by DeMiguel, Garlappi, and Uppal (2009) is just transferred to these cases. For the interpretation of the following theorem we need to allow to split up the capital in the risky assets and in a risk free asset (for which we assume interest rate 0). We use the following notation.

Definition 4.2. *A c-weighted $\frac{1}{n}$ portfolio is a portfolio that invests equal weights c in each of the n assets of the $\frac{1}{n}$ portfolio.*

Thus, for a c-weighted $\frac{1}{n}$ portfolio a fraction $n \cdot c$ is invested in the risky assets, and $1 - n \cdot c$ in the riskless asset. The $\frac{1}{n}$-weighted $\frac{1}{n}$ portfolio then is the $\frac{1}{n}$ portfolio as introduced before. Note that all c-weighted $\frac{1}{n}$ portfolios have the same Sharpe ratio,

$$S_n = \frac{c\mathbf{1}_n'\mu}{\sqrt{c^2\mathbf{1}_n'\Sigma\mathbf{1}_n}} = \frac{\mathbf{1}_n'\mu}{\sqrt{\mathbf{1}_n'\Sigma\mathbf{1}_n}}.$$

Theorem 4.3. *Given N risky assets, let $S_*^2 = \mu'\Sigma^{-1}\mu$ be the squared Sharpe ratio of the mean-variance portfolio, and for $n = 3, \dots N$ let $S_n^2 = (\mathbf{1}_n'\mu)^2/\mathbf{1}_n'\Sigma\mathbf{1}_n$ be the squared Sharpe ratio of the $\frac{1}{n}$ portfolio, which is the equal weights portfolio for any n fixed assets and $\mathbf{1}_n$ is the vector with n entries equal to 1 and $N - n$ entries equal to 0. Given returns data $R_1, \dots, R_M$, and assume $\mathbf{1}_n'\mu > 0$, then:*

1. *If μ is unknown and Σ is known, the sample-based mean-variance strategy has a lower expected loss than the best c-weighted $\frac{1}{n}$ portfolio if:*

$$S_*^2 - S_n^2 - \frac{N}{M} > 0. \tag{4.2}$$

2. *If μ is known and Σ is unknown, the sample-based mean-variance strategy has a lower expected loss than the best c-weighted $\frac{1}{n}$ portfolio if:*

$$kS_*^2 - S_n^2 > 0, \tag{4.3}$$

where

$$k = \left(\frac{M}{M-N-2}\right)\left(2 - \frac{M(M-2)}{(M-N-1)(M-N-4)}\right) < 1.$$

3. *If both μ and Σ are unknown, the sample-based mean-variance strategy has a lower expected loss than the best c-weighted $\frac{1}{n}$ portfolio if:*

$$kS_*^2 - S_n^2 - h > 0, \tag{4.4}$$

where

$$h = \frac{NM(M-2)}{(M-N-1)(M-N-2)(M-N-4)} > 0.$$

Proof. We first calculate in (a) the expected loss of the sample mean-variance portfolio in the three cases, then in (b) we calculate the loss for the c-weighted $\frac{1}{n}$ portfolio, and we compare the losses in (c).

(a) To compare the expected losses we first need the expected loss of the sample mean-variance portfolio for the three cases above, which is a detailed reformulation of the proof in DeMiguel, Garlappi, and Uppal (2009) and Kan and Zhou (2007).

1. When Σ is known and μ is unknown we need to find the conditional expected loss:

$$L(x^*, \hat{x}_{mvo}|\Sigma) = U(x^*) - \mathrm{E}[U(\hat{x}_{mvo})|\Sigma],$$

where $\hat{x}_{mvo} = f(R_1, \ldots, R_M)$ is just a function of the given data. For the sample-mean-variance rule we take the optimal solution and plug in Σ and the sample mean $\hat{\mu} = \frac{1}{M}\sum_{i=1}^{M} R_i$ with $\hat{\mu} \sim \mathcal{N}(\mu, \frac{1}{M}\Sigma)$, thus we get $\hat{x}_{mvo} = \frac{1}{\gamma}\Sigma^{-1}\hat{\mu}$. Then we can calculate the expected utility given the covariance matrix.

$$\begin{aligned}\mathrm{E}[U(\hat{x}_{mvo})|\Sigma] &= \mathrm{E}[\hat{x}_{mvo}'\mu - \frac{\gamma}{2}\hat{x}_{mvo}'\Sigma\hat{x}_{mvo}] = \mathrm{E}[(\frac{1}{\gamma}\Sigma^{-1}\hat{\mu})'\mu - \frac{\gamma}{2}(\frac{1}{\gamma}\Sigma^{-1}\hat{\mu})'\Sigma\frac{1}{\gamma}\Sigma^{-1}\hat{\mu}] \\ &= \frac{1}{\gamma}\mu'\Sigma^{-1}\mu - \frac{\gamma}{2\gamma^2}\mathrm{E}[\hat{\mu}'\Sigma^{-1}\Sigma\Sigma^{-1}\hat{\mu}] = \frac{1}{\gamma}\mu'\Sigma^{-1}\mu - \frac{1}{2\gamma}\mathrm{E}[\hat{\mu}'\Sigma^{-1}\hat{\mu}]\end{aligned}$$

From Theorem 3.3.3 in Anderson (1984), we know that if the m-component vector $Y \sim \mathcal{N}(\nu, C)$ then $Y'C^{-1}Y$ is distributed according to $\chi_m^2(\nu'C^{-1}\nu)$, the noncentral χ^2-distribution with m degrees of freedom and noncentral parameter $\nu'C^{-1}\nu$. Transferring this, we get that

$$\hat{\mu}'\Sigma^{-1}\hat{\mu} \sim \chi_N^2(M\mu'\Sigma^{-1}\mu)/M, \tag{4.5}$$

which can be derived in a few lines:

$$\hat{\mu}'\left(\frac{1}{M}\Sigma\right)^{-1}\hat{\mu} \sim \chi_N^2\left(\mu'\left(\frac{1}{M}\Sigma\right)^{-1}\mu\right)$$

is equivalent to

$$M\hat{\mu}'\Sigma^{-1}\hat{\mu} \sim \chi^2_N\left(M\mu'\Sigma^{-1}\mu\right),$$

which leads to the desired result. For the expected utility we need the expected value of the noncentral χ^2-distribution $\chi^2_N(M\mu'\Sigma^{-1}\mu)$, which is $N + M\mu'\Sigma^{-1}\mu$. Then we get:

$$\mathrm{E}[U(\hat{x}_{mvo})|\Sigma] = \frac{1}{\gamma}\mu'\Sigma^{-1}\mu - \frac{1}{2\gamma}\frac{N + M\mu'\Sigma^{-1}\mu}{M} = \frac{1}{\gamma}S_*^2 - \frac{1}{2\gamma}\frac{N + MS_*^2}{M} = \frac{S_*^2}{2\gamma} - \frac{N}{2\gamma M}.$$

In the end we can calculate the expected loss.

$$\begin{aligned} L(x^*, \hat{x}_{mvo}|\Sigma) = U(x^*) - \mathrm{E}[U(\hat{x}_{mvo})|\Sigma] &= \frac{S_*^2}{2\gamma} - \frac{S_*^2}{2\gamma} + \frac{N}{2\gamma M} \\ &= \frac{N}{2\gamma M}. \end{aligned}$$

2. In this case the mean μ is known and the covariance Σ is unknown, and we want to calculate the expected loss of the sample mean-variance rule,

$$L(x^*, \hat{x}_{mvo}|\mu) = U(x^*) - \mathrm{E}[U(\hat{x}_{mvo})|\mu].$$

Now, we need the sample covariance $\hat{\Sigma} = \frac{1}{M}\sum_{i=1}^{M}(R_i - \hat{\mu})(R_i - \hat{\mu})'$ with $M\hat{\Sigma} \sim \mathcal{W}_N(M-1, \Sigma)$, where $\mathcal{W}_N(\nu, C)$ denotes the Wishart distribution with ν degrees of freedom and covariance matrix C. Thus, we get the investment vector $\hat{x}_{mvo} = \frac{1}{\gamma}\hat{\Sigma}^{-1}\mu$ for the sample mean-variance strategy. This leads to the following expected utility:

$$\begin{aligned} \mathrm{E}[U(\hat{x}_{mvo})|\mu] &= \mathrm{E}[\hat{x}'_{mvo}\mu - \frac{\gamma}{2}\hat{x}'_{mvo}\Sigma\hat{x}_{mvo}] = \mathrm{E}[(\frac{1}{\gamma}\hat{\Sigma}^{-1}\mu)'\mu - \frac{\gamma}{2}(\frac{1}{\gamma}\hat{\Sigma}^{-1}\mu)'\Sigma\frac{1}{\gamma}\hat{\Sigma}^{-1}\mu] \\ &= \frac{1}{\gamma}\mathrm{E}[\mu'\hat{\Sigma}^{-1}\mu] - \frac{1}{2\gamma}\mathrm{E}[\mu'\hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}\mu]. \end{aligned}$$

Now, we use the expected values of the terms with the inverse of the sample covariance matrix, which can be achieved by a reformulation: $\hat{\Sigma}^{-1} = \Sigma^{-\frac{1}{2}}W^{-1}\Sigma^{-\frac{1}{2}}$, where $W = \Sigma^{-\frac{1}{2}}\hat{\Sigma}\Sigma^{-\frac{1}{2}}$ from which we can derive the distribution. Using the reformulation for

$$M\hat{\Sigma} \sim \mathcal{W}_N(M-1, \Sigma)$$

yields

$$M\Sigma^{-\frac{1}{2}}\hat{\Sigma}\Sigma^{-\frac{1}{2}} \sim \mathcal{W}_N(M-1, \Sigma^{-\frac{1}{2}}\Sigma\Sigma^{-\frac{1}{2}})$$

resulting in

$$M\Sigma^{-\frac{1}{2}}\hat{\Sigma}\Sigma^{-\frac{1}{2}} \sim \mathcal{W}_N(M-1, \mathrm{I}_N).$$

Thus, we can reformulate the expected utility as follows:

$$\begin{aligned} \mathrm{E}[U(\hat{x}_{mvo})|\mu] &= \frac{1}{\gamma}\mathrm{E}[\mu'\Sigma^{-\frac{1}{2}}W^{-1}\Sigma^{-\frac{1}{2}}\mu] - \frac{1}{2\gamma}\mathrm{E}[\mu'\Sigma^{-\frac{1}{2}}W^{-1}\Sigma^{-\frac{1}{2}}\Sigma\Sigma^{-\frac{1}{2}}W^{-1}\Sigma^{-\frac{1}{2}}\mu] \\ &= \frac{1}{\gamma}\mathrm{E}[\mu'\Sigma^{-\frac{1}{2}}W^{-1}\Sigma^{-\frac{1}{2}}\mu] - \frac{1}{2\gamma}\mathrm{E}[\mu'\Sigma^{-\frac{1}{2}}W^{-2}\Sigma^{-\frac{1}{2}}\mu]. \end{aligned}$$

Thus, we substitute the problem of the expectation of the product of the inverse sample correlation matrix with the second moment of the inverse of the new matrix W. However, these are known, see e.g. Haff (1979):

$$\begin{aligned}
\mathrm{E}[W^{-1}] &= \frac{M}{M-N-2}\mathrm{I}_N, \\
\mathrm{E}[W^{-2}] &= \frac{M}{M-N-2}\,\frac{M(M-1)}{(M-N-1)(M-N-4)}\mathrm{I}_N.
\end{aligned}$$

Inserting these values into the expected utility formula, and rearranging, leads to:

$$\begin{aligned}
\mathrm{E}[U(\hat{x}_{mvo})|\mu] &= \frac{M}{M-N-2}\mu'\Sigma^{-\frac{1}{2}}\mathrm{I}_N\Sigma^{-\frac{1}{2}}\mu\left(\frac{1}{\gamma}-\frac{1}{2\gamma}\,\frac{M(M-1)}{(M-N-1)(M-N-4)}\right) \\
&= \frac{M}{M-N-2}\mu'\Sigma^{-1}\mu\frac{1}{2\gamma}\left(2-\frac{M(M-1)}{2(M-N-1)(M-N-4)}\right) \\
&= \frac{M}{M-N-2}\left(2-\frac{M(M-1)}{2(M-N-1)(M-N-4)}\right)\frac{S_*^2}{2\gamma} \\
&= k\frac{S_*^2}{2\gamma}.
\end{aligned}$$

Then, we get for the expected loss:

$$L(x^*,\hat{x}_{mvo}|\mu) = U(x^*) - \mathrm{E}[U(\hat{x}_{mvo})|\mu] = \frac{S_*^2}{2\gamma} - k\frac{S_*^2}{2\gamma} = (1-k)\frac{S_*^2}{2\gamma}.$$

3. In the third and most realistic case we know neither μ nor Σ, thus we are looking for the unconditioned expected loss with $\hat{x}_{mvo} = \frac{1}{\gamma}\hat{\Sigma}^{-1}\hat{\mu}$.

$$L(x^*,\hat{x}_{mvo}) = U(x^*) - \mathrm{E}[U(\hat{x}_{mvo})]$$

Using the same computations for the inverse of the sample covariance matrix as in the second case and considering that the sample versions of the mean and the covariance matrix are independent, we get the following expected utility:

$$\begin{aligned}
\mathrm{E}[U(\hat{x}_{mvo})] &= \mathrm{E}[\hat{x}'_{mvo}\mu - \frac{\gamma}{2}\hat{x}'_{mvo}\Sigma\hat{x}_{mvo}] \\
&= \frac{1}{\gamma}\mathrm{E}[\hat{\mu}'\hat{\Sigma}^{-1}\mu] - \frac{1}{2\gamma}\mathrm{E}[\hat{\mu}'\hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}\hat{\mu}] \\
&= \frac{1}{\gamma}\mathrm{E}[\hat{\mu}'\Sigma^{-\frac{1}{2}}W^{-1}\Sigma^{-\frac{1}{2}}\mu] - \frac{1}{2\gamma}\mathrm{E}[\hat{\mu}'\Sigma^{-\frac{1}{2}}W^{-2}\Sigma^{-\frac{1}{2}}\hat{\mu}] \\
&= \frac{1}{\gamma}\,\frac{M}{M-N-2}\mu'\Sigma^{-1}\mu - \frac{1}{2\gamma}\mathrm{E}[W^{-2}]\mathrm{E}[\hat{\mu}'\Sigma^{-1}\hat{\mu}].
\end{aligned}$$

Applying the independence, the first term is just the product of the individual expected values. In the second term we can pull out the expectation of W^{-2}. After that, using the knowledge of the distribution of the last expectation from 4.5, we result in:

$$\begin{aligned}
\mathrm{E}[U(\hat{x}_{mvo})] &= \frac{S_*^2}{\gamma}\frac{M}{M-N-2} - \frac{1}{2\gamma}\frac{M(M-1)(N+MS_*^2)}{(M-N-1)(M-N-2)(M-N-4)} \\
&= \frac{S_*^2}{2\gamma}k - \frac{NM(M-2)}{2\gamma(M-N-1)(M-N-2)(M-N-4)} \\
&= \frac{S_*^2}{2\gamma}k - \frac{1}{2\gamma}h.
\end{aligned}$$

In the end, we get the expected loss:

$$L(x^*, \hat{x}_{mvo}) = U(x^*) - \mathrm{E}[U(\hat{x}_{mvo})] = \frac{S_*^2}{2\gamma} - \frac{S_*^2}{2\gamma}k + \frac{1}{2\gamma}h = \frac{1}{2\gamma}[(1-k)S_*^2 + h].$$

(b) Now, we calculate the expected loss of the c-weighted $\frac{1}{n}$ portfolios. Following the proof by DeMiguel, Garlappi, and Uppal (2009) we formulate the $\frac{1}{n}$ rule generally as some c-weighted equal weights strategy:

$$x_{ew}(c,n) = c\mathbf{1}_n,\ c \in \mathbb{R}^+.$$

Whithout loss of generality let $\mathbf{1}_n$ be the vector where the first n entries are one and the last $N-n$ are qual to zero. Besides, when normalized this strategy is in fact a $\frac{1}{n}$ strategy:

$$w_{ew} = \frac{x_{ew}(c,n)}{|\mathbf{1}_n' x_{ew}(c,n)|} = \frac{c\mathbf{1}_n}{|\mathbf{1}_n' c\mathbf{1}_n|} = \frac{\mathbf{1}_n}{n} = \frac{1}{n}\mathbf{1}_n.$$

Now, the expected loss using the equal weight rule is:

$$\begin{aligned}
L(x^*, x_{ew}(c,n)) = U(x^*) - \mathrm{E}[U(x_{ew}(c,n))] &= U(x^*) - x_{ew}(c,n)'\mu + \frac{\gamma}{2}x_{ew}(c,n)'\Sigma x_{ew}(c,n) \\
&= U(x^*) - c\mathbf{1}_n'\mu + \frac{\gamma}{2}c^2\mathbf{1}_n'\Sigma\mathbf{1}_n.
\end{aligned}$$

We do not need the expectation in the second term because the $\frac{1}{n}$ rule does not use estimated values and therefore the result holds for all three cases. Then, the lowest possible loss is achieved, when minimizing the expected loss over $c \in \mathbb{R}^+$, which can be chosen freely. The expected loss is convex with respect to c such that we find the optimal solution for $\min_{c\in\mathbb{R}^+} L(x^*, x_{ew}(c,n))$ by the first order condition,

$$-\mathbf{1}_n'\mu + \gamma c\mathbf{1}_n'\Sigma\mathbf{1}_n = 0.$$

Rearranging terms leads to the optimal solution,

$$c^* = \frac{\mathbf{1}_n'\mu}{\gamma\mathbf{1}_n'\Sigma\mathbf{1}_n}.$$

Thus we get the lowest bound for applying the $\frac{1}{n}$ rule:

$$\begin{aligned}
\min_{c\in\mathbb{R}^+} L(x^*, x_{ew}(c,n)) &= \frac{S_*^2}{2\gamma} - \frac{\mathbf{1}_n'\mu}{\gamma\mathbf{1}_n'\Sigma\mathbf{1}_n}\mathbf{1}_n'\mu + \frac{\gamma}{2}\left(\frac{\mathbf{1}_n'\mu}{\gamma\mathbf{1}_n'\Sigma\mathbf{1}_n}\right)^2\mathbf{1}_n'\Sigma\mathbf{1}_n \\
&= \frac{S_*^2}{2\gamma} - \frac{(\mathbf{1}_n'\mu)^2}{\gamma\mathbf{1}_n'\Sigma\mathbf{1}_n} + \frac{(\mathbf{1}_n'\mu)^2}{2\gamma\mathbf{1}_n'\Sigma\mathbf{1}_n} \\
&= \frac{S_*^2}{2\gamma} - \frac{1}{\gamma}S_n^2 + \frac{1}{2\gamma}S_n^2 \\
&= \frac{1}{2\gamma}(S_*^2 - S_n^2).
\end{aligned}$$

(c) After that, we are ready to compare the expected losses in the three different cases. Remember, we want to know when the loss of the sample mean-variance rule is lower than the loss of the $\frac{1}{n}$ rule.

1. In the first case μ is known and Σ has to be estimated. We are searching for conditions under which the expected loss of the mean-variance strategy is smaller than the loss of the best c-weighted $\frac{1}{n}$ portfolio,

$$L(x^*, \hat{x}_{mvo}|\Sigma) < \min_{c \in \mathbb{R}^+} L(x^*, x_{ew}(c, n)).$$

Inserting the results from parts (a) and (b) this yields in the first case

$$\frac{N}{2\gamma M} < \frac{1}{2\gamma}(S_*^2 - S_n^2),$$

which leads to the desired result,

$$0 < S_*^2 - S_n^2 - \frac{N}{M}.$$

2. In the second case μ has to be estimated and Σ is known. Thus, we calculate the condition so that

$$L(x^*, \hat{x}_{mvo}|\mu) < \min_{c \in \mathbb{R}^+} L(x^*, x_{ew}(c, n)).$$

Using the results from the second case of part (a), and part (b) this holds if

$$(1-k)\frac{S_*^2}{2\gamma} < \frac{1}{2\gamma}(S_*^2 - S_n^2),$$

and leads to

$$0 < kS_*^2 - S_n^2.$$

3. In the third case both μ and Σ have to be estimated. There, we have

$$L(x^*, \hat{x}_{mvo}) < \min_{c \in \mathbb{R}^+} L(x^*, x_{ew}(c, n)).$$

And applying the results from part (a) and (b) yields

$$\frac{1}{2\gamma}[(1-k)S_*^2 + h] < \frac{1}{2\gamma}(S_*^2 - S_n^2),$$

resulting in

$$0 < kS_*^2 - S_n^2 - h.$$

□

In all three cases it is more likely for the sample mean-variance method to outperform the naive investment rule if the number of data M is high and the number of assets N is low. This can be seen by finding the minimum value M^* for which we have:

$$M^* = \inf\{M : L(x^*, \hat{x}_{mvo}) < \min_{c \in \mathbb{R}^+} L(x^*, x_{ew}(c, n))\},$$

i.e. for which minimum M is the sample mean-variance portfolio at least as good as the best c-weighted $\frac{1}{n}$ portfolio.

1. In the case where μ has to be estimated and Σ is known, we can get a closed form solution for M^*. From the condition under which the sample mean-variance outperforms the $\frac{1}{n}$ rule,
$$0 < S_*^2 - S_n^2 - \frac{N}{M},$$
we get directly
$$M > \frac{N}{S_*^2 - S_n^2}.$$
Therefore $M^* = \inf\{M : M > \frac{N}{S_*^2 - S_n^2}\}$, and the conclusion that high values for M and small values for N support the performance of the sample mean-variance strategy follows immediately.

2. Here, when μ is known and Σ has to be estimated, we cannot get a closed form solution for the minimum value M^*. However, we know that k is increasing in M and decreasing in N. From the condition in this case,
$$0 < kS_*^2 - S_n^2,$$
we see, that again the sample mean-variance strategy is more likely to outperform the $\frac{1}{n}$ rule if M is large and N is small, as
$$k > \frac{S_n^2}{S_*^2}.$$

3. In the third case, where both μ and Σ have to be estimated, we have
$$0 < kS_*^2 - S_n^2 - h,$$
which yields
$$0 < kS_*^2 - S_n^2 - h < kS_*^2 - S_n^2.$$
As $h > 0$, the last inequalities hold and this implies that as in the second case we need high M and low N for the sample mean-variance strategy to outperform the $\frac{1}{n}$ investment.

Then, we examine the impact of using fewer assets. First, we compare the full naive strategy with N assets against the portfolio with $n < N$ assets. As there is no estimation we are left with only one case. The question is, when does the $\frac{1}{n}$ strategy outperform the $\frac{1}{N}$ strategy? This is, when
$$\min_{c \in \mathbb{R}^+} L(x^*, x_{ew}(c, n)) < \min_{c \in \mathbb{R}^+} L(x^*, x_{ew}(c, N)),$$
which is equivalent to
$$\frac{1}{2\gamma}(S_*^2 - S_n^2) < \frac{1}{2\gamma}(S_*^2 - S_N^2)S_*^2 - S_n^2 < S_*^2 - S_N^2.$$
Rearranging terms yields
$$S_n^2 > S_N^2.$$

Indeed, it is possible for the $\frac{1}{n}$ strategy to outperform the $\frac{1}{N}$ strategy if the Sharpe ratio of the strategy with fewer assets is higher than the Sharpe ratio of the naive strategy with all assets. This is the case when it is possible, via selection, to either increase the mean or decrease the variance of the portfolio, or both. Thus, it is dependent on the actual available assets. It is more likely to find a $\frac{1}{n}$ strategy outperforming the $\frac{1}{N}$ strategy, if the available assets show a large variety. Moreover, the optimal n depends highly on the given assets.

However, this already demonstrates that choosing fewer than all assets for the naive strategy can have positive effects. If so, it is clear that the results from DeMiguel, Garlappi, and Uppal (2009) transfer to the naive strategy with fewer assets. We illustrate this by using the scenarios from the paper by DeMiguel, Garlappi, and Uppal (2009) extended by a second naive portfolio with fewer assets. Note, that we use the Sharpe ratios on the same time-scale as the observations in the theorem above. Thus, if we used annualized Sharpe ratios, we would have to scale with Δt such that $M\ \Delta t$ is the number of years of observation, which is needed. For example (4.2) would be replaced by

$$S_*^2 - S_n^2 - \frac{N}{M\Delta t} > 0,$$

and in (4.3) the Δt term would cancel out.

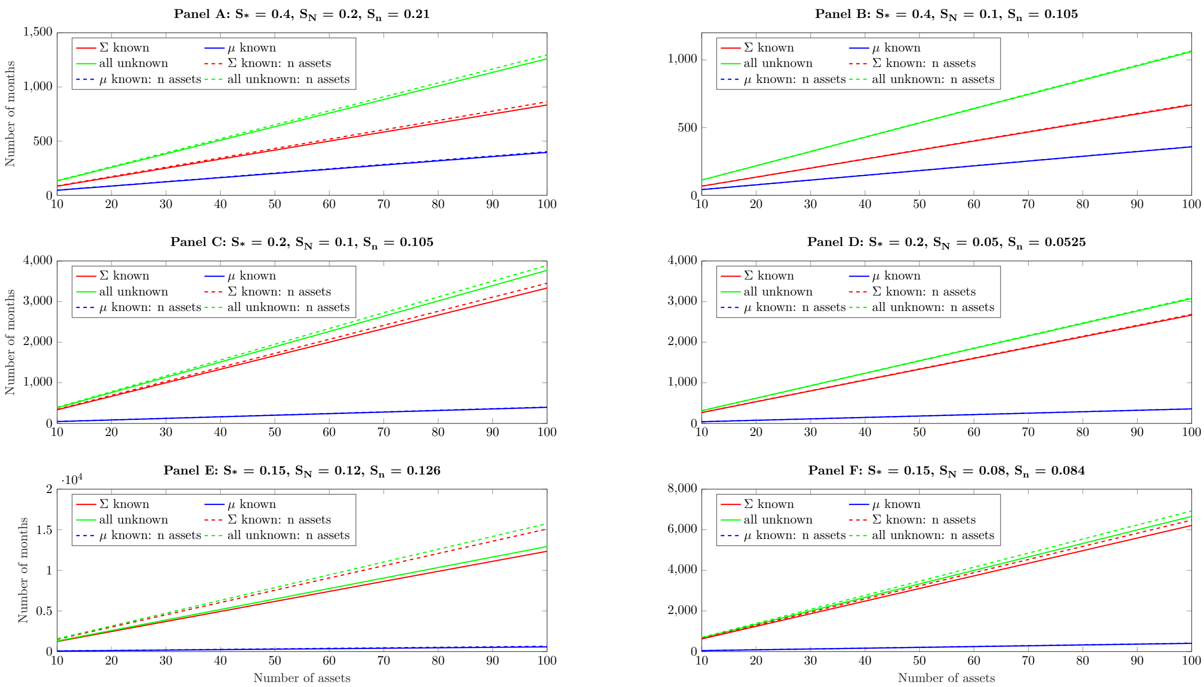


Figure 4.1: Estimation Period needed for the mean-variance portfolio to outperform the $\frac{1}{N}$ and $\frac{1}{n}$ portfolios

In Figure 4.1, we see the graphs from the paper by DeMiguel, Garlappi, and Uppal (2009) as solid lines in red, blue and green, corresponding to the cases 1, 2, and 3. Remember that the third case is the most realistic one, where neither the mean nor the covariance matrix are known. We see in all panels that the corresponding solid green line is always above the other solid lines. This tells us that the estimation error is very high and a higher amount of data is needed to overcome this disadvantage. As already pointed out in DeMiguel, Garlappi, and Uppal (2009), the estimation error coming from the sample mean is more influential than the error from the sample covariance. This can be seen by comparing the solid red line, which represents the case where μ has to be estimated and the solid blue line, where only the covariance matrix has to be estimated. The red line is always above the blue line, thus we always need more data to overcome the loss by estimating the mean than by estimating the covariance.

If we take a look at the sizes of the number of needed data, it is astonishing how much data is really needed. Here are some examples: In Panel A, where the Sharpe ratio of the equal weights strategy with all assets is only half of the Sharpe ratio of the optimal strategy, we are looking at the third case (green solid line). Here, for only ten assets, we already need 132 data points, which is 11 years of monthly data . For 50 available assets, it is already the amount of 633 data points, over 50 years of monthly data, and for 100 assets, we need 1260 data points, over 100 years of monthly data. In Panel B, where the Sharpe ratio of the full naive strategy is reduced to a fourth of the Sharpe ratio of the optimal strategy, the amount of needed data does not decrease substantially. For ten assets 112 data points are needed, for 50 assets 534 data points, and for 100 assets 1061 data points are needed. In the panels C and D we have the same ratio of the Sharpe ratios, $\frac{1}{2}$ for panels A and C and $\frac{1}{4}$ for panels B and D, but the magnitude of the Sharpe ratio becomes smaller. Thus, small errors have a higher impact on the result and therefore even more data is needed, see Table 4.1. The last two panels E and F are calibrated to data of the U.S. stock market and show the impact when the Sharpe ratios get even closer and thus the effect of estimation error makes the number of needed data again higher as in the cases before. For 50 assets we need 6470 data points in panel E, whereas in panel F, where the Sharpe ratio of the $\frac{1}{N}$ strategy is only 0.8, we still need 3331 data points.

	Panel A	Panel B	Panel C	Panel D	Panel E	Panel F
10 assets	132	112	383	313	1301	672
25 assets	320	270	947	773	3239	1669
50 assets	633	534	1888	1540	6470	3331
100 assets	1260	1061	3769	3074	12932	6655

Table 4.1: Estimation Period needed for the mean-variance portfolio to outperform the $\frac{1}{N}$ portfolio, for the case where neither μ nor Σ are known.

To compare these results to a $\frac{1}{n}$ portfolio with fewer assets, we assume that it is possible to gain an increase in Sharpe ratio of 5 percent from the full $\frac{1}{N}$ portfolio. As expected, the dashed lines for the $\frac{1}{n}$ portfolio lie above the other lines. However, it is striking again, that we observe a huge increase of needed data in realistic settings and only a small increase in the Sharpe ratio. The changes for the settings in Panel E can be seen in Table 4.2.

	10 assets	25 assets	50 assets	100 assets
Panel E - N assets	1301	3239	6470	12932
Panel E - n assets	1583	3943	7877	15746

Table 4.2: Estimation Period needed for the mean-variance portfolio to outperform the $\frac{1}{N}$ and the $\frac{1}{n}$ portfolio with fewer assets, which achieves an increase in Sharpe ratio of 5 percent, for the case where neither μ nor Σ are known. For available assets $N = 10, 25, 50$ and 100.

In the end, this supports that, even if the gains in Sharpe ratio are rather small, it is worth trying to find a naive strategy with fewer assets. To further support this idea we examine the $\frac{1}{n}$ strategy on real and simulated data.

4.4 Data Experiments

The following tables show the performance of the sample mean-variance portfolio, the full $\frac{1}{N}$ portfolio and the portfolio choosing the 24 historically best (by Sharpe ratio) assets in an equal weight portfolio. The rolling sample procedure from Section 3.4 is used to produce the portfolio returns. Since the sample mean-variance strategy can also produce negative weights, we must exclude cases, where the portfolio returns become smaller than -1, which results in negative portfolio wealth. Thus, the performance measures like Sharpe ratio or CEQ return are no longer adequate and cannot be used to compare the strategy to other portfolios.

For real data we will use two kinds of tables to compare the performance of the portfolios. One of them shows the averaged annual performance criteria for the whole period and for the last three five-year-periods. In addition, we use ranking tables, where the averaged annual performance criteria are compared and the portfolios are ranked according to the results of the comparison. All individual rankings are then combined to an overall ranking. The diversification-risk ratio can only be used for portfolios without shortselling, therefore this performance criterion is not included in the overall ranking, when at least one of the portfolios may include negative weights.

For simulated data we focus on the ranking tables, but this time the ranking compares the averaged annual performance criteria for the whole period and each simulation set. For each simulation set the portfolios are ranked and the overall ranking combines all these ranking results. The number in brackets right after the ranking number tells us how often the portfolio outperformed both other strategies for the respective performance criterion. The performance values for each simulation set and each portfolio are displayed in a figure. There are diagrams for each performance criterion, where the portfolios are characterized by different colors. Further details can be seen in cutouts containing the first hundred simulations, which are added in Appendix D.

4.4.1 Real Data

First, we use the data set provided by Kenneth French with 48 industry portfolios, see Appendix A for a detailed data description. Note that the performance in the tables below is calculated for the monthly data and then annualized if required. We checked the results for different estimation windows, namely 60, 120, 180, 240, 300 and 360 months. The sample mean-variance strategy only leads to reasonable results starting at 180 months estimation period.

In Table 4.3 we see the performance of the three portfolios as annualized data in the setting with estimation period of 180 months, which leaves 360 months for testing. The naive strategy with all assets is the benchmark we want to catch up with naive strategies with fewer assets. The first attempt is to use exactly half the number of assets, which is 24 assets in this example.

Comparing all performance criteria for the whole period the three portfolios can be ranked, see Table 4.4. The sample mean-variance takes always the last place and the full naive method nearly always takes the first place. Only in terms of standard deviation the naive method with 24 assets gets the first place. In total, this leads to the following ranking: the full naive at the first place, the naive portfolio with 24 assets at the second place and the sample mean-variance method takes the last place. This ranking is also valid for the monthly data given in Table 4.5.

The diversification-risk ratio is excluded from ranking as it does not deliver results for the sample mean-variance strategy, due to negative weights, which are not allowed for the diversification measure. However, we observe that the diversification-risk ratio of the full naive portfolio is always greater than the ratio for the naive portfolio with 24 assets.

Then, we take a look at the performance with increasing estimation periods, 240 (see Table 4.6), 300 (see Table 4.7) and 360 (see Table 4.9) months. As we are interested in rankings the annual tables are enough for the comparison. Since we use the same data set, the testing periods shift accordingly and result in fewer data. Thus, the average performances change, but the three displayed periods can be reported in all settings and compared directly.

The ranking for 240 and 300 months of estimation period (Table 4.8) coincides with the ranking for 180 months. When using 360 months for the estimation, there is a change in ranking for the standard deviation. Here, the full naive strategy performs slightly worse than the sample mean-variance optimization and gets the last place, but this does not affect the overall ranking. One can explain this change in ranking (Table 4.10) by examining the standard deviations of the last three five-years-periods. The naive portfolio is always a hint on the market behavior and the last periods suffer from quite high volatility compared to other periods. Thus, the naive portfolio shows lower average standard deviation when more periods are available. However, in the last case with 360 months estimation period, only the last periods are left for testing. Thus, the average performance of the naive portfolio cannot profit from other periods with possibly lower volatility. On the other hand the sample mean-variance method is likely to get good estimates and therefore can reduce the risk in terms of volatility. However, we see that this does not help for downside risk measures used in the Sortino or omega ratio.

Considering the diversification-risk ratio, we observe that the ranking does not change with increasing estimation period. The full naive portfolio always performs better than the naive portfolio with 24 assets.

In Table 4.9 we see that the mean-variance strategy no longer leads to very extreme results. However, the overall performance decreases heavily. Although more data is used to estimate the needed values, the outcome is even worse. This might be a hint that estimation over longer periods is surely statistically better, but the problem of forecasting with historical values becomes more of a problem. So, if there are changes in the market or in the distribution of some assets within shorter periods, then the use of more estimation data is lapsed. These effects can then only be examined with simulated data, where the underlying processes are known. The performance of the naive methods is still good, and this time highly superior to the mean-variance method. However, with the increase of the estimation window we begin to see some other effects, which arise from the smaller number of testing time periods. The performance of both naive methods begins to go down, due to less testing data. In the end, we see that pure enlargement of the estimation window does not produce higher Sharpe ratios in the case of the mean-variance portfolio. However, the naive strategies are not strongly affected by the change of settings. Even the strategy with 24 assets, which suffers like the mean-variance method from false estimated values or the problems that historical data does not need to be a good forecast, stays near the performance of the full naive portfolio.

Now we take a look at the portfolio turnovers of the different strategies. Remember, the portfolio turnover examines the change in weights. Following a special strategy is determined by the fixing of weights for each rebalancing period. However, as asset prices change the percentage of wealth invested the individual assets change as well. To follow a special strategy the investment must be updated such that the desired weights are preserved. For this, trading is required and the portfolio turnover defined above is the average of the amount of this trading, in terms of the difference in the weights. The best portfolio turnover is always achieved by the full naive strategy, closely followed by the naive portfolio with 24 assets. The mean-variance portfolio

180/360/annual	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq return	div-risk ratio
naive: 24 assets	14.89	15.72	94.90	152.10	206.60	13.89	13.66	106.56
1st period	5.59	14.69	38.37	59.07	131.98	74.12	4.51	165.11
2nd period	3.79	17.31	22.05	29.52	119.20	2.33	2.29	82.88
3rd period	21.23	15.31	139.89	262.46	281.66	1.21	20.06	85.92
naive: 48 assets	15.84	16.24	97.69	158.45	211.35	2.40	14.53	121.33
1st period	10.84	14.70	74.36	123.64	170.79	8.84	9.76	191.21
2nd period	5.27	18.87	28.17	39.93	126.18	1.16	3.49	85.52
3rd period	22.98	16.12	143.78	273.11	281.63	1.34	21.68	95.26
sample mvo	7.24	25.17	28.82	40.22	125.34	16.22	4.08	-
1st period	-22.60	30.61	-74.45	-86.56	56.25	14.04	-27.29	-
2nd period	-0.48	29.77	-1.62	-2.05	98.51	13.14	-4.91	-
3rd period	5.25	23.77	22.30	31.23	118.56	15.54	2.43	-

Table 4.3: Real data: Annual performance (in percent, except turnover) of the naive portfolio with 24 assets, the full naive portfolio and the sample mean-variance portfolio.
Estimation window: 180 months of daily data, test window: 12 months of monthly data, total testing period: 360 months

annual ranking	mean	std	Sharpe	Sortino	omega	turnover	ceq	overall	div-risk
naive: 24 assets	2	1	2	2	2	2	2	2	2
naive: 48 assets	1	2	1	1	1	1	1	1	1
sample mvo	3	3	3	3	3	3	3	3	-

Table 4.4: Real data: Ranking on averaged annual performance of the naive portfolio with 24 assets, the full naive portfolio and the sample mean-variance portfolio.
Estimation window: 180 months, test window: 12 months, total testing period: 360 months

180/360/monthly	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq return	div-risk ratio
naive: 24 assets	1.24	4.54	27.39	43.91	206.60	1.16	1.14	1278.70
1st period	0.47	4.24	11.08	17.05	131.98	6.18	0.38	1981.40
2nd period	0.32	5.00	6.37	8.52	119.20	0.19	0.19	994.55
3rd period	1.77	4.42	40.38	75.77	281.66	0.10	1.67	1031.10
naive: 48 assets	1.32	4.69	28.20	45.74	211.35	0.20	1.21	1456.00
1st period	0.90	4.24	21.47	35.69	170.79	0.74	0.81	2294.60
2nd period	0.44	5.45	8.13	11.53	126.18	0.10	0.29	1026.30
3rd period	1.92	4.65	41.51	78.84	281.63	0.11	1.81	1143.10
sample mvo	0.60	7.27	8.32	11.61	125.34	1.35	0.34	-
1st period	-1.88	8.84	-21.49	-24.99	56.25	1.17	-2.27	-
2nd period	-0.04	8.59	-0.47	-0.59	98.51	1.09	-0.41	-
3rd period	0.44	6.86	6.44	9.02	118.56	1.29	0.20	-

Table 4.5: Real data: Monthly performance (in percent, except turnover) of the naive portfolio with 24 assets, the full naive portfolio and the sample mean-variance portfolio.
Estimation window: 180 months of daily data, test window: 12 months of monthly data, total testing period: 360 months

240/300/annual	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq return	div-risk ratio
naive: 24 assets	13.06	14.39	90.95	145.70	197.68	4.98	12.03	118.40
1st period	6.19	13.88	44.94	68.20	138.26	17.84	5.22	154.92
2nd period	3.90	16.38	24.00	32.53	121.26	3.42	2.56	93.12
3rd period	20.48	14.49	142.51	263.79	283.91	1.22	19.43	88.07
naive: 48 assets	14.64	15.70	93.40	152.82	202.04	2.65	13.41	129.69
1st period	10.84	14.70	74.36	123.64	170.79	8.84	9.76	191.21
2nd period	5.27	18.87	28.17	39.93	126.18	1.16	3.49	85.52
3rd period	22.98	16.12	143.78	273.11	281.63	1.34	21.68	95.26
sample mvo	5.85	22.34	26.23	36.07	123.35	33.47	3.35	-
1st period	-9.62	21.94	-44.21	-55.48	72.09	26.30	-12.03	-
2nd period	6.52	29.49	22.29	28.10	123.18	89.70	2.17	-
3rd period	1.33	19.76	6.79	8.96	105.35	13.12	-0.62	-

Table 4.6: Real data: Annual performance (in percent, except turnover) of the portfolios over 300 months. Estimation window: 240 months of daily data, test window: 12 months of monthly data, total testing period: 300 months

300/240/annual	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq return	div-risk ratio
naive: 24 assets	13.30	14.53	91.74	144.39	198.73	4.37	12.25	119.41
1st period	7.91	14.19	56.21	90.08	150.48	13.85	6.90	159.96
2nd period	2.39	16.27	14.79	19.54	112.81	1.10	1.06	90.56
3rd period	20.95	14.10	149.81	283.74	298.67	1.14	19.95	87.49
naive: 48 assets	15.27	16.10	95.00	153.83	204.11	3.09	13.97	126.10
1st period	10.84	14.70	74.36	123.64	170.79	8.84	9.76	191.21
2nd period	5.27	18.87	28.17	39.93	126.18	1.16	3.49	85.52
3rd period	22.98	16.12	143.78	273.11	281.63	1.34	21.68	95.26
sample mvo	6.81	18.84	36.20	53.17	132.19	21.71	5.03	-
1st period	-2.34	18.99	-12.44	-18.45	91.35	37.94	-4.15	-
2nd period	-0.35	21.62	-1.62	-2.01	98.50	25.78	-2.68	-
3rd period	6.99	16.74	42.15	63.13	135.69	17.89	5.59	-

Table 4.7: Real data: Annual performance (in percent, except turnover) of the portfolios over 240 months. Estimation window: 300 months of daily data, test window: 12 months of monthly data, total testing period: 240 months

annual ranking	mean	std	Sharpe	Sortino	omega	turnover	ceq	overall	div-risk
naive: 24 assets	2	1	2	2	2	2	2	2	2
naive: 48 assets	1	2	1	1	1	1	1	1	1
sample mvo	3	3	3	3	3	3	3	3	-

Table 4.8: Real data: Ranking on averaged annual performance of the naive portfolio with 24 assets, the full naive portfolio and the sample mean-variance portfolio.
Estimation window: 240/300 months, test window: 12 months, total testing period: 300/240 months

usually shows very high turnover. Also, the naive portfolio with fewer assets has periods with high turnover values, especially in the setting with 180 months estimation period. Remember that this method chooses the best 24 assets by Sharpe ratio, therefore the statistical robustness plays a role as well as the forecast problem. Moreover, when the assets only differ slightly the ordering might change and the chosen assets may be totally replaced in the following period, which causes high portfolio turnover.

When looking at the CEQ return, we see that the sample mean-variance strategy is the least favorable. The difference between the two naive strategies is not too large, but still the naive method with all assets is more favorable. As we look at the omega ratio, we see again that the naive strategies outperform the sample mean-variance strategy.

Performance criteria alone do not tell us how the strategies would have evolved over time. Therefore, we will examine the process of the portfolio values (wealth) in different settings, supposing the initial wealth had been 100 for each strategy, see figures 4.2, 4.3, 4.4, 4.5, and 4.6. We see that both naive strategies stay close together. Although the naive strategy with 24 assets invests only in half the number of assets this strategy is still able to follow the market movement. On the one hand, taking fewer assets is a chance to exclude bad evolving assets, but on the other hand, it is not so robust to changes in asset behavior. Therefore, sometimes it performs better than the full naive strategy and sometimes worse. However, over time both strategies are quite similar.

So far we only chose assets by their historical Sharpe ratio, and we know that the Sharpe ratio is not always a good forecast, even if the estimation error is small. If we have additional information the selection of assets can be improved, see Chapter 5.

The sample mean-variance strategy performs poorly compared to the naive portfolios with respect to long terms and to long estimation periods. For the first months it seems to perform well but then there might be some change in the market, which cannot be well estimated in the mean-variance setting and the returns go down. In contrast, the naive strategies are less affected and seem to have a higher trend upwards.

In Table 4.11 the performance of different sized equal weight portfolios is shown. The performance values are the annual averages over the whole period of 360 months. In this example, we see that there is no natural best number of assets in an equal weights portfolio. As a matter of fact, taking 21 or 27 assets would increase the performance of the naive portfolio. This is strongly dependent on the given data. Therefore, taking half the number of assets is reasonable. In the following chapters we will be able to improve the selection of the 24 assets and thus improve the performance. When mean and variance are the only characteristics, then it is reasonable to take all assets, as the mean, the Sharpe ratio and the CEQ return rise with more assets, independently of the rising standard deviation towards the end.

This clear rising of the mean cannot be generally true. Here, we are given assets, which are industry portfolios, meaning they are already a combination of assets and therefore they perform quite well. In a different setting with e.g. stocks available in one country, there are always stocks which perform badly at least for some period. Including these assets would not lead to a rise of mean. For Sortino and omega ratio, which take the difference of upside and downside deviation into account, the highest values are achieved for the full naive portfolio. However, in between, there is always an up and down of performance. Thus, the constraint of taking fewer assets does not lead to a preferable number of assets.

The portfolio turnover on the other hand does not hint at any systematic choice of assets. Unfortunately, the case with half the number of assets leads in this example to a high portfolio

360/180/annual	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq return	div-risk ratio
naive: 24 assets	10.46	14.60	71.80	111.09	170.81	7.25	9.39	115.47
1st period	6.93	14.29	48.94	76.92	142.17	18.83	5.91	161.01
2nd period	2.77	15.85	17.65	23.46	115.33	0.99	1.52	88.26
3rd period	21.66	13.24	165.06	323.73	321.75	0.61	20.79	85.20
naive: 48 assets	13.03	16.70	78.26	125.80	181.28	3.95	11.64	126.16
1st period	10.84	14.70	74.36	123.64	170.79	8.84	9.76	191.21
2nd period	5.27	18.87	28.17	39.93	126.18	1.15	3.49	85.52
3rd period	22.98	16.12	143.78	273.11	281.63	1.34	21.68	95.26
sample mvo	5.08	16.58	30.73	45.64	125.67	18.44	3.71	-
1st period	0.20	19.80	1.01	1.46	100.73	16.31	-1.76	-
2nd period	3.13	16.40	19.26	27.25	115.80	26.10	1.79	-
3rd period	11.91	12.89	93.16	159.21	194.56	11.53	11.08	-

Table 4.9: Real data: Annual performance (in percent, except turnover) of the portfolios over 180 months. estimation window: 360 months of daily data, test window: 12 months of monthly data, total testing period: 180 months

annual ranking	mean	std	Sharpe	Sortino	omega	turnover	ceq	overall	div-risk
naive: 24 assets	2	1	2	2	2	2	2	2	2
naive: 48 assets	1	3	1	1	1	1	1	1	1
sample mvo	3	2	3	3	3	3	3	3	-

Table 4.10: Real data: Ranking on averaged annual performance of the naive portfolio with 24 assets, the full naive portfolio and the sample mean-variance portfolio.
Estimation window: 360 months, test window: 12 months, total testing period: 180 months

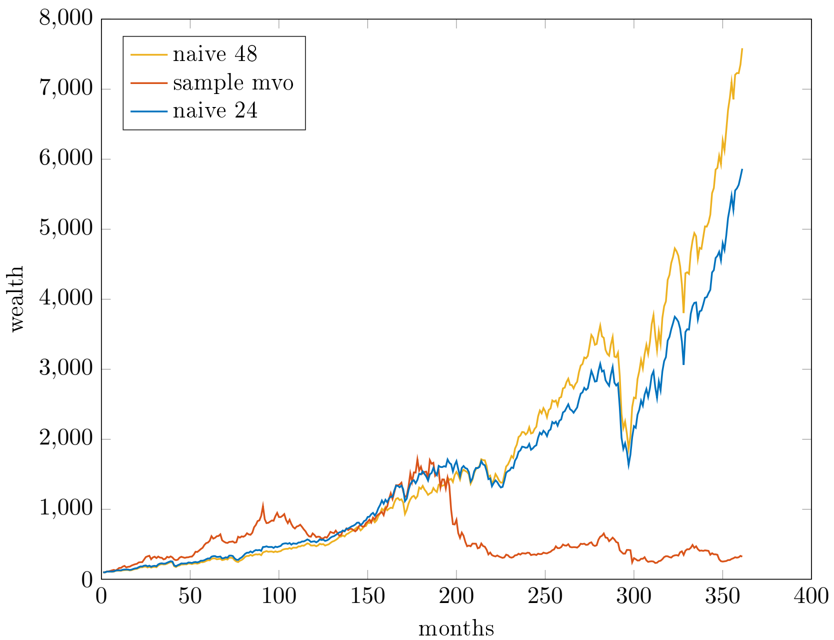


Figure 4.2: Real data: Portfolio prices starting at 100 for the naive strategies with 24 and 48 assets and the sample mean-variance portfolio. Estimation period: 180 months

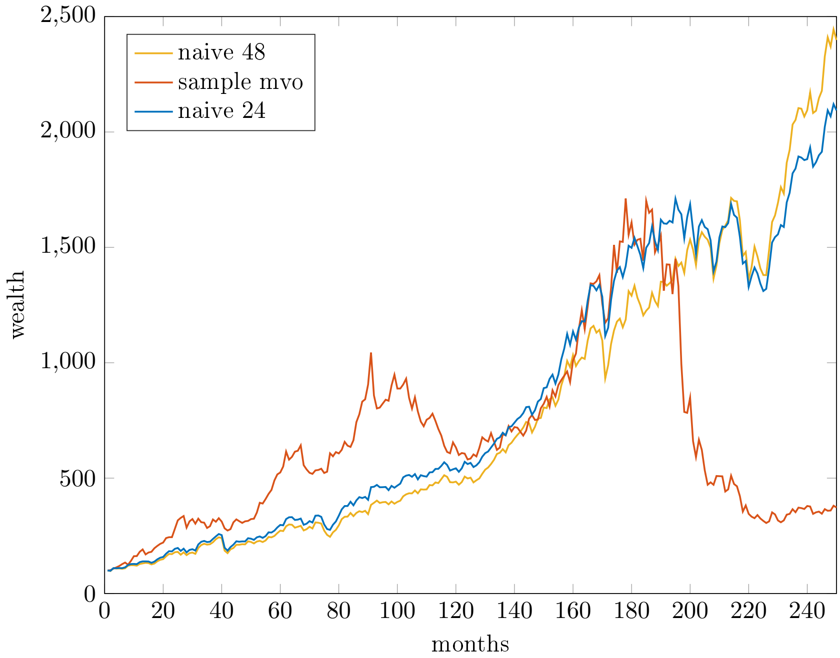


Figure 4.3: Real data: Portfolio prices starting at 100 for the naive strategies with 24 and 48 assets and the sample mean-variance portfolio. Estimation period: 180 months, Part 1 (1 to 250 months)

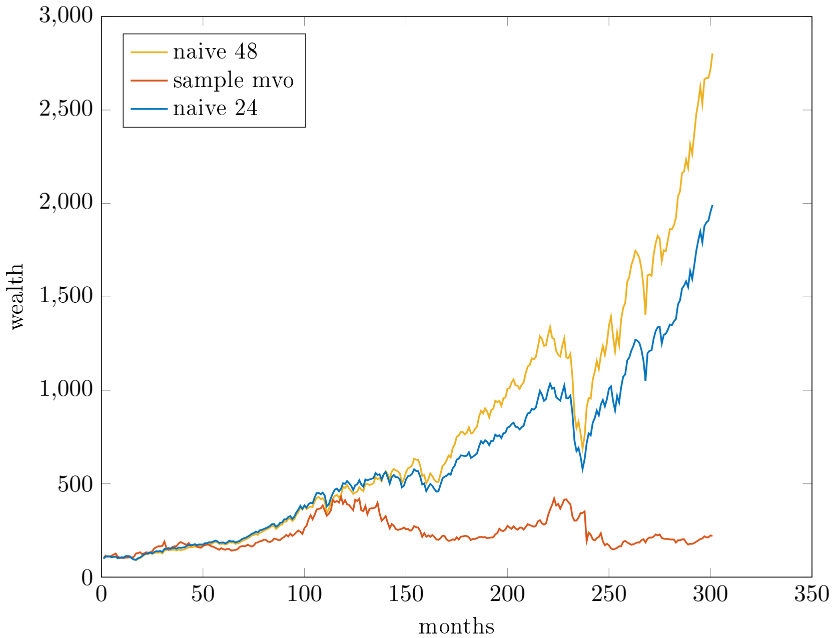


Figure 4.4: Real data: Portfolio prices starting at 100 for the naive strategies with 24 and 48 assets and the sample mean-variance portfolio. Estimation period: 240 months

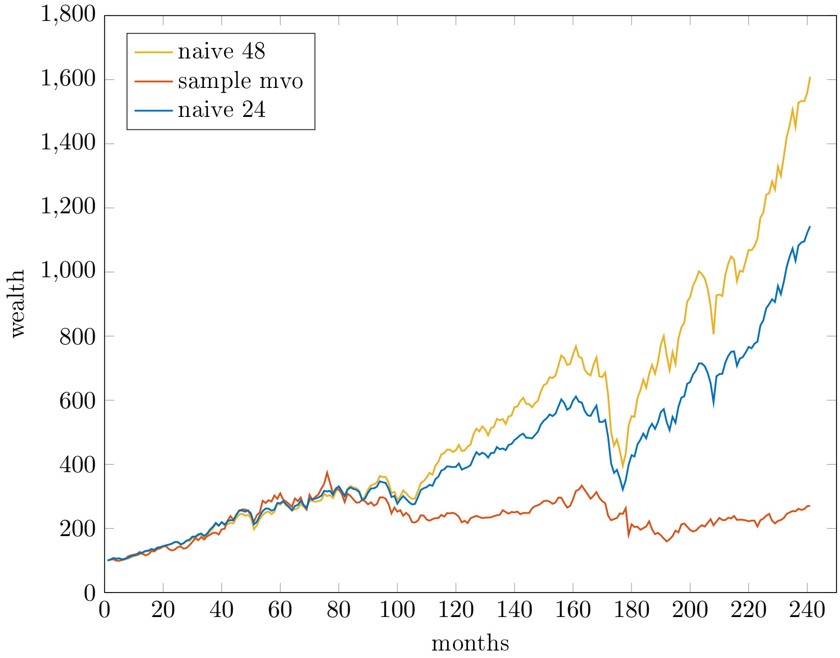


Figure 4.5: Real data: Portfolio prices starting at 100 for the naive strategies with 24 and 48 assets and the sample mean-variance portfolio. Estimation period: 300 months

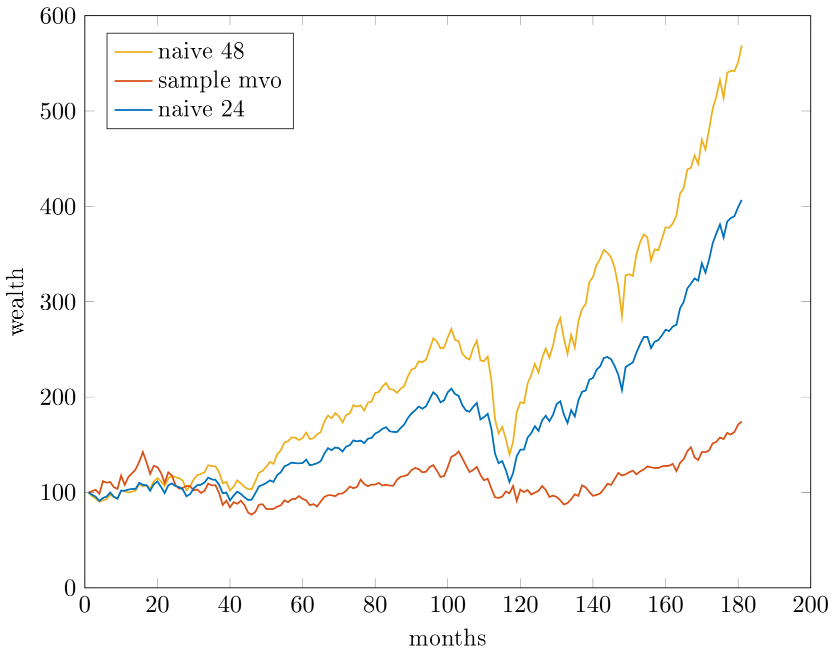


Figure 4.6: Real data: Portfolio prices starting at 100 for the naive strategies with 24 and 48 assets and the sample mean-variance portfolio. Estimation period: 360 months

turnover, where in general the turnover is rather low. This effect results from the way the assets are selected. Here, the assets with the best historical Sharpe ratios are chosen, and if the Sharpe ratios are close, minor changes lead to choosing different assets, which produces a high turnover.

If we look at the composition of portfolios over time, it can be seen that it changes a lot for 9 and 24 assets, whereas it stays quite stable for 18 assets. We get the best diversification-risk ratio for the full naive strategy in this setting. However, we have already seen that this is not the case when longer estimation periods are used. One can see that in general, more assets are better than fewer.

However, taking only more assets does not always improve the performance. This goes alongside Markowitz (1952): “The adequacy of diversification is not thought by investors to depend solely on the number of different securities held.”

180/360/annual	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq return	div-risk ratio
3 assets	11.6970	17.8455	65.6370	103.9759	167.0766	1.5670	10.1047	73.55
6 assets	12.0694	16.2730	74.2711	115.4507	178.2073	6.0228	10.7453	97.09
9 assets	13.3744	15.6386	85.6408	135.3175	193.3628	19.8496	12.1516	108.83
12 assets	14.1519	15.2384	92.9993	149.7657	202.9045	2.5425	12.9908	110.57
15 assets	14.5187	15.5847	93.2896	150.4579	203.3794	2.5859	13.3043	110.67
18 assets	14.6727	15.6368	93.9645	150.1235	204.6212	1.8629	13.4501	109.47
21 assets	14.8975	15.6215	95.4980	153.0550	207.1423	9.1514	13.6773	109.25
24 assets	14.8926	15.7150	94.8987	152.1029	206.5952	13.8897	13.6578	106.56
27 assets	15.0356	15.8102	95.2330	153.4384	206.9269	3.6023	13.7858	107.67
30 assets	15.0475	15.8421	95.1164	152.9090	207.0731	3.0359	13.7926	108.40
33 assets	15.2439	15.8872	96.0843	154.6791	207.6140	2.7568	13.9819	109.54
36 assets	15.5175	16.0398	96.8787	156.2191	209.6243	3.1675	14.2312	112.22
39 assets	15.4529	16.1102	96.0534	154.4230	208.1497	5.8072	14.1552	112.71
42 assets	15.5557	16.1037	96.7316	155.5978	209.2794	8.7308	14.2591	110.23
45 assets	15.7508	16.2280	97.1946	157.1082	210.3042	2.8397	14.4341	117.78
48 assets	15.8448	16.2426	97.6869	158.4524	211.3500	2.4040	14.5257	121.33

Table 4.11: Real data: Annual performance (in percent, except turnover) over the whole period of the naive portfolios with different sizes.
Estimation window: 180 months of daily data, test window: 12 months of monthly data, total testing period: 360 months

4.4.2 Simulated Data

Nevertheless, all findings above must be checked empirically with simulated data. Losing the characteristics of real data, we are able to simulate different settings and can conclude how data characteristics affect the portfolio strategies. Exploring real data for these characteristics may help to choose the right portfolio.

We start with three models considering independent assets. These are not only the easiest models, but we also want to exclude the effects of covariance, at the moment. Dependence structures will be examined in Chapter 5. This aims at checking robustness with misspecified models. We include moderate model uncertainty in Model 2 and 3 with switching parameters.

For 1000 simulation sets we count the number of success of each strategy compared to the others. Success means the portfolio was able to outperform the other strategies in a performance measure. This number is given in brackets behind the ranking positions. As before, we have to exclude cases, where the mean-variance strategy produces negative portfolio returns, thus also the number of non-degenerate simulations is reported in brackets after the word "ranking".

Simulation Model 1
Let us start with the most simplified model where the assets are independently normally distributed with mean and variance not changing over time. See Appendix B.1 for a detailed description of Model 1. As we have seen in Section 4.3, 60 months of daily data is not sufficient to get good moments estimates. Still, the performance of the sample mean-variance portfolio is good and results in the first place in the overall ranking. This is mainly due to the static simulation, where parameters do not change over time. In addition, there are several assets with negative mean which can be avoided by the mean-variance strategy. In contrast to this, the full naive portfolio is bound to take every asset, which decreases its performance. And this is the reason for the high portfolio turnover of the naive strategy, which is always greater than the turnover of the naive portfolio with fewer assets. Apart from the volatility, the full naive method is listed in third place.

For 60 months estimation period in Table 4.12, it is obvious that the mean, the Sharpe ratio, the Sortino ratio, the omega ratio and the CEQ return of the naive portfolio with 24 assets are always greater than the mean, the Sharpe ratio, the Sortino ratio, the omega ratio and the CEQ return of the full naive portfolio. However, the volatility of the full naive strategy is always smaller than the volatility of the naive method with only 24 assets and the sample mean-variance portfolio. The reason why we cannot make such general statements on the performance of the sample mean-variance strategy is, that it still performs extremely for some simulation sets. As in the exemplary diagram for the 180 months case (Figure 4.7), there are also peaks in performance in the 60 months case. From these figures, we see that usually the mean, the Sharpe ratio, the Sortino ratio, the omega ratio and the CEQ return are greater than the corresponding values of both naive strategies. The turnover of the sample mean-variance strategy is usually smaller than the turnover of the naive strategy but greater than the turnover of the naive strategy with 24 assets. However, the volatility is usually greater than the volatility of both naive strategies. That in this case the diversification-risk ratio of the full naive strategy is greater than the one of the naive strategy with 24 assets in more than two-thirds of the cases, is not unexpected. As the assets are independently simulated it usually has the higher diversification measure and the Ω-risk is also often smaller. With increasing estimation periods the ranking does not change and the general statements from above remain valid, except for the volatility of the sample mean-variance strategy, which is often greater than the volatility of both naive strategies. However, there are also extreme losses in some cases. As anticipated the number of usable data sets increases with better estimation. Although the absolute values of the winning values in brackets increase, the percentage more or less stays the same. Only the percentage of the diversification-risk ratio for the full naive strategy decreases from 72 to 63 percent.

Simulation Model 2
Next, we examine the performance for Simulation Model 2, which considers independent assets with switching parameters (each year), see Section B.2 for details. Here, we get a totally different ranking, where the full naive method and the sample mean-variance change places. The increase in parameter uncertainty clearly favors the naive strategy. This matches with what Pflug, Pichler, and Wozabal (2012) found in their paper *'The 1/N investment strategy is optimal under high model ambiguity'*. Still, the naive method with 24 assets stays in the middle of the other strategies and is thus robust, if we do not know how the market and its assets behave.

Corresponding to the 60 months estimation period case in Table 4.13, we can make several

Estimation window: 60 months, testing window: 12 months, total testing period: 480 months. 846 simulation sets.

ranking (846)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	overall	div-risk
naive: 24 assets	2 (1)	2 (0)	2 (147)	2 (32)	2 (11)	1 (841)	2 (4)	2	2 (227)
naive: 48 assets	3 (0)	1 (846)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3	1 (619)
plug-in mvo	1 (845)	3 (0)	1 (699)	1 (814)	1 (835)	2 (5)	1 (842)	1	-

Estimation window: 120 months, testing window: 12 months, total testing period: 420 months. 933 simulation sets.

ranking (933)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	overall	div-risk
naive: 24 assets	2 (5)	2 (0)	2 (64)	2 (14)	2 (6)	1 (926)	2 (6)	2	2 (288)
naive: 48 assets	3 (0)	1 (932)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3	1 (645)
plug-in mvo	1 (928)	3 (1)	1 (869)	1 (919)	1 (927)	2 (7)	1 (927)	1	-

Estimation window: 180 months, testing window: 12 months, total testing period: 360 months. 952 simulation sets.

ranking (952)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	overall	div-risk
naive: 24 assets	2 (7)	2 (0)	2 (23)	2 (12)	2 (7)	1 (946)	2 (8)	2	2 (320)
naive: 48 assets	3 (0)	1 (951)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3	1 (632)
plug-in mvo	1 (945)	3 (1)	1 (929)	1 (940)	1 (945)	2 (6)	1 (944)	1	-

Estimation window: 240 months, testing window: 12 months, total testing period: 300 months. 964 simulation sets.

ranking (964)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	overall	div-risk
naive: 24 assets	2 (8)	2 (0)	2 (19)	2 (9)	2 (9)	1 (958)	2 (10)	2	2 (334)
naive: 48 assets	3 (0)	1 (963)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3	1 (630)
plug-in mvo	1 (956)	3 (1)	1 (945)	1 (955)	1 (955)	2 (6)	1 (954)	1	-

Estimation window: 300 months, testing window: 12 months, total testing period: 240 months. 968 simulation sets.

ranking (968)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	overall	div-risk
naive: 24 assets	2 (8)	2 (0)	2 (13)	2 (11)	2 (8)	1 (961)	2 (9)	2	2 (347)
naive: 48 assets	3 (0)	1 (967)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3	1 (621)
plug-in mvo	1 (960)	3 (1)	1 (955)	1 (957)	1 (960)	2 (7)	1 (959)	1	-

Estimation window: 360 months, testing window: 12 months, total testing period: 180 months. 974 simulation sets.

ranking (974)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	overall	div-risk
naive: 24 assets	2 (9)	2 (0)	2 (13)	2 (17)	2 (15)	1 (964)	2 (10)	2	2 (365)
naive: 48 assets	3 (0)	1 (973)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3	1 (609)
plug-in mvo	1 (965)	3 (1)	1 (961)	1 (957)	1 (959)	2 (10)	1 (964)	1	-

Table 4.12: Model 1: Annual ranking. Comparison of the naive portfolio with 24 assets, the full naive portfolio and the unconstrained mean-variance portfolio.
Settings: Simulation model 1 with 48 independent assets with static parameters. Estimation window: 60 - 360 months, testing window: 12 months, total testing period: 480 - 180 months.
The number in brackets after the ranking indicates how often the portfolio outperformed both other strategies in terms of this performance criterion.

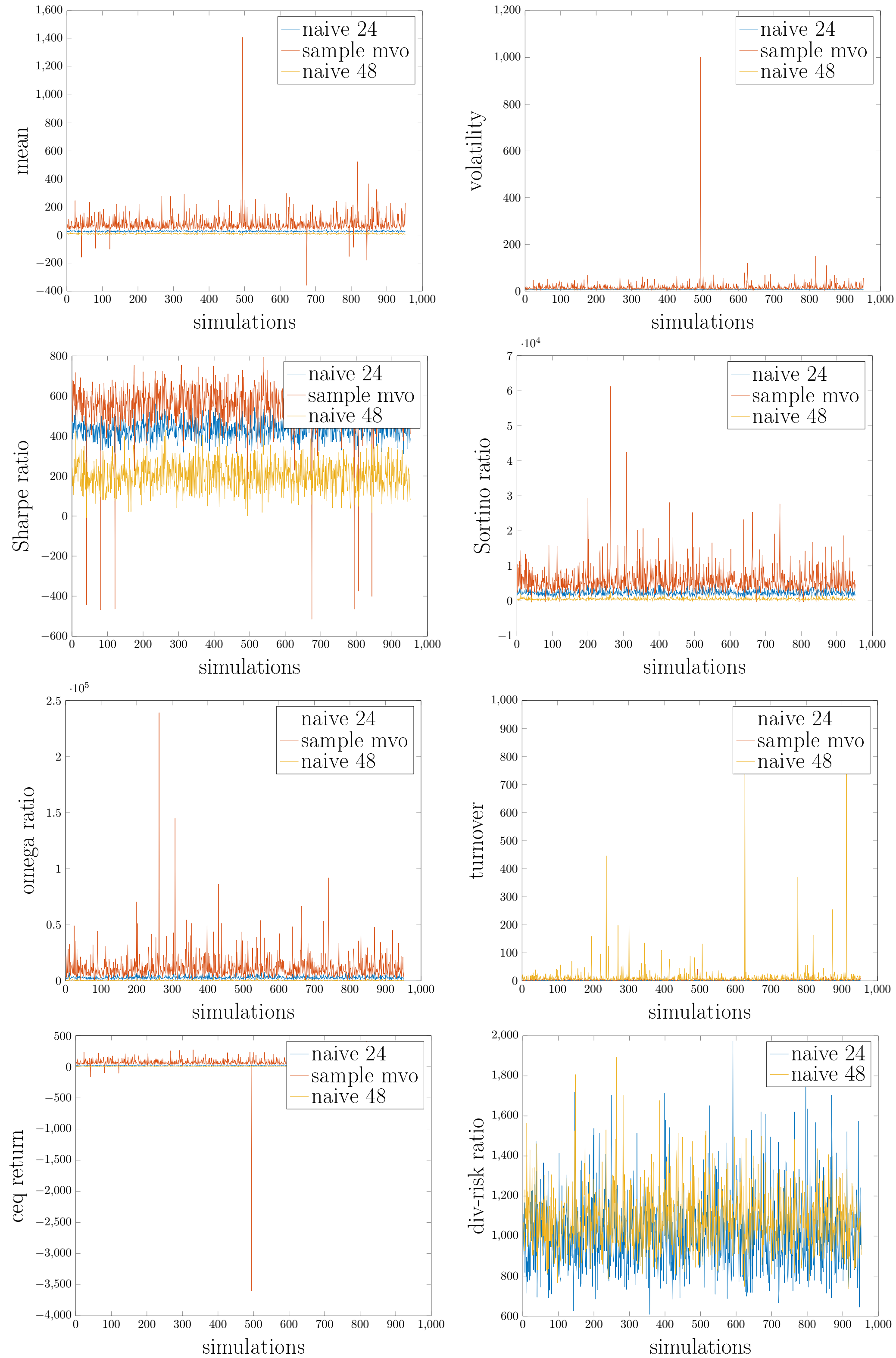


Figure 4.7: Model 1: Annual performance (in percent, except turnover). Simulation model 1, 48 assets, 180 months estimation window, 360 months total testing period. 952 simulation sets. (For more details, the corresponding graphs with only the first 100 simulations can be found in Appendix D)

Estimation window: 60 months, testing window: 12 months, total testing period: 480 months. 864 simulation sets.

ranking (864)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	overall	div-risk
naive: 24 assets	2 (128)	2 (0)	2 (0)	2 (0)	2 (0)	2 (259)	1 (204)	2	2 (0)
naive: 48 assets	3 (320)	1 (846)	1 (864)	1 (864)	1 (864)	1 (596)	3 (398)	1	1 (864)
plug-in mvo	1 (416)	3 (0)	3 (0)	3 (0)	3 (0)	3 (9)	2 (262)	3	-

Estimation window: 120 months, testing window: 12 months, total testing period: 420 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	overall	div-risk
naive: 24 assets	1 (202)	2 (0)	2 (0)	2 (0)	2 (0)	2 (298)	1 (231)	2	2 (0)
naive: 48 assets	2 (385)	1 (1000)	1 (1000)	1 (1000)	1 (1000)	1 (677)	3 (461)	1	1 (1000)
plug-in mvo	3 (413)	3 (0)	3 (0)	3 (0)	3 (0)	3 (25)	2 (308)	3	-

Estimation window: 180 months, testing window: 12 months, total testing period: 360 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	overall	div-risk
naive: 24 assets	2 (210)	2 (0)	2 (0)	2 (0)	2 (0)	2 (295)	1 (227)	2	2 (0)
naive: 48 assets	3 (370)	1 (1000)	1 (1000)	1 (1000)	1 (1000)	1 (660)	3 (431)	1	1 (1000)
plug-in mvo	1 (420)	3 (0)	3 (0)	3 (0)	3 (0)	3 (45)	2 (342)	3	-

Estimation window: 240 months, testing window: 12 months, total testing period: 300 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	overall	div-risk
naive: 24 assets	2 (228)	2 (0)	2 (0)	2 (2)	2 (0)	2 (311)	1 (223)	2	2 (0)
naive: 48 assets	1 (405)	1 (1000)	1 (1000)	1 (998)	1 (1000)	1 (626)	3 (453)	1	1 (1000)
plug-in mvo	3 (367)	3 (0)	3 (0)	3 (0)	3 (0)	3 (63)	2 (324)	3	-

Estimation window: 300 months, testing window: 12 months, total testing period: 240 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	overall	div-risk
naive: 24 assets	3 (253)	2 (0)	2 (1)	2 (1)	2 (2)	2 (314)	1 (245)	2	2 (0)
naive: 48 assets	1 (406)	1 (1000)	1 (999)	1 (998)	1 (998)	1 (616)	3 (441)	1	1 (1000)
plug-in mvo	2 (341)	3 (0)	3 (0)	3 (1)	3 (0)	3 (74)	2 (314)	3	-

Estimation window: 360 months, testing window: 12 months, total testing period: 180 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	overall	div-risk
naive: 24 assets	2 (272)	2 (0)	2 (2)	2 (6)	2 (3)	2 (276)	1 (267)	2	2 (0)
naive: 48 assets	3 (399)	1 (1000)	1 (997)	1 (993)	1 (996)	1 (614)	3 (426)	1	1 (1000)
plug-in mvo	1 (329)	3 (0)	3 (1)	3 (1)	3 (1)	3 (110)	2 (307)	3	-

Table 4.13: Model 2: Annual ranking. Comparison of the naive portfolio with 24 assets, the full naive portfolio and the unconstrained mean-variance portfolio.
Settings: Simulation model 2 with 48 independent assets with switching parameters. Estimation window: 60 - 360 months, testing window: 12 months, total testing period: 480 - 180 months.
The number in brackets after the ranking indicates how often the portfolio outperformed both other strategies in terms of this performance criterion.

general statements: The mean-variance volatility is always greater than the volatility of both naive strategies. The volatility of the naive portfolio with half the number of assets is always greater than the naive volatility. Moreover, it has smaller values considering the Sharpe, Sortino and omega ratio. However, the naive method with 24 assets gets the first place in the CEQ return ranking. It is not more often on the first place than the other strategies, but it takes the second place more often, than the others and in total this leads to the first place. Whereas the full naive strategy outperforms the naive portfolio with 24 assets in terms of the volatility, the Sharpe and Sortino ratio as well as the diversification risk ratio.

The graphs in this case are similar to the ones in Figure 4.8. Since the mean of the sample mean-variance method was usually greater in Model 1, we see here that its performance is widely spread around the naive means. It is no longer possible to correctly evaluate the assets, thus the outcome is a matter of chance. Therefore, the volatility also rises dramatically. For the Sharpe, Sortino and omega ratio we observe a clear distinction between the portfolios, which is even more striking in the 60 months case than in the 180 months case. The portfolio turnover of the naive strategy is usually smaller than the turnover of the other strategies. Moreover, the turnover of the sample mean-variance method is usually greater than the turnover of both naive strategies. The CEQ return is more or less the result of the mean and the volatility of the portfolios and therefore leads to a similar diagram as the mean. With increasing estimation periods the mean-variance strategy is able to improve its performance. The Markov chain, driving the parameters in the simulation, has only a finite and small state space, therefore the evaluation of assets becomes better with longer estimation periods. This means that the values of the sample mean-variance portfolio approach the values of the naive strategies. Thus, the general statements vanish and are replaced by usual behavior. Only the mean sometimes changes the ranking. It is astonishing that the sample mean-variance portfolio has the best mean performance with lowest estimation period, namely in nearly half the cases. Only the diversification-risk ratio stays the same independent of estimation periods. The naive strategy always outperforms the naive with 24 assets. An example for the performance values in Model 2 can be seen in Figure 4.8 with all performance criteria for the 180 months estimation period.

Simulation Model 3

Finally, we examine the performance in Model 3, considering independent assets with switching market dependent parameters (each year), see Appendix B.3 for details. Although the parameters change over time, they only change within a certain range representing the good, bad and normal state of the market. Therefore, the moment estimation captures the assets behavior quite good and the mean-variance strategy should still work quite well.

This is reflected in the rankings as the sample mean-variance strategy again gets the first place in the overall ranking. Only in terms of volatility and turnover it is on the third and second place. The naive method gets the last place except for the volatility. Nevertheless, the naive strategy with 24 assets defends its second place and improves the turnover for the first place.

In the case of the 60 months estimation period in Table 4.14, the mean of the sample mean-variance strategy is always greater than the mean of the naive strategy, which is always greater than the mean of the full naive strategy. The standard deviation of the naive portfolio is always smaller than the standard deviation of the other strategies. The Sharpe, Sortino and omega ratio and the CEQ return of the naive portfolio with 24 assets is always greater than the values for the full naive portfolio. Only the omega ratio of the sample mean-variance strategy is always greater

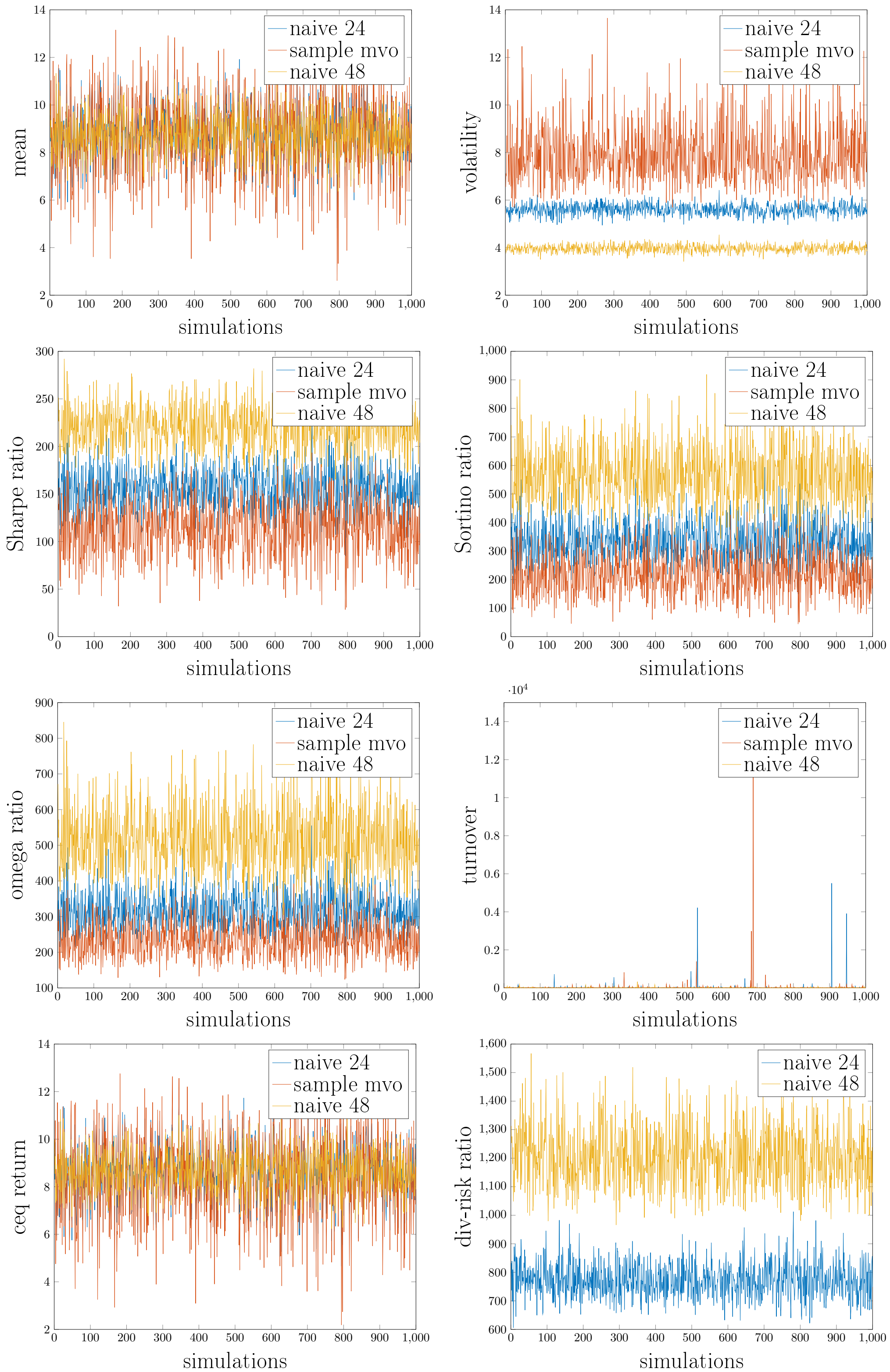


Figure 4.8: Model 2: Annual performance (in percent, except turnover). Simulation model 2, 48 assets, 180 months estimation window, 360 months total testing period. 1000 simulation sets. (For more details, the corresponding graphs with only the first 100 simulations can be found in Appendix D)

Estimation window: 60 months, testing window: 12 months, total testing period: 480 months. 828 simulation sets.

ranking (828)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	overall	div-risk
naive: 24 assets	2 (0)	2 (0)	2 (173)	2 (78)	2 (29)	1 (817)	2 (6)	2	2 (174)
naive: 48 assets	3 (0)	1 (828)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3	1 (654)
plug-in mvo	1 (828)	3 (0)	1 (655)	1 (750)	1 (799)	2 (11)	1 (822)	1	-

Estimation window: 120 months, testing window: 12 months, total testing period: 420 months. 908 simulation sets.

ranking (908)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	overall	div-risk
naive: 24 assets	2 (0)	2 (0)	2 (71)	2 (22)	2 (4)	1 (897)	2 (6)	2	2 (228)
naive: 48 assets	3 (0)	1 (908)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3	1 (680)
plug-in mvo	1 (908)	3 (0)	1 (837)	1 (886)	1 (904)	2 (11)	1 (902)	1	-

Estimation window: 180 months, testing window: 12 months, total testing period: 360 months. 936 simulation sets.

ranking (936)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	overall	div-risk
naive: 24 assets	2 (3)	2 (0)	2 (37)	2 (10)	2 (5)	1 (921)	2 (5)	2	2 (246)
naive: 48 assets	3 (0)	1 (936)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3	1 (690)
plug-in mvo	1 (933)	3 (0)	1 (899)	1 (926)	1 (931)	2 (15)	1 (931)	1	-

Estimation window: 240 months, testing window: 12 months, total testing period: 300 months. 953 simulation sets.

ranking (953)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	overall	div-risk
naive: 24 assets	2 (3)	2 (0)	2 (25)	2 (16)	2 (6)	1 (938)	2 (4)	2	2 (264)
naive: 48 assets	3 (0)	1 (953)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3	1 (689)
plug-in mvo	1 (950)	3 (0)	1 (928)	1 (937)	1 (947)	2 (15)	1 (949)	1	-

Estimation window: 300 months, testing window: 12 months, total testing period: 240 months. 962 simulation sets.

ranking (962)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	overall	div-risk
naive: 24 assets	2 (7)	2 (0)	2 (13)	2 (18)	2 (12)	1 (946)	2 (7)	2	2 (288)
naive: 48 assets	3 (0)	1 (962)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3	1 (674)
plug-in mvo	1 (955)	3 (0)	1 (949)	1 (944)	1 (950)	2 (16)	1 (955)	1	-

Estimation window: 360 months, testing window: 12 months, total testing period: 180 months. 971 simulation sets.

ranking (971)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	overall	div-risk
naive: 24 assets	2 (8)	2 (0)	2 (17)	2 (25)	2 (15)	1 (959)	2 (8)	2	2 (313)
naive: 48 assets	3 (0)	1 (971)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3	1 (658)
plug-in mvo	1 (963)	3 (0)	1 (954)	1 (946)	1 (956)	2 (12)	1 (963)	1	-

Table 4.14: Model 3: Annual ranking. Comparison of the naive portfolio with 24 assets, the full naive portfolio and the unconstrained mean-variance portfolio.
Settings: Simulation model 3 with 48 independent assets with market switching parameters. Estimation window: 60 - 360 months, testing window: 12 months, total testing period: 480 - 180 months.
The number in brackets after the ranking indicates how often the portfolio outperformed both other strategies in terms of this performance criterion.

than the ratio of the full naive portfolio. These relations stay the same with rising estimation period and are only slightly weakened so that they are true for most of the cases. Again, the mean-variance portfolio shows extreme outcomes starting with 180 months estimation period. Only the statement of larger volatility always remains true.

In Figure 4.9 we see that a ranking of portfolios is still observable, excluding the extreme cases. The clear distance between the two naive strategies, especially in the mean comes from the selection possibility of the naive strategy with 24 asset. It can avoid generally bad performing assets, whereas the full naive strategy is forced to take them. However, in the diversification-risk ratio, we see that the full naive portfolio is the more diversified portfolio. This certainly leads to better performance in a real data setting, as the evolution of assets is not predictable. A nice observation is that the performance of the naive method with fewer assets is close to the performance of the sample-mean-variance strategy. An example for the performance values in Model 3 can be seen in Figure 4.9 with all performance criteria for the 180 months estimation period.

4.4.3 Concluding Remark

In the real data setting the naive portfolio is not consistently outperformed by the unconstrained mean-variance portfolio, meaning that both naive strategies are ranked on a better place for most of the performance criteria. The naive portfolio with 24 assets stays close to the full naive strategy, but is usually ranked on the second place. Yet, there is no best choice for the number of assets in a naive portfolio with fewer assets. The diversification-risk ratio also favors the full naive portfolio. The mean-variance optimization fails to produce reasonable strategies for estimation periods lower than 180 months. Although the increase of estimation period should improve its performance considerably, this cannot be observed for this example.

For the simulated data it depends whether the model favors the mean-variance strategy (static parameters) or not (switching parameters). For the first, the mean-variance portfolio outperforms the other strategies in the overall ranking. For the latter, the full naive portfolio outperforms the other strategies when the parameters can change completely. However, in the case of market switching the mean-variance portfolio again takes first place. So, whatever the model, the naive portfolio with fewer assets is ranked in the middle. Additionally, we observe that the unconstrained mean-variance strategy still produces extreme positions, as can be seen in the peaks in the performance graphs.

Thus, we are going to investigate several versions of the mean-variance strategy to choose an appropriate benchmark.

4.5 Mean-Variance Optimization

DeMiguel, Garlappi, and Uppal (2009) use the unconstrained plug-in mean-variance method for the theoretical results as the other mean-variance strategies can only be performed numerically. In the following chapters, we will concentrate on mean-variance strategies without shortselling and thus will show some results on the comparison of three mean-variance methods: the sample (unconstrained) mean-variance, the (constrained) mean-variance without shortselling and the minimum-variance optimization.

The last one can be seen as a special case, where the mean is ignored or restricted to the same value for all assets. The advantage is that only the covariance matrix has to be estimated, which

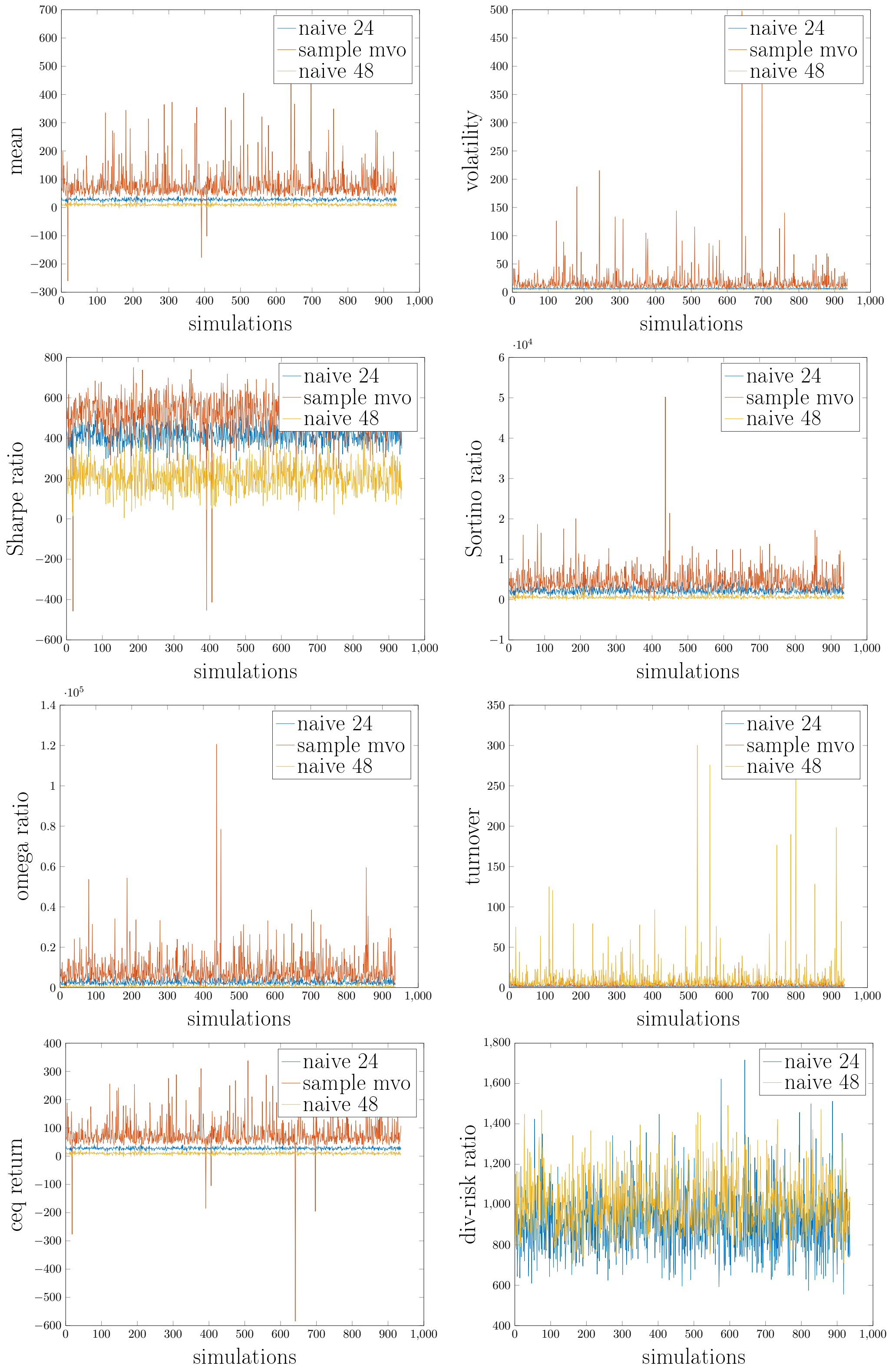


Figure 4.9: Model 3: Annual performance (in percent, except turnover). Simulation model 3, 48 assets, 180 months estimation window, 360 months total testing period. 1000 simulation sets. (For more details, the corresponding graphs with only the first 100 simulations can be found in Appendix D)

can be done quite well, whereas the mean is hard to estimate and leads to high estimation errors. We have seen similar effects in Section 4.3. One advantage of the other mean-variance strategies is that they do not lead to unreasonable portfolio returns and therefore can be used in every setting. The diversification-risk ratio is omitted in this section, as also the minimum-variance strategy leads to negative weights.

4.5.1 Data Experiments

In the following, we see results on the first, second and third simulation model. The relation of the three mean-variance methods stay the same for different estimation periods, as all three depend mainly on the same moment estimation. Therefore, it suffices to show the results for the case of 180 months estimation period.

In Table 4.15, we see the results for Simulation Model 1 with independent assets and static parameters. The mean-variance portfolio without shortselling always has a greater mean, Sharpe, Sortino and omega ratio and a greater CEQ return than the minimum-variance portfolio. Whereas the unconstrained mean-variance portfolio usually has the highest mean, Sharpe, Sortino and omega ratio and the highest CEQ return. However, its volatility is always greater than the volatility of the mean-variance portfolio without shortselling, which in turn is always greater than the volatility of the minimum-variance strategy. The mean-variance portfolio without shortselling always produces the smallest turnover, whereas the minimum-variance strategy usually has the highest turnover. See Figure 4.10 for the graphs of all averaged annual performance data for all simulation sets.

ranking (952)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	overall
plug-in mvo	1 (945)	3 (0)	1 (852)	1 (803)	1 (851)	2 (0)	1 (944)	1
mvo w/o ss	2 (7)	2 (0)	2 (100)	2 (149)	2 (101)	1 (952)	2 (8)	2
min vo	3 (0)	1 (952)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3

Table 4.15: Model 1: Mean-variance ranking of the sample, no shortselling and minimum-variance portfolio. Simulation model 1, 48 assets, 180 months estimation window, 360 months total testing period. 952 simulation sets.

In Table 4.16, we see the results for Simulation Model 2 with independent assets and switching parameters (each year). There, we can make the following general statements: The minimum-variance portfolio outperforms both other portfolio strategies in terms of the volatility, the Sharpe, Sortino and omega ratio. For the volatility, the Sortino and omega ratio we always have that the mean-variance portfolio without shortselling outperforms the sample mean variance strategy. For the Sharpe ratio the sample mean-variance method is often smaller than the other methods. For the mean there is no clear ranking observable, only that the mean values of the no-shortselling mean-variance and the sample mean-variance portfolio spread around the values of the minimum-variance portfolio. This carries over to the CEQ return results. Moreover, the spreading of the sample mean-variance portfolio increases caused by the high volatility. For the turnover it is more likely that the sample mean-variance produces higher turnover than the mean-variance without shortselling (around 77 percent) and the minimum-variance portfolio (around 86 percent). Whereas the no-shortselling portfolio is more likely to have a higher turnover than the minimum-variance portfolio (62 percent). See Figure 4.11 for the corresponding graphs.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	overall
plug-in mvo	2 (411)	3 (0)	3 (0)	3 (0)	3 (0)	3 (47)	3 (337)	3
mvo w/o ss	1 (159)	2 (0)	2 (0)	2 (0)	2 (0)	2 (356)	2 (193)	2
min vo	3 (430)	1 (1000)	1 (1000)	1 (1000)	1 (1000)	1 (597)	1 (470)	1

Table 4.16: Model 2: Mean-variance ranking of the sample, no shortselling and minimum-variance portfolio. Simulation model 2, 48 assets, 180 months estimation window, 360 months total testing period. 1000 simulation sets.

In Table 4.17, we see the results for Simulation Model 3, independent assets and switching market parameters (each year). In terms of standard deviation, we still have the absolute ordering from sample mean-variance over no-shortselling mean-variance to minimum-variance portfolio with decreasing values. The mean-variance portfolio without shortselling outperforms the minimum-variance portfolio in the mean, the Sharpe, Sortino and omega ratio as well as in the turnover and the CEQ return. The sample mean variance is usually better in terms of mean and CEQ return. However, it produces nearly overlapping values with the no-shortselling values for the Sharpe, Sortino and omega ratio. Additionally, it nearly always has a higher turnover than the mean-variance strategy without shortselling. See Figure 4.12 for the corresponding diagrams.

ranking (936)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	overall
plug-in mvo	1 (933)	3 (0)	1 (807)	1 (756)	1 (821)	2 (1)	1 (931)	1
mvo w/o ss	2 (3)	2 (0)	2 (129)	2 (180)	2 (115)	1 (935)	2 (5)	2
min vo	3 (0)	1 (936)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3

Table 4.17: Model 3: Mean-variance ranking of the sample, no shortselling and minimum-variance portfolio. Simulation model 3, 48 assets, 180 months estimation window, 360 months total testing period. 936 simulation sets.

4.5.2 Concluding Remark

In the end, we prefer to use the constrained mean-variance portfolio without shortselling, usually just called mean-variance portfolio in the remaining simulation studies. It does not lead to the undesired negative portfolio returns, which make a comparison harder. It is not only concentrated on minimizing the variance and therefore leads to better performance. The constrained mean-variance portfolio lies in between the minimum-variance portfolio and the unconstrained mean-variance portfolio and thus, it is the most challenging benchmark of the considered mean-variance strategies.

4.6 Data Experiments with Constrained Mean-Variance Portfolio

We end this chapter with a comparison of the performance of the constrained mean-variance portfolio with the full naive portfolio and the one with only 24 assets. Now we can get results in all estimation period settings and we can apply the diversification-risk ratio to them.

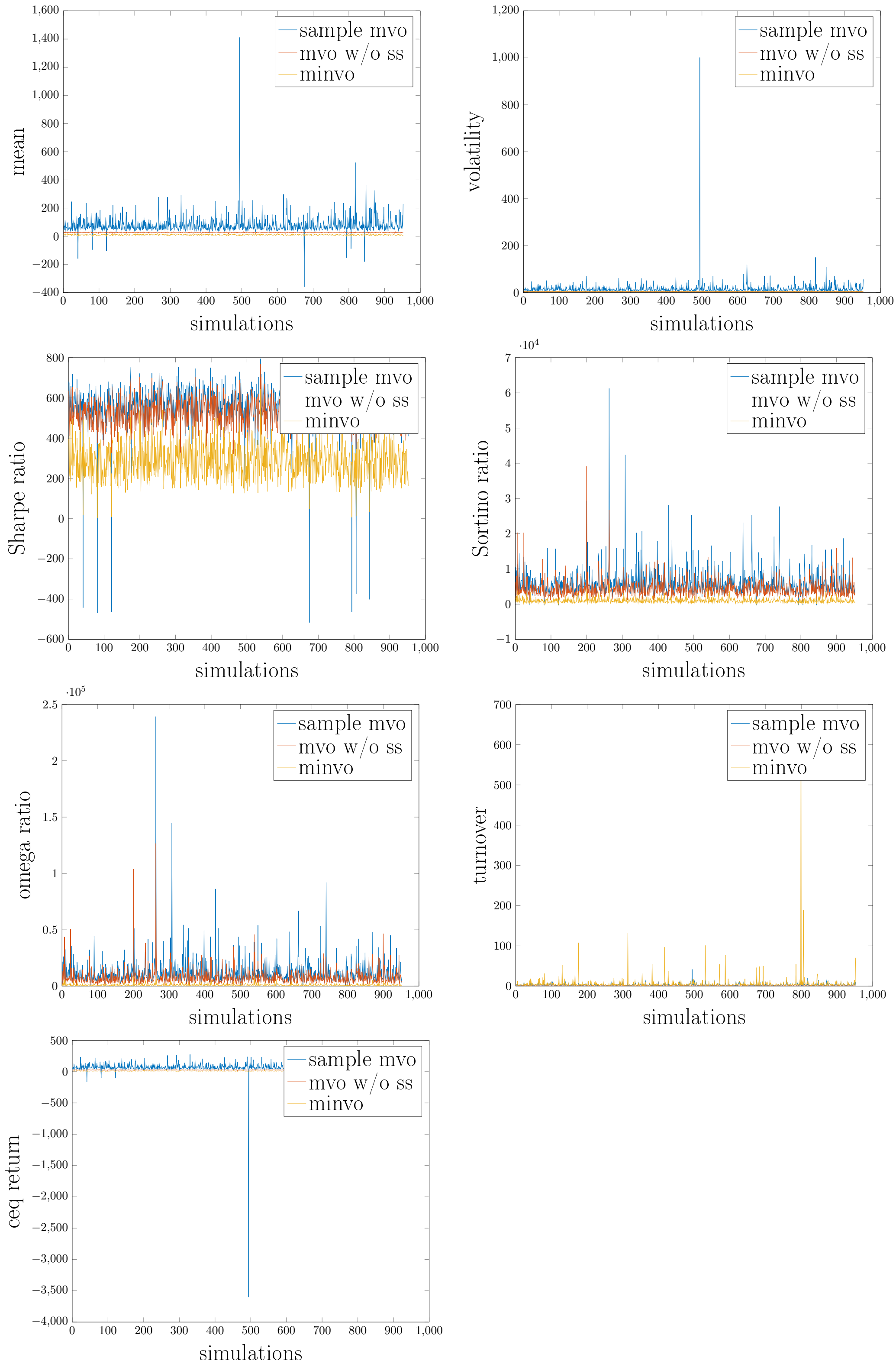


Figure 4.10: Model 1: Mean-variance performance (in percent, except turnover) of the sample, no short-selling and minimum-variance portfolio. Simulation model 1, 48 assets, 180 months estimation window, 360 months total testing period. 952 simulation sets. (For more details, the corresponding graphs with only the first 100 simulations can be found in Appendix D)

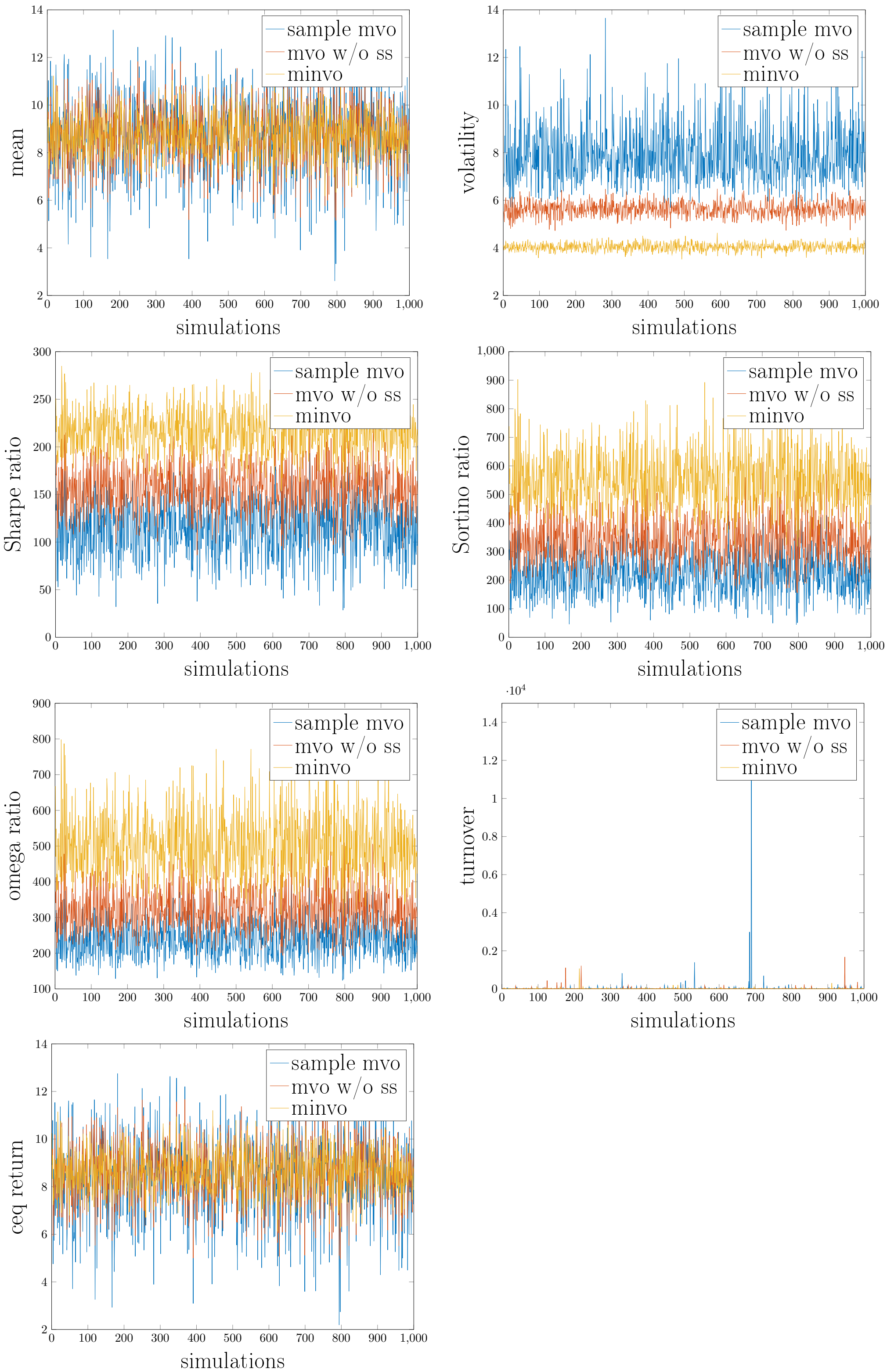


Figure 4.11: Model 2: Mean-variance performance (in percent, except turnover) of the sample, no short-selling and minimum-variance portfolio. Simulation model 2, 48 assets, 180 months estimation window, 360 months total testing period. 1000 simulation sets. (For more details, the corresponding graphs with only the first 100 simulations can be found in Appendix D)

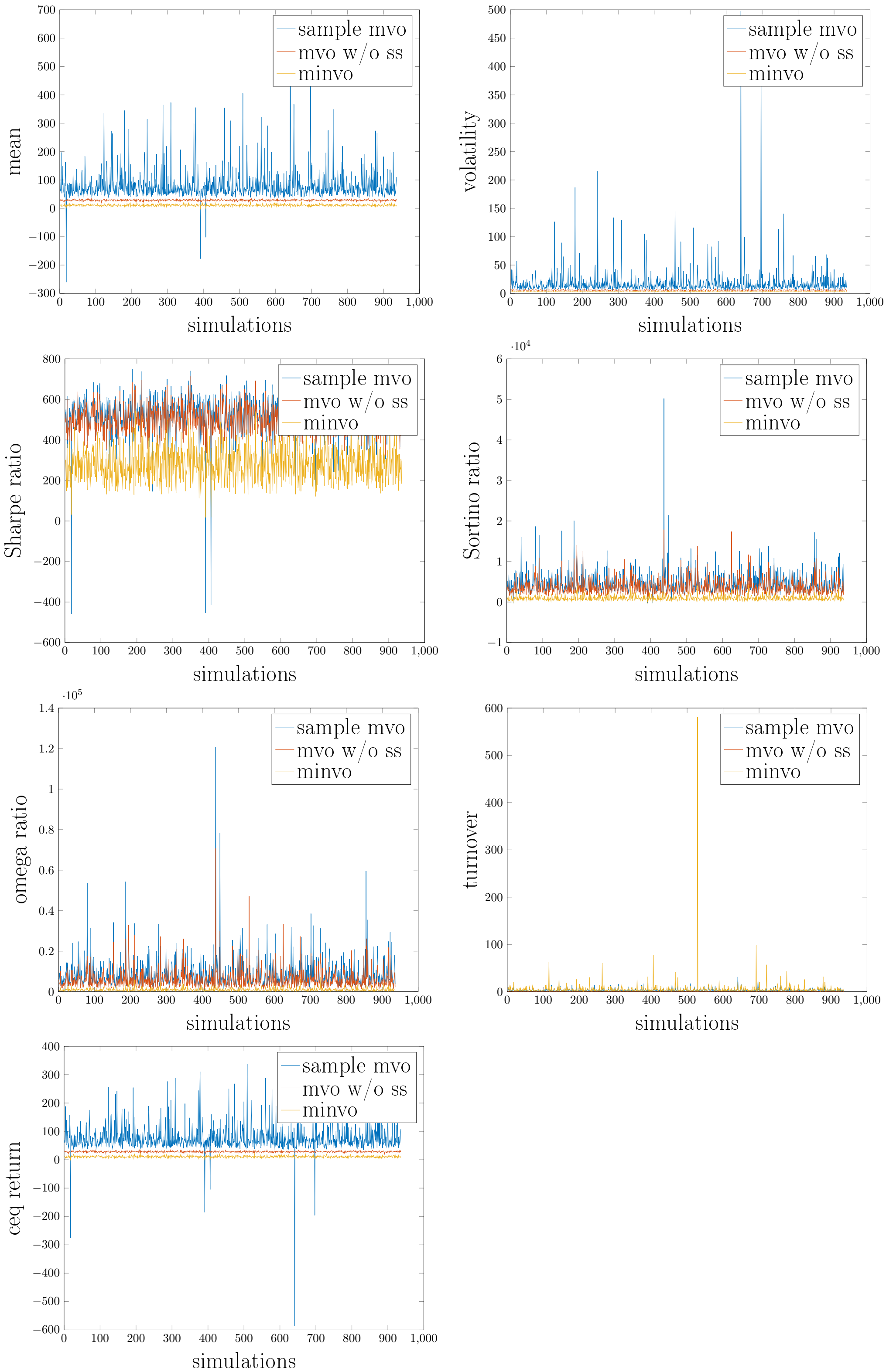


Figure 4.12: Model 3: Mean-variance performance (in percent, except turnover) of the sample, no short-selling and minimum-variance portfolio. Simulation model 3, 48 assets, 180 months estimation window, 360 months total testing period. 1000 simulation sets. (For more details, the corresponding graphs with only the first 100 simulations can be found in Appendix D)

4.6.1 Real Data

First we start with the real data example. Although we are able to get results for all considered estimation periods, we only show the results for 60, 180 and 360 months, because the others do not provide more insight. We see the annual performances over the whole period and the last three five-year-periods in Table 4.18 for 60 months estimation period, in Table 4.20 for 180 months estimation period and in Table 4.22 for 360 months estimation period. The corresponding rankings are given in Tables 4.19, 4.21 and 4.23. How prices of the portfolios would evolve over time starting with 100 for each strategy, can be seen in Figure 4.13 together with the number of assets in the portfolios for each investment period.

In the 60 months case the performance of the full naive and the naive portfolio with 24 assets is quite similar. Only the full naive method is slightly better. The mean-variance optimization without shortselling does not only produce reasonable results, but performs much better than the sample mean-variance portfolio without constraints. As expected, the full naive portfolio shows the highest diversification-risk ratio and the mean-variance portfolio the smallest value. The resulting ordering suddenly changes with larger estimation periods. The mean-variance strategy performs better and better, resulting in the overall second place for 180 months and in the first place for 360 months. Clearly, the decreasing estimation risk strengthens the method. Whereas the full naive portfolio stays quite stable. First it takes first place and then falls back to second place. Since it has no problems with estimation, it is on the other hand not able to optimize the strategy and therefore it falls behind in the last case. The naive method with 24 assets gains its good performance from the small impact of estimation error, but rapidly looses good performance as it has neither the robustness from investing in all assets nor the possibility of optimizing the strategy. Thus it falls to third place for 180 and 360 months. Highly astonishing is the superiority of the mean-variance strategy in terms of the diversification risk ratio in the last two cases. As this strategy takes the whole covariance matrix into account, it can also improve in diversification. A look at the components of this ratio shows that not only a reduced Ω-risk leads to the better ratio but also an increase in diversification measure.

4.6.2 Simulated Data

Next, we see the results for the simulated data for 60, 180 and 360 months estimation period.

Simulation Model 1
The first simulation model with independent assets and static parameters is in favor of the mean-variance procedures. Although 60 months estimation period is usually not long enough to get good estimates, the constrained mean variance portfolio is at the first place in the over all ranking.

Corresponding to the 60 months estimation period case in Table 4.24, we get the following general statements. The full naive strategy always results in lower values than the other strategies in terms of the mean, the Sharpe, Sortino and omega ratio, as well as the CEQ return. It usually has the lowest volatility, definitely lower than the naive portfolio with 24 assets. It also results in the highest turnover, whereas the constrained mean-variance always has the lowest. Regarding the diversification-risk ratio, the full naive portfolio has greater values than the naive method with 24 assets in around 70 percent of the cases. The mean-variance strategy without shortselling has greater values than the naive with fewer assets in about 80 percent of the cases,

60/480/annual	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq return	div-risk ratio
naive: 24 assets	16.74	17.05	98.26	166.91	211.32	3.79	15.28	108.68
1st period	8.02	15.63	51.75	81.72	144.72	13.94	6.80	174.17
2nd period	6.39	18.34	35.14	48.42	132.57	8.82	4.71	89.57
3rd period	22.15	15.67	142.56	272.73	281.99	1.18	20.92	99.10
naive: 48 assets	16.73	17.06	98.19	167.81	212.03	1.98	15.28	112.94
1st period	10.84	14.70	74.36	123.64	170.79	8.84	9.76	191.21
2nd period	5.27	18.87	28.17	39.93	126.18	1.16	3.49	85.52
3rd period	22.98	16.12	143.78	273.11	281.63	1.34	21.68	95.26
mvo w/o ss	17.38	18.32	95.01	168.76	206.57	2.57	15.71	97.16
1st period	10.04	17.78	56.91	94.23	152.76	5.82	8.46	157.81
2nd period	5.47	18.55	29.75	42.32	124.67	6.20	3.75	97.12
3rd period	20.94	14.49	145.80	277.68	268.64	0.64	19.89	76.28

Table 4.18: Real data: Annual performance of the naive portfolio with 24 assets, the full naive portfolio and the mean-variance portfolio without shortselling.
estimation window: 60 months of daily data, test window: 12 months of monthly data, total testing period: 480 months

annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
naive: 24 assets	2	1	1	3	2	3	2	2	2
naive: 48 assets	3	2	2	2	1	1	3	1	1
mvo w/o ss	1	3	3	1	3	2	1	3	3

Table 4.19: Real data: Ranking of the naive portfolio with 24 assets, the full naive portfolio and the mean-variance portfolio without shortselling.
estimation window: 60 months of daily data, test window: 12 months of monthly data, total testing period: 480 months

180/360/annual	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq return	div-risk ratio
naive: 24 assets	14.89	15.72	94.90	152.10	206.60	13.89	13.66	106.56
1st period	5.59	14.69	38.37	59.07	131.98	74.12	4.51	165.11
2nd period	3.79	17.31	22.05	29.52	119.20	2.33	2.29	82.88
3rd period	21.23	15.31	139.89	262.46	281.66	1.21	20.06	85.92
naive: 48 assets	15.84	16.24	97.69	158.45	211.35	2.40	14.53	121.33
1st period	10.84	14.70	74.36	123.64	170.79	8.84	9.76	191.21
2nd period	5.27	18.87	28.17	39.93	126.18	1.16	3.49	85.52
3rd period	22.98	16.12	143.78	273.11	281.63	1.34	21.68	95.26
mvo w/o ss	14.21	14.86	95.71	159.64	206.28	2.00	13.10	134.44
1st period	5.38	14.05	38.64	61.19	133.55	4.67	4.40	170.79
2nd period	7.13	16.90	42.55	59.21	139.46	1.36	5.70	110.16
3rd period	20.26	18.12	112.75	210.20	240.76	3.80	18.61	107.67

Table 4.20: Real data: Annual performance of the naive portfolio with 24 assets, the full naive portfolio and the mean-variance portfolio without shortselling.
Estimation window: 180 months of daily data, test window: 12 months of monthly data, total testing period: 360 months

annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
naive: 24 assets	2	2	3	3	2	3	2	3	3
naive: 48 assets	1	3	1	2	1	2	1	2	1
mvo w/o ss	3	1	2	1	3	1	3	1	2

Table 4.21: Real data: Ranking of the naive portfolio with 24 assets, the full naive portfolio and the mean-variance portfolio without shortselling.
Estimation window: 180 months of daily data, test window: 12 months of monthly data, total testing period: 360 months

360/180/annual	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq return	div-risk ratio
naive: 24 assets	10.46	14.60	71.80	111.09	170.81	7.25	9.39	115.47
1st period	6.93	14.29	48.94	76.92	142.17	18.83	5.91	161.01
2nd period	2.77	15.85	17.65	23.46	115.33	0.99	1.52	88.25
3rd period	21.66	13.24	165.06	323.73	321.75	0.61	20.79	85.20
naive: 48 assets	13.03	16.70	78.26	125.80	181.28	3.95	11.64	126.16
1st period	10.84	14.70	74.36	123.64	170.79	8.84	9.76	191.21
2nd period	5.27	18.87	28.17	39.93	126.18	1.16	3.49	85.52
3rd period	22.98	16.12	143.78	273.11	281.63	1.34	21.68	95.26
mvo w/o ss	10.42	11.62	89.93	144.66	194.76	4.04	9.75	179.08
1st period	6.94	12.50	56.01	91.23	151.79	9.12	6.16	187.75
2nd period	6.27	12.15	52.06	71.60	149.50	1.65	5.53	147.42
3rd period	18.05	9.92	183.46	393.43	361.04	0.67	17.56	196.85

Table 4.22: Real data: Annual performance of the naive portfolio with 24 assets, the full naive portfolio and the mean-variance portfolio without shortselling.
estimation window: 360 months of daily data, test window: 12 months of monthly data, total testing period: 180 months

annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
naive: 24 assets	2	2	3	3	3	3	3	3	3
naive: 48 assets	1	3	2	2	2	1	1	2	2
mvo w/o ss	3	1	1	1	1	2	2	1	1

Table 4.23: Real data: Ranking of the naive portfolio with 24 assets, the full naive portfolio and the mean-variance portfolio without shortselling.
estimation window: 360 months of daily data, test window: 12 months of monthly data, total testing period: 180 months

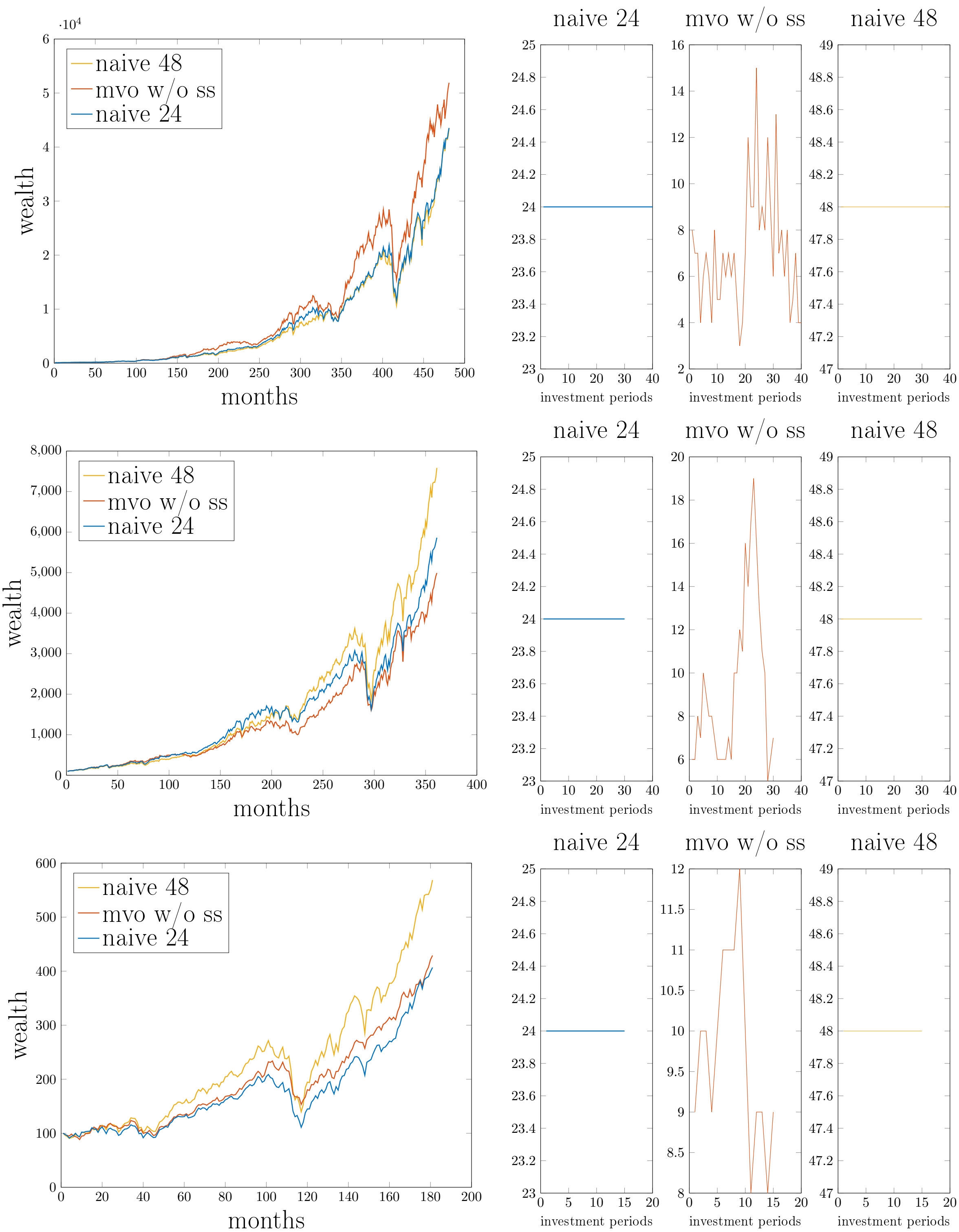


Figure 4.13: Real data: Portfolio prices starting at 100 for the naive strategies with 24 and 48 assets and the mean-variance portfolio without shortselling. Number of assets in a portfolio. Estimation periods: 60, 180, and 360 months

whereas it only outperforms the full naive method in 66 percent of the cases.

The graphs are similar to the exemplary graphs for 180 months in Figure 4.14. It is striking that the values from the constrained mean-variance and the naive method with fewer assets are on the same level for the mean and the CEQ return. The results for larger estimation periods changes only slightly and the overall ranking, as well as the definite statements above, stay the same. Only the statement about the turnover in the 360 months case changes in a way that the full naive portfolio always leads to the highest values, but the naive method with fewer assets only result in higher values than the mean-variance method with fewer assets in nearly all cases.

Simulation Model 2

The second model, with changing parameters, is better for the full naive method, see Table 4.25. Thus, it always takes the first place in rankings except for the mean. It always has the smallest volatility and the largest values for the Sharpe, Sortino and omega ratio as well as the diversification risk ratio. The naive method with fewer assets and the mean-variance portfolio without shortselling again share the same range of values. The turnover is no longer dominated by the full naive method and no general structure is observable.

These statements remain for the 180 months case, but for the 360 months case the gap between the full naive and the other strategies (see Figure 4.15) becomes smaller and therefore the statements are no longer absolute. Only the standard deviation of the full naive method stays the smallest for all cases. In addition, the diversification-risk ratio always remains above the values from the naive method with fewer assets.

The ranking for the naive method with 24 assets and the constrained mean-variance portfolio would rather be the same, only due to small changes they swap between second and third place. For the CEQ return ranking it seems that the full naive portfolio outperforms the other strategies, but it falls more often to the third place than the others. Although the difference in CEQ ranking values is small for all methods, the full naive strategy is always on the third place. Only, the mean-variance portfolio and the naive portfolio with fewer assets swap between first and second place.

Simulation Model 3

In the third model it should be more a question of estimation period, whether the mean variance portfolio outperforms the rest, due to the market state changes. As we can see in Table 4.26, it is clear from the start that the mean-variance without shortselling performs better and gets the first place in ranking.

We observe again a gap between the full naive performance values and the values of the other strategies, see Figure 4.16. This is always true for the mean, the Sharpe, Sortino and omega ratio as well as the CEQ return. In terms of the volatility, there is a gap between the full naive and the naive method with 24 assets. The mean-variance volatility values lie in between and sometimes overlap the other strategies' values. Regarding the diversification-risk ratio we see that the naive method with fewer assets is outperformed by the other strategies in around 75 percent of the cases whereas the mean-variance portfolio achieves to outperform the full naive method in an increasing amount of first only 43 percent, then 59 and in the end in 64 percent of the cases. There we see at least an increase in performance with rising estimation period. However, the performance of the full naive and the mean-variance portfolio is quite similar for the diversification-risk ratio and thus swap between the first and second place.

Estimation window: 60 months, testing window: 12 months, total testing period: 480 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
naive: 24 assets	2 (346)	3 (0)	2 (18)	2 (35)	2 (22)	2 (0)	2 (339)	3 (88)	2
naive: 48 assets	3 (0)	1 (988)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	2 (296)	3
mvo w/o ss	1 (654)	2 (12)	1 (982)	1 (965)	1 (978)	1 (1000)	1 (661)	1 (616)	1

Estimation window: 180 months, testing window: 12 months, total testing period: 360 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
naive: 24 assets	2 (420)	3 (0)	2 (4)	2 (15)	2 (4)	2 (0)	2 (409)	3 (97)	2
naive: 48 assets	3 (0)	1 (958)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	2 (205)	3
mvo w/o ss	1 (580)	2 (42)	1 (996)	1 (985)	1 (996)	1 (1000)	1 (591)	1 (698)	1

Estimation window: 360 months, testing window: 12 months, total testing period: 180 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
naive: 24 assets	2 (449)	3 (0)	2 (5)	2 (28)	2 (17)	2 (2)	2 (440)	3 (126)	2
naive: 48 assets	3 (0)	1 (924)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	2 (211)	3
mvo w/o ss	1 (551)	2 (76)	1 (995)	1 (972)	1 (983)	1 (998)	1 (560)	1 (663)	1

Table 4.24: Model 1: Annual ranking. Comparison of the naive portfolio with 24 assets, the full naive portfolio and the constrained mean-variance portfolio.
Settings: Simulation model 1 with 48 independent assets with static parameters. Estimation window: 60, 180 and 360 months, testing window: 12 months, total testing period: 480, 360 and 180 months.
The number in brackets after the ranking indicates how often the portfolio outperformed both other strategies in terms of this performance criterion.

Estimation window: 60 months, testing window: 12 months, total testing period: 480 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
naive: 24 assets	2 (273)	2 (0)	2 (0)	2 (0)	2 (0)	3 (210)	1 (268)	2 (0)	2
naive: 48 assets	3 (370)	1 (1000)	1 (1000)	1 (1000)	1 (1000)	1 (560)	3 (412)	1 (1000)	1
mvo w/o ss	1 (357)	3 (0)	3 (0)	3 (0)	3 (0)	2 (230)	2 (320)	3 (0)	3

Estimation window: 180 months, testing window: 12 months, total testing period: 360 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
naive: 24 assets	2 (303)	2 (0)	2 (0)	3 (0)	3 (0)	3 (212)	2 (287)	3 (0)	3
naive: 48 assets	3 (382)	1 (1000)	1 (1000)	1 (1000)	1 (1000)	1 (535)	3 (419)	1 (1000)	1
mvo w/o ss	1 (315)	3 (0)	3 (0)	2 (0)	2 (0)	2 (253)	1 (294)	2 (0)	2

Estimation window: 360 months, testing window: 12 months, total testing period: 180 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
naive: 24 assets	3 (339)	3 (0)	3 (2)	3 (4)	3 (2)	3 (210)	2 (323)	3 (0)	3
naive: 48 assets	2 (400)	1 (1000)	1 (996)	1 (988)	1 (993)	1 (514)	3 (423)	1 (995)	1
mvo w/o ss	1 (261)	2 (0)	2 (2)	2 (8)	2 (5)	2 (276)	1 (254)	2 (5)	2

Table 4.25: Model 2: Annual ranking. Comparison of the naive portfolio with 24 assets, the full naive portfolio and the constrained mean-variance portfolio.
Settings: Simulation model 2 with 48 independent assets with switching parameters. Estimation window: 60, 180 and 360 months, testing window: 12 months, total testing period: 480, 360 and 180 months.
The number in brackets after the ranking indicates how often the portfolio outperformed both other strategies in terms of this performance criterion.

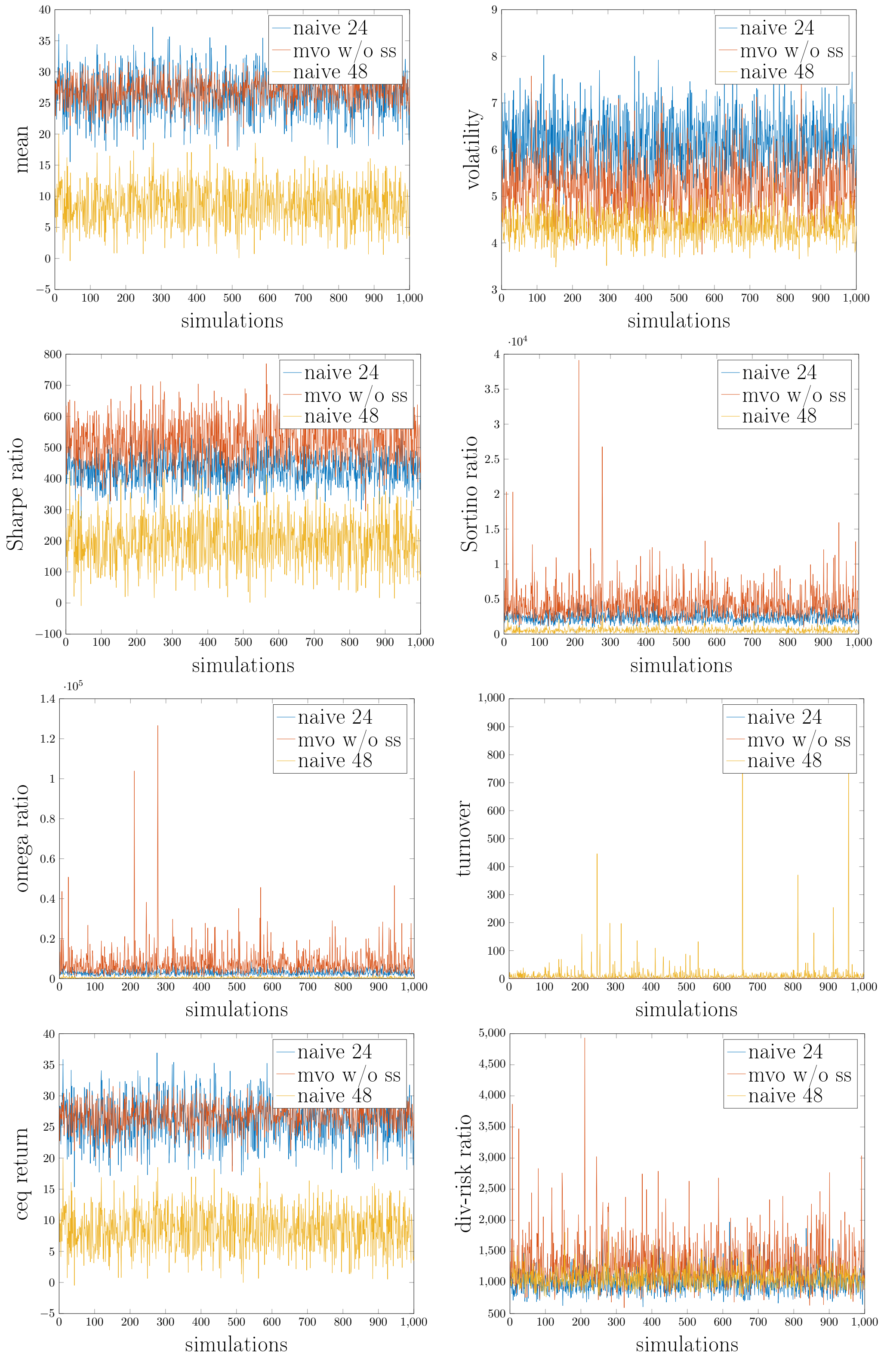


Figure 4.14: Model 1: Annual performance (in percent, except turnover). Simulation model 1, 48 assets, 180 months estimation window, 360 months total testing period. 1000 simulation sets. (For more details, the corresponding graphs with only the first 100 simulations can be found in Appendix D)

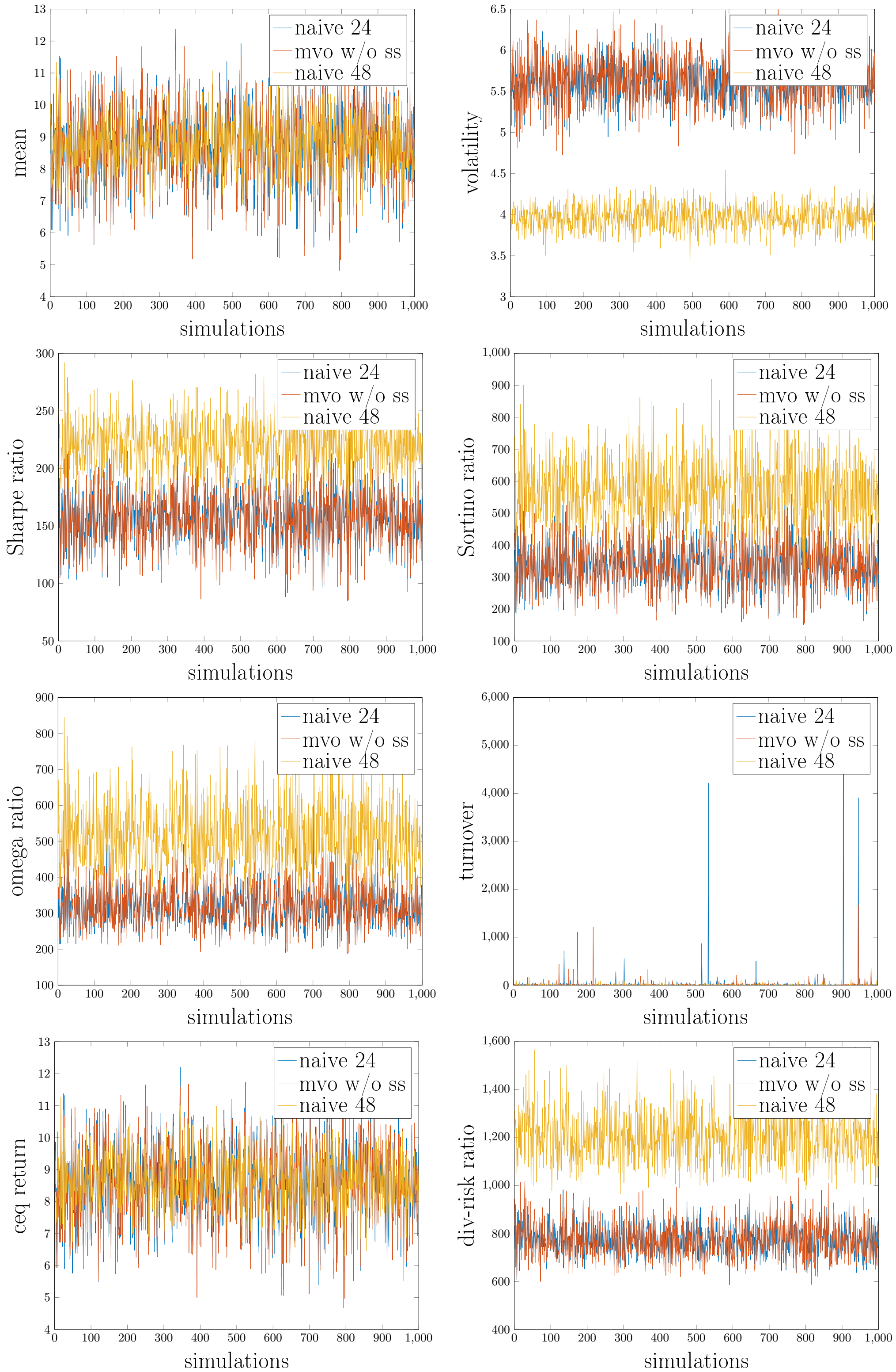


Figure 4.15: Model 2: Annual performance (in percent, except turnover). Simulation model 2, 48 assets, 180 months estimation window, 360 months total testing period. 1000 simulation sets. (For more details, the corresponding graphs with only the first 100 simulations can be found in Appendix D)

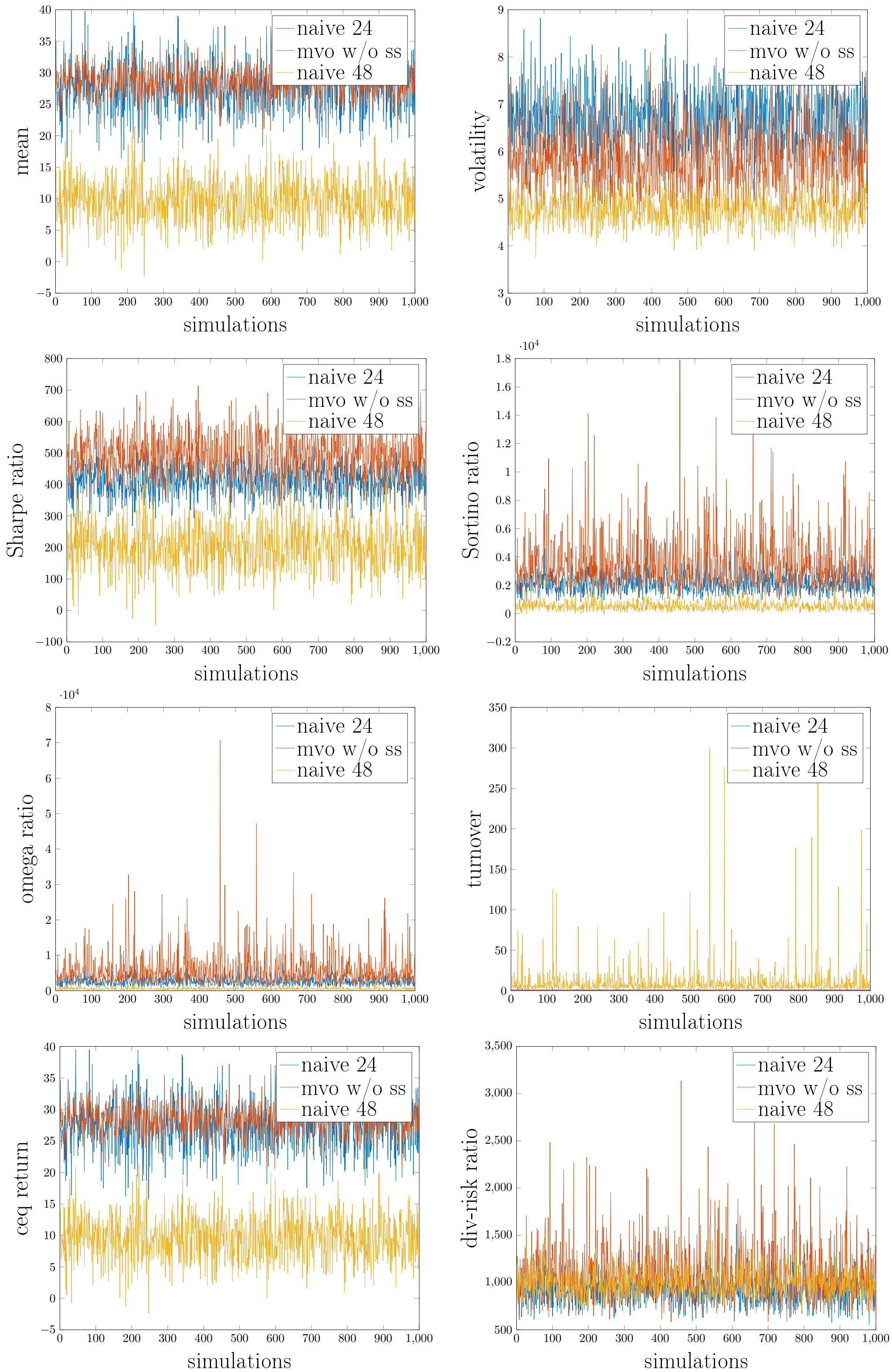


Figure 4.16: Model 3: Annual performance (in percent, except turnover). Simulation model 3, 48 assets, 180 months estimation window, 360 months total testing period. 1000 simulation sets. (For more details, the corresponding graphs with only the first 100 simulations can be found in Appendix D)

Estimation window: 60 months, testing window: 12 months, total testing period: 480 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
naive: 24 assets	2 (265)	3 (0)	2 (45)	2 (73)	2 (48)	2 (8)	2 (260)	3 (100)	2
naive: 48 assets	3 (0)	1 (992)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	1 (504)	3
mvo w/o ss	1 (735)	2 (8)	1 (955)	1 (927)	1 (952)	1 (992)	1 (740)	2 (396)	1

Estimation window: 180 months, testing window: 12 months, total testing period: 360 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
naive: 24 assets	2 (376)	3 (0)	2 (13)	2 (41)	2 (19)	2 (2)	2 (369)	3 (100)	2
naive: 48 assets	3 (0)	1 (966)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	2 (358)	3
mvo w/o ss	1 (624)	2 (34)	1 (987)	1 (959)	1 (981)	1 (998)	1 (631)	1 (542)	1

Estimation window: 360 months, testing window: 12 months, total testing period: 180 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
naive: 24 assets	2 (399)	3 (0)	2 (17)	2 (60)	2 (41)	2 (8)	2 (388)	3 (119)	2
naive: 48 assets	3 (0)	1 (939)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	2 (295)	3
mvo w/o ss	1 (601)	2 (61)	1 (983)	1 (940)	1 (959)	1 (992)	1 (612)	1 (586)	1

Table 4.26: Model 3: Annual ranking. Comparison of the naive portfolio with 24 assets, the full naive portfolio and the constrained mean-variance portfolio.
Settings: Simulation model 3 with 48 independent assets with market switching parameters. Estimation window: 60, 180 and 360 months, testing window: 12 months, total testing period: 480, 360 and 180 months.
The number in brackets after the ranking indicates how often the portfolio outperformed both other strategies in terms of this performance criterion.

4.6.3 Concluding Remark

As we have seen so far, it is possible to keep beneficial characteristics of the naive portfolio, when choosing fewer assets. In the real data setting, the constrained mean-variance portfolio can be compared to the other portfolios for all estimation periods. We finally observe an improvement of its performance with an increasing estimation period, resulting in the first place in the overall ranking for 360 months. By disallowing shortselling, the average performance of the mean-variance portfolio is considerably improved. Therefore, this will be our benchmark portfolio in the following.

In the simulation settings, where the constrained mean-variance portfolio performs well (models 1 and 3) the naive portfolio with 24 assets gets close to the performance of this benchmark. In addition, it is able to outperform the full naive method in some settings. In Simulation Model 2, it falls from the second to the third place with rising estimation period. There, the constrained mean-variance portfolio can improve its performance considerably. However, the naive portfolio with fewer asset is usually on the third place regarding the diversification-risk ratio, whereas the mean-variance portfolio is on the first place, where it performs well.

4.7 Summary

In this chapter we have seen that reducing the number of assets in a naive portfolio can have beneficial effects.

In a constant rebalancing setting we were able to show that the expected growth rate of all naive portfolios is the same independent of the number of assets. Considering the number of data, which is needed for the mean-variance portfolio to outperform the naive portfolio, we could state similar results for the naive portfolio with fewer assets. In some cases it can be even harder for the mean-variance strategy to outperform the naive portfolio with fewer assets than the full naive portfolio.

Then, we included the first data experiments on real and simulated data comparing the full naive portfolio, the naive portfolio with half the number of assets and the unconstrained mean-variance portfolio using several performance criteria.

In the real data setting the full naive portfolio performs better than the other strategies, closely followed by the naive portfolio with 24 assets.

We used independent assets in the simulations to exclude the covariance effects, in this first analysis.

In Model 1 with static parameters, the mean-variance strategy outperformed the naive strategies, followed by the naive portfolio with fewer assets.

Using switching parameters in Model 2, we included model uncertainty, which lead to the first place in ranking for the full naive portfolio. Again, the naive strategy with fewer assets followed close behind.

Model 3 was intended to simulate market dependent parameters for three market states. There the mean-variance portfolio again was able to outperform the other strategies. As before, the naive portfolio with fewer assets followed in second place.

The main problem with these data experiments was that the performance of the mean-variance portfolio often could not be used for comparison. This was due to negative weights, causing negative wealth, where our performance criteria were no longer appropriate.

Therefore, we proceeded with a comparison of three mean-variance strategies, the unconstrained mean-variance strategy from above, the constrained mean-variance strategy without shortselling and the minimum-variance strategy. There, the constrained mean-variance optimization always led to reasonable results and performed better than the minimum-variance optimization. This motivated us to use the constrained mean-variance portfolio in the following as a challenging benchmark.

In the subsequent data experiments we observed a considerable improvement of the mean-variance performance, and all data sets could be used for comparison. The rankings for models 1 and 3 stayed the same, with the mean-variance strategy in first place followed by the naive portfolio with fewer assets. In Model 2, the mean-variance strategy was able to improve from the third to the second place with increasing estimation period, leaving the naive method with fewer assets in third place.

So far, the selection of assets was only done by comparing the historical Sharpe ratios, and in the simulation models we excluded any dependence structure. If we know more about the assets, in particular on their dependence (correlation structure), we can include this knowledge and expect to improve the choice of assets. We will see in Chapter 5, how we can exploit the knowledge about dependencies between the assets.

5 Naive Diversification with Dependence Structure

We have seen so far, that in some cases it is beneficial to use a smaller number of assets in the naive strategy. Up to now, it was not clear how to choose these assets besides increasing the Sharpe ratio. However, if we possess additional knowledge about the structure in the market such as correlation, we can utilize this knowledge for the selection.

First, we examine the cluster portfolio using the correlation structures in the market for real and simulated data in Section 5.2. In Section 5.3 we will establish a framework, that supports the empirical results found in the first section. This framework is based on the most extreme dependence structure, namely the comonotonicity. In our setting with normally distributed asset returns, this corresponds to a perfect positive correlation within a group. Then, we can show that under certain structural assumptions it is beneficial to invest in the cluster portfolio rather than in the naive portfolio. We include additional data experiments in comonotonic settings in Section 5.4. A summary of this chapter is given in Section 5.5.

5.1 Correlation in the Market

It is commonly known, that an often high correlation exists between assets in stock markets, see e.g. Goldstein, McCarthy, and Orlov (2019). So we can exploit this information by choosing only one asset from each group of highly correlated assets and allocate a naive portfolio on them. Practically this is done by estimating the correlation coefficients and grouping the assets via a cluster algorithm. Then we can choose from each cluster one of the assets to achieve a well diversified portfolio with fewer than all available assets. That this works quite well is shown in the following experiments on real and simulated data.

5.2 Data Experiments

5.2.1 Real Data

First, we compare the cluster portfolios using Pearson correlation with the full naive portfolio and the sample mean-variance portfolio without shortselling on real data. The results are given in Tables 5.1, 5.3 and 5.5, and the ranking results in Tables 5.2, 5.4 and 5.6. The performance of the portfolios for this real data set seems to depend highly on the length of the estimation period. In Section 5.2.2, we will see that the number of correlation groups differs a lot throughout the given data period. It also depends on the estimation period, which is used to find the clusters.

The combination of a naive portfolio with fewer assets, which are selected by clustering, improves the performance enormously when using a fixed number of assets. Then, the cluster portfolio is ranked on the first place for 60 and 360 months estimation period and on the second place for 180 months. The alternative idea to improve performance by letting the cluster algorithm itself find the right number of clusters, does not perform well in this example. This cluster portfolio takes the last place for 60 and 180 months, and only performs better as the full naive portfolio in the 360 months case. The free cluster method detects a lot of clusters and leads to large portfolios with usually more than 24 assets, see Figure 5.1. However, the constrained mean-variance portfolio always consists of fewer assets than both cluster portfolios. This also contributes to a lower trading volume. Yet, there are some important observations: In contrast to the results from Chapter 4, the cluster portfolios improve the performance of the naive portfolio with 24 assets. This is not consistently true for each performance criterium in each case, but in most of the cases and definitely in terms of the diversification-risk ratio.

60/480/annual	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq return	div-risk ratio
Pearson cluster: 24	16.69	16.44	101.61	173.63	219.46	4.04	15.34	130.38
1st period	11.26	13.23	85.81	144.58	190.09	276.34	10.39	215.53
2nd period	6.63	17.63	37.93	53.42	135.75	22.51	5.08	114.12
3rd period	22.80	14.85	154.83	303.66	302.81	0.89	21.70	118.53
Pearson cluster: free	16.23	16.51	98.42	167.47	214.53	3.06	14.87	126.67
1st period	10.65	13.52	79.39	130.24	178.55	187.71	9.73	198.60
2nd period	5.25	18.27	28.99	40.92	126.78	18.59	3.58	102.60
3rd period	22.71	14.56	157.33	313.19	316.30	1.01	21.65	145.81
naive: 48 assets	16.73	17.06	98.19	167.81	212.03	1.98	15.28	112.94
1st period	10.84	14.70	74.36	123.64	170.79	106.11	9.76	191.21
2nd period	5.27	18.87	28.17	39.93	126.18	13.86	3.49	85.52
3rd period	22.98	16.12	143.78	273.11	281.63	1.34	21.68	95.26
mvo w/o ss	17.38	18.32	95.01	168.76	206.57	2.57	15.71	97.16
1st period	10.04	17.78	56.91	94.23	152.76	69.87	8.46	157.81
2nd period	5.47	18.55	29.75	42.32	124.67	74.41	3.75	97.12
3rd period	20.94	14.49	145.80	277.68	268.64	0.64	19.89	76.28

Table 5.1: Real data: Annual performance of the Pearson cluster portfolio with 24 and free number of assets, the full naive portfolio and the sample mean-variance portfolio without shortselling.
estimation window: 60 months of daily data, test window: 12 months of monthly data, total testing period: 480 months

annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	3	1	1	1	1	4	2	1	1
Pearson cluster: free	4	2	2	4	2	3	4	2	4
naive: 48 assets	2	3	3	3	3	1	3	3	2
mvo w/o ss	1	4	4	2	4	2	1	4	3

Table 5.2: Real data: Ranking of the Pearson cluster portfolio with 24 and free number of assets, the full naive portfolio and the sample mean-variance portfolio without shortselling.
estimation window: 60 months of daily data, test window: 12 months of monthly data, total testing period: 480 months

5.2.2 Correlation Structure over Time

If we examine the data for clusters as above, we find that, although it consists of branch portfolios, there are still correlation clusters to be found. In Figure 5.2, we see the monthly change of the number of clusters for 60 and 180 months of estimation period. Note, that with 180 months estimation period only 360 months are left for cluster detection. Thus, only the cluster numbers 121 up to 480 in Figure 5.2(a) correspond to the same estimation periods as the cluster numbers in Figure 5.2(b).

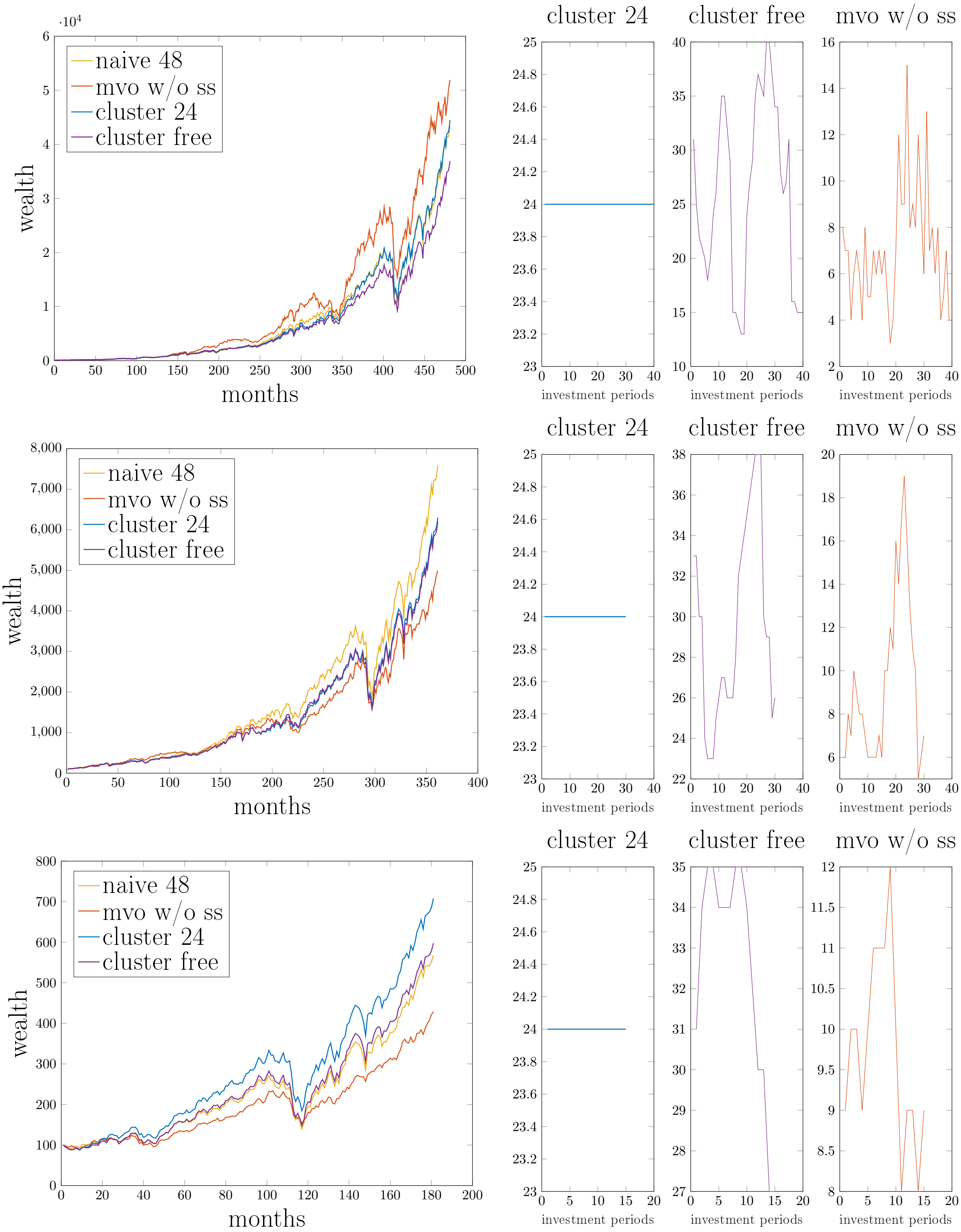


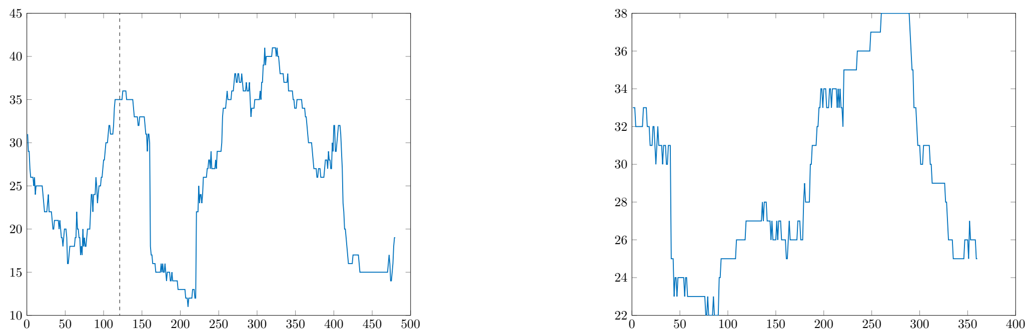
Figure 5.1: Real data: Portfolio prices starting at 100 for the Pearson cluster portfolio with 24 and free number of assets, the naive portfolio with 48 assets and the mean-variance portfolio without shortselling. Number of assets in a portfolio. Estimation periods: 60, 180, and 360 months

180/360/annual	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq return	div-risk ratio
Pearson cluster: 24	15.14	15.72	96.47	153.91	211.27	3.41	13.90	141.78
1st period	9.81	13.37	73.99	121.11	171.81	92.31	8.92	225.60
2nd period	6.56	18.43	35.91	50.40	134.71	83.64	4.86	98.50
3rd period	22.13	15.81	141.13	263.95	274.68	1.82	20.88	110.68
Pearson cluster: free	15.13	15.99	94.75	151.98	209.21	1.51	13.85	133.32
1st period	10.56	14.03	75.89	124.44	173.96	31.55	9.57	206.13
2nd period	5.61	19.30	29.32	41.92	127.99	17.40	3.75	86.08
3rd period	22.28	15.70	143.16	268.86	279.58	1.48	21.05	109.49
naive: 48 assets	15.84	16.24	97.69	158.45	211.35	2.40	14.53	121.33
1st period	10.84	14.70	74.36	123.64	170.79	8.84	9.76	191.21
2nd period	5.27	18.87	28.17	39.93	126.18	1.16	3.49	85.52
3rd period	22.98	16.12	143.78	273.11	281.63	1.34	21.68	95.26
mvo w/o ss	14.21	14.86	95.71	159.64	206.28	2.00	13.10	134.44
1st period	5.38	14.05	38.64	61.19	133.55	4.67	4.40	170.79
2nd period	7.13	16.90	42.55	59.21	139.46	1.36	5.70	110.16
3rd period	20.26	18.12	112.75	210.20	240.76	3.80	18.61	107.67

Table 5.3: Real data: Annual performance of the Pearson cluster portfolio with 24 and free number of assets, the full naive portfolio and the sample mean-variance portfolio without shortselling.
estimation window: 180 months of daily data, test window: 12 months of monthly data, total testing period: 360 months

annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	2	2	2	3	2	4	2	1	2
Pearson cluster: free	3	3	4	4	3	1	3	3	4
naive: 48 assets	1	4	1	2	1	3	1	4	1
mvo w/o ss	4	1	3	1	4	2	4	2	3

Table 5.4: Real data: Ranking of the Pearson cluster portfolio with 24 and free number of assets, the full naive portfolio and the sample mean-variance portfolio without shortselling.
estimation window: 180 months of daily data, test window: 12 months of monthly data, total testing period: 360 months



(a) 60 months estimation period (b) 180 months estimation period

Figure 5.2: Real data: Monthly Pearson cluster with 60 and 180 estimation periods.

We observe that the clusters change over time, not only in size, but also in composition. In Figure 5.3, we see some pictures of the visualization of the cluster compositions. The rows of the pictures correspond to the estimation period used, 60 months in the first row and 180 months in the second row. The columns correspond to the same estimation period. Each picture can be explained as follows: Each column corresponds to one asset. The first row gives the colors

360/180/annual	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq return	div-risk ratio
Pearson cluster: 24	14.34	15.66	91.84	150.22	202.21	9.21	13.12	149.20
1st period	13.42	13.53	100.03	178.09	212.83	19.36	12.50	236.89
2nd period	7.36	18.13	40.93	58.64	139.58	5.53	5.72	102.47
3rd period	22.25	14.93	150.30	289.04	291.63	1.12	21.14	110.59
Pearson cluster: free	13.30	16.29	81.87	131.94	187.57	2.82	11.98	137.92
1st period	10.83	13.76	79.37	131.30	179.28	5.54	9.88	217.01
2nd period	6.29	19.07	33.28	47.97	131.68	1.48	4.48	90.68
3rd period	22.78	15.50	148.27	283.03	289.19	1.11	21.58	108.05
naive: 48 assets	13.03	16.70	78.26	125.80	181.28	3.95	11.64	126.16
1st period	10.84	14.70	74.36	123.64	170.79	8.84	9.76	191.21
2nd period	5.27	18.87	28.17	39.93	126.18	1.16	3.49	85.52
3rd period	22.98	16.12	143.78	273.11	281.63	1.34	21.68	95.26
mvo w/o ss	10.42	11.62	89.93	144.66	194.76	4.04	9.75	179.08
1st period	6.94	12.50	56.01	91.23	151.79	9.12	6.16	187.75
2nd period	6.27	12.15	52.06	71.60	149.50	1.65	5.53	147.42
3rd period	18.05	9.92	183.46	393.43	361.04	0.67	17.56	196.85

Table 5.5: Real data: Annual performance of the Pearson cluster portfolio with 24 and free number of assets, the full naive portfolio and the sample mean-variance portfolio without shortselling.
estimation window: 360 months of daily data, test window: 12 months of monthly data, total testing period: 180 months

annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	1	2	1	1	1	4	1	2	1
Pearson cluster: free	2	3	3	3	3	1	2	3	3
naive: 48 assets	3	4	4	4	4	2	3	4	4
mvo w/o ss	4	1	2	2	2	3	4	1	2

Table 5.6: Real data: Ranking of the Pearson cluster portfolio with 24 and free number of assets, the full naive portfolio and the sample mean-variance portfolio without shortselling.
estimation window: 360 months of daily data, test window: 12 months of monthly data, total testing period: 180 months

associated with the asset in this column. The rest of the column presents the cluster combination by the colors of each asset which is part of this cluster.

Thus, we observe that although the data consists of branch portfolios, there are still correlation structures present. Not only do they change over time, but they also depend on the estimation period. Note, that we actually produced a film with all cluster visualizations in which we saw periods where nearly every cluster changed. Moreover, we observed longer periods where nothing changed. This supports the idea of clustering the assets as well as the need for readjustment and rebalancing over time.

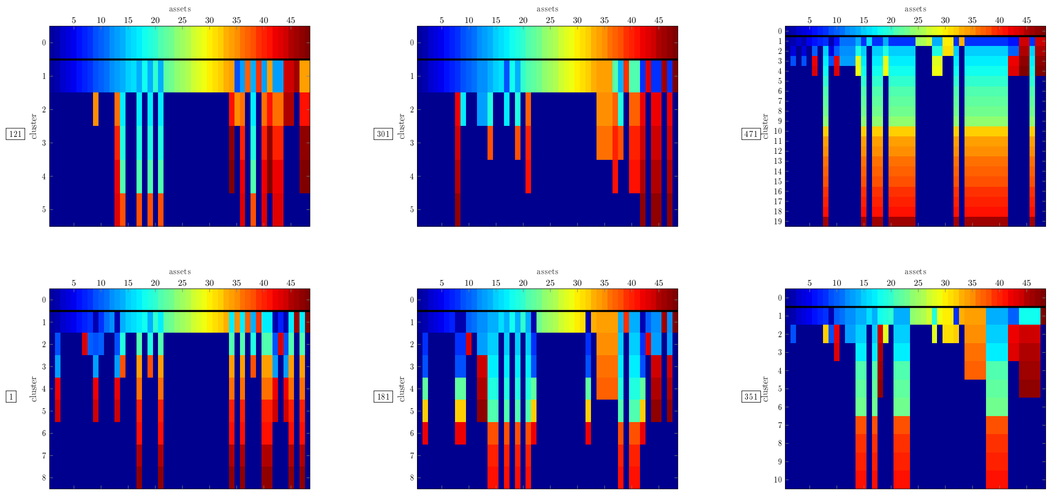


Figure 5.3: Examples of cluster changes. 60 months estimation period in row 1 and 180 months estimation period in row 2.

5.2.3 Simulated Data

Once more, we are going to examine the effects of the new cluster portfolio with simulated data. Thus, we turn to the simulation models 4 and 6, see Sections B.4 and B.6, which consider correlation structures. In Model 4, we have static parameters, and we have switching parameters in Model 6. For each model we have versions for 20, 24 and 28 groups. With this, we want to examine the impact of wrong cluster selection by the fixed cluster portfolio. Once the number of groups in the model is higher, once it is lower than the fixed number of 24 clusters and in one version it is equal.

Simulation Model 4

For Simulation Model 4 the ranking is stable over estimation periods for the 20 and 24 groups. For 28 groups there are only minor changes so that the over all ranking stays the same for the tested estimation periods. This model again favors the mean-variance methods, as the parameters are static, thus the constrained mean-variance strategy is always ranked first place. The full naive method is again forced to take all assets and takes last place. The cluster portfolios stay in the middle, where the variant with fixed number of assets ranks higher when the true number of groups is greater than half the number of assets. In addition, it ranks lower, when the true number of groups is lower than 24.

Simulation Model 4 with 20 groups

For the 20 groups case see Table 5.7 and the graphs in Figure 5.4. In general the mean-variance mean and CEQ return are greater then the corresponding values of the full naive portfolio for the 20 groups case. The corresponding values from the cluster portfolios are squeezed in between, nearly always greater than the naive and seldom greater than the mean-variance values. The cluster with free number of groups generates higher values than the free cluster method in over 90 percent of the cases. It is surprising that in about 80 percent of the cases the mean-variance strategy leads to a lower standard deviation compared to the other methods, even over

90 percent when using 360 months estimation period, whereas the naive strategy usually has higher standard deviation. This can be attributed to the fact that the constrained mean-variance portfolio is quite sparse and robust regarding slight parameter changes. The cluster methods lies in between, with the fixed cluster generating higher standard deviation than the free method in about 70 percent of the cases. The mean-variance portfolio usually has the highest values for the Sharpe, Sortino and omega ratio. In the 60 months estimation period it always has a higher omega ratio than the naive method. Again, the free cluster method has usually higher values than the fixed cluster method. Moreover, the naive strategy usually has the lowest values. The mean-variance portfolio nearly always produces the lowest turnover. Only for one data set its turnover is higher compared to the free cluster method. In over 80 percent of the cases the mean-variance strategy has the highest diversification-risk ratio, which increases over 90 percent in the 360 months case, whereas the full naive portfolio only rarely results in higher values. Again, the cluster portfolios lie between the other methods, where the free method is superior to the fixed one in this setting.

Estimation window: 60 months, testing window: 12 months, total testing period: 480 months. 1000 simulation sets.

annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	3 (0)	3 (49)	3 (2)	3 (2)	3 (0)	3 (0)	3 (0)	3 (18)	3
Pearson cluster: free	2 (28)	2 (139)	2 (96)	2 (116)	2 (20)	2 (1)	2 (27)	2 (166)	2
naive: 48 assets	4 (0)	4 (35)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4 (2)	4
mvo w/o ss	1 (972)	1 (777)	1 (902)	1 (882)	1 (980)	1 (999)	1 (973)	1 (814)	1

Estimation window: 180 months, testing window: 12 months, total testing period: 360 months. 1000 simulation sets.

annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	3 (0)	3 (31)	3 (2)	3 (4)	3 (1)	3 (0)	3 (0)	3 (17)	3
Pearson cluster: free	2 (32)	2 (96)	2 (81)	2 (83)	2 (30)	2 (1)	2 (32)	2 (116)	2
naive: 48 assets	4 (0)	4 (25)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4 (1)	4
mvo w/o ss	1 (968)	1 (848)	1 (917)	1 (913)	1 (969)	1 (999)	1 (968)	1 (866)	1

Estimation window: 360 months, testing window: 12 months, total testing period: 180 months. 1000 simulation sets.

annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	3 (1)	3 (17)	3 (3)	3 (2)	3 (1)	3 (0)	3 (1)	3 (13)	3
Pearson cluster: free	2 (37)	2 (63)	2 (44)	2 (42)	2 (27)	2 (1)	2 (36)	2 (89)	2
naive: 48 assets	4 (0)	4 (16)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4
mvo w/o ss	1 (962)	1 (904)	1 (953)	1 (956)	1 (972)	1 (999)	1 (963)	1 (898)	1

Table 5.7: Model 4.20: Annual ranking. Comparison of the cluster portfolio with 24 assets, the free cluster portfolio, the full naive portfolio and the constrained mean-variance portfolio.
Settings: Simulation model 4.20 with 48 correlated assets with static parameters and 20 groups. Estimation window: 60, 180 and 360 months, testing window: 12 months, total testing period: 480, 360 and 180 months.
The number in brackets after the ranking indicates how often the portfolio outperformed both other strategies in terms of this performance criterion.

An example for the performance values in Model 4 with 20 groups can be seen in Figure 5.4 with all performance criteria for the 180 months estimation period.

Simulation Model 4 with 24 groups

For the models with 24 groups the two cluster methods coincide and therefore we will show only three portfolios there, see Table 5.8 and the graphs in Figure 5.5. For all estimation periods the mean and the CEQ return of the mean-variance portfolio is always greater and the turnover

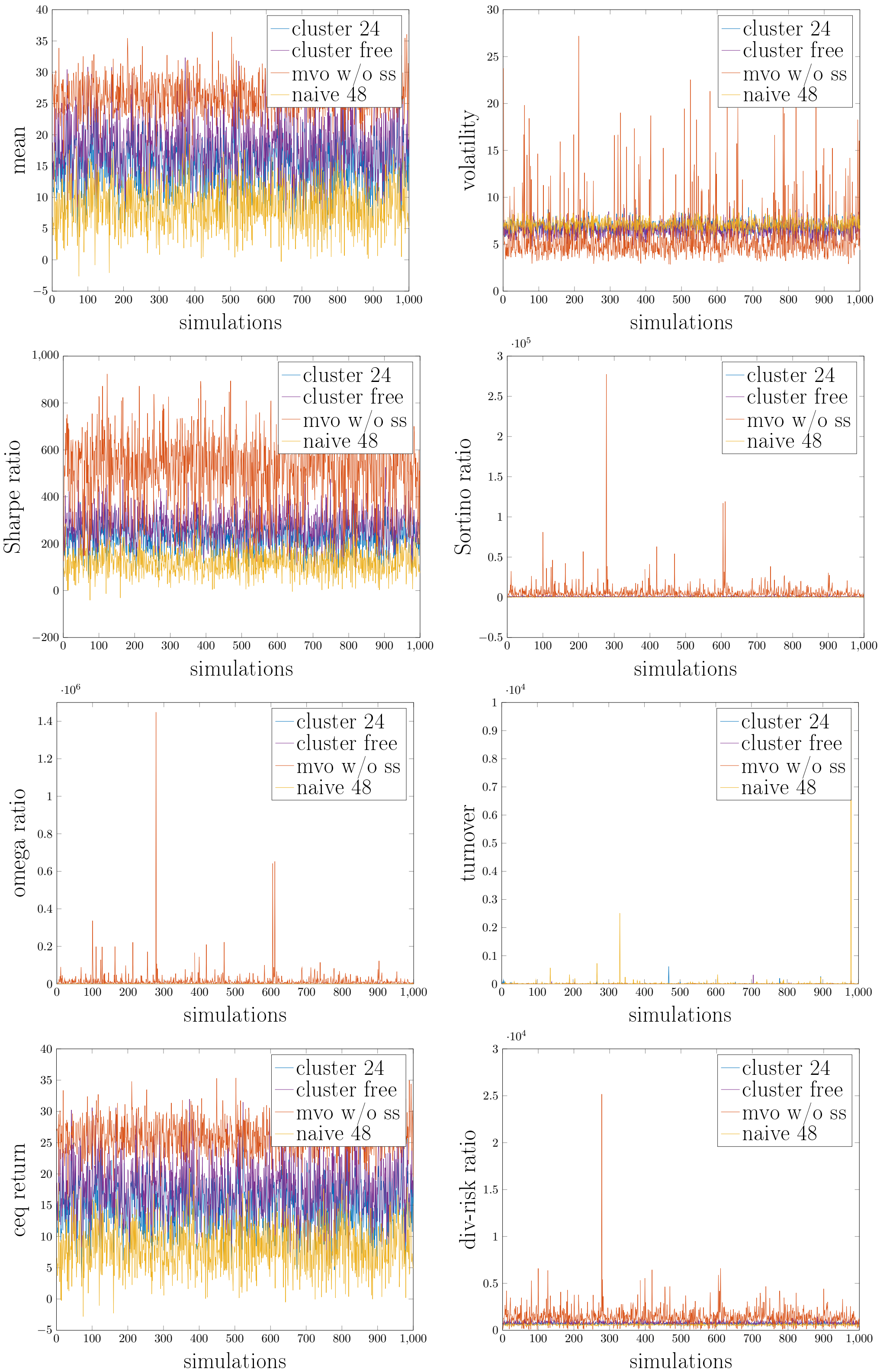


Figure 5.4: Model 4.20: Annual performance (in percent, except turnover). Simulation model 4, 48 assets, 20 correlation groups, 180 months estimation window, 360 months total testing period. 1000 simulation sets. (For more details, the corresponding graphs with only the first 100 simulations can be found in Appendix D)

always lower than the corresponding values of the naive portfolio. Also the cluster portfolio nearly always leads to higher values for the mean, Sharpe, Sortino and omega ratio, as well as the CEQ return and the diversification risk ratio. The omega ratio of the constrained mean-variance portfolio is always higher than the naive portfolio values in the 60 months case. In over 80 up to over 90 percent of the cases (with rising estimation length) the volatility of the mean-variance portfolio is smaller than the values of the other methods. The cluster portfolio has nearly always smaller standard deviation than the full naive portfolio. For 60 and 180 months estimation period the cluster turnover is always greater than the mean-variance turnover, and nearly always for 360 months. The naive turnover is nearly always greater than the cluster turnover.

Estimation window: 60 months, testing window: 12 months, total testing period: 480 months. 1000 simulation sets.

Annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster	2 (15)	2 (183)	2 (60)	2 (70)	2 (16)	2 (0)	2 (15)	2 (136)	2
naive: 48 assets	3 (0)	3 (1)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3
mvo w/o ss	1 (985)	1 (816)	1 (940)	1 (930)	1 (984)	1 (1000)	1 (985)	1 (864)	1

Estimation window: 180 months, testing window: 12 months, total testing period: 360 months. 1000 simulation sets.

Annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster	2 (18)	2 (120)	2 (46)	2 (48)	2 (16)	2 (0)	2 (17)	2 (84)	2
naive: 48 assets	3 (0)	3 (1)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3 (1)	3
mvo w/o ss	1 (982)	1 (879)	1 (954)	1 (952)	1 (984)	1 (1000)	1 (983)	1 (915)	1

Estimation window: 360 months, testing window: 12 months, total testing period: 180 months. 1000 simulation sets.

Annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster	2 (31)	2 (83)	2 (22)	2 (22)	2 (13)	2 (1)	2 (28)	2 (62)	2
naive: 48 assets	3 (0)	3 (1)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3
mvo w/o ss	1 (969)	1 (916)	1 (978)	1 (978)	1 (987)	1 (999)	1 (972)	1 (938)	1

Table 5.8: Model 4.24: Annual ranking. Comparison of the cluster portfolio, the full naive portfolio and the constrained mean-variance portfolio.
Settings: Simulation model 4.24 with 48 correlated assets with static parameters and 24 groups. Estimation window: 60, 180 and 360 months, testing window: 12 months, total testing period: 480, 360 and 180 months.
The number in brackets after the ranking indicates how often the portfolio outperformed both other strategies in terms of this performance criterion.

Simulation Model 4 with 28 groups

For the 28 groups case see Table 5.9, and the graphs in Figure 5.6. Here, we have lots of general statements assuring the superiority of the fixed cluster method: The mean, the Sharpe, Sortino and omega ratio as well as the CEQ return are always greater than the corresponding values of the naive portfolio. In addition, it always turns out that the mean-variance mean and the CEQ return are greater than the corresponding values of the naive portfolio. The mean-variance turnover always has the lowest value compared to all other methods in the 60 months case. For the other cases it is only true for the naive and the free cluster method and the relation towards the fixed cluster method is only nearly always true. Opposed to the 20 groups case, the fixed cluster portfolio outperforms the free cluster method in terms of all performance criteria except the diversification-risk ratio. Whereas the mean-variance method is superior to the cluster portfolios in about 90 percent of the cases only in terms of the mean and CEQ return. In around 80 percent of the cases the mean variance portfolio has higher omega ratios

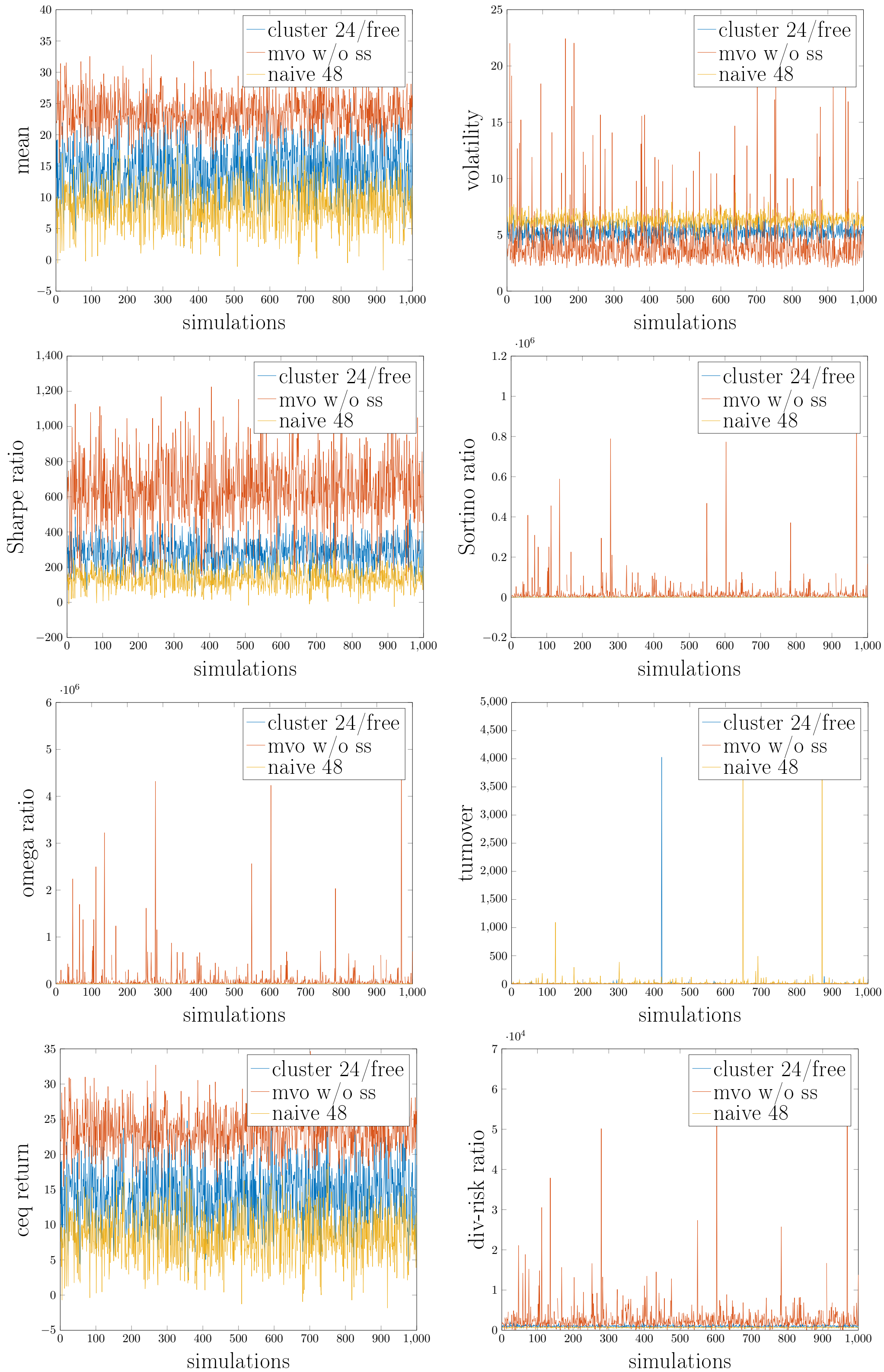


Figure 5.5: Model 4.24: Annual performance (in percent, except turnover). Simulation model 4, 48 assets, 24 correlation groups, 180 months estimation window, 360 months total testing period. 1000 simulation sets. (For more details, the corresponding graphs with only the first 100 simulations can be found in Appendix D)

than the cluster portfolios. It is only superior in around 50 up to around 60 percent of the cases in terms of the volatility, the Sharpe, Sortino and diversification-risk ratio for the 60 months case. However, it increases the performance for these criteria with rising estimation length up to around 80 percent of the cases.

Estimation window: 60 months, testing window: 12 months, total testing period: 480 months. 1000 simulation sets.

annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	2 (56)	3 (44)	1 (383)	1 (408)	2 (193)	2 (0)	2 (64)	2 (235)	2
Pearson cluster: free	3 (0)	1 (455)	3 (38)	3 (46)	3 (19)	3 (0)	3 (0)	1 (294)	3
naive: 48 assets	4 (0)	4 (40)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4 (2)	4
mvo w/o ss	1 (944)	2 (461)	2 (579)	2 (546)	1 (788)	1 (1000)	1 (936)	3 (469)	1

Estimation window: 180 months, testing window: 12 months, total testing period: 360 months. 1000 simulation sets.

annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	2 (69)	3 (63)	2 (288)	2 (290)	2 (164)	2 (3)	2 (72)	3 (180)	2
Pearson cluster: free	3 (1)	1 (293)	3 (25)	3 (28)	3 (17)	3 (0)	3 (1)	2 (194)	3
naive: 48 assets	4 (0)	4 (22)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4 (2)	4
mvo w/o ss	1 (930)	2 (622)	1 (687)	1 (682)	1 (819)	1 (997)	1 (927)	1 (624)	1

Estimation window: 360 months, testing window: 12 months, total testing period: 180 months. 1000 simulation sets.

annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	2 (104)	3 (41)	2 (153)	2 (147)	2 (116)	2 (8)	2 (104)	3 (99)	2
Pearson cluster: free	3 (1)	2 (152)	3 (27)	3 (29)	3 (21)	3 (0)	3 (1)	2 (114)	3
naive: 48 assets	4 (0)	4 (16)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4 (5)	4
mvo w/o ss	1 (895)	1 (791)	1 (820)	1 (824)	1 (863)	1 (992)	1 (895)	1 (782)	1

Table 5.9: Model 4.28: Annual ranking. Comparison of the cluster portfolio with 24 assets, the free cluster portfolio, the full naive portfolio and the constrained mean-variance portfolio.
Settings: Simulation model 4.28 with 48 correlated assets with static parameters and 28 groups. Estimation window: 60, 180 and 360 months, testing window: 12 months, total testing period: 480, 360 and 180 months.
The number in brackets after the ranking indicates how often the portfolio outperformed both other strategies in terms of this performance criterion.

Simulation Model 6
In Simulation Model 6 we have high correlation groups and annual changing parameters, see Appendix B.6 for details. We use the parameter switching to examine the effects of model uncertainty.

Simulation Model 6 with 20 groups
Again we start with the variant with 20 groups. Corresponding to the 60 months estimation period case in Table 5.10, we have no general statements. This can also be seen in Figure 5.7, the performance is similarly distributed for all portfolios, but the performance of the mean-variance portfolio is more overlapping in the 60 months case. The most remarkable figure is the superiority of the free cluster method in terms of the standard deviation, where it outperforms the other strategies in 70 percent of the cases. In over 50 percent of the cases it outperforms the other strategies by the diversification-risk ratio. Indeed, it outperforms each strategy in over 70 percent of the cases, the naive strategy in nearly 90 percent of the cases. In terms of other performance criteria the free method and the fixed cluster portfolio outperform each other nearly equally often, but always in favor of the free method. For the Sharpe, Sortino and omega ratio

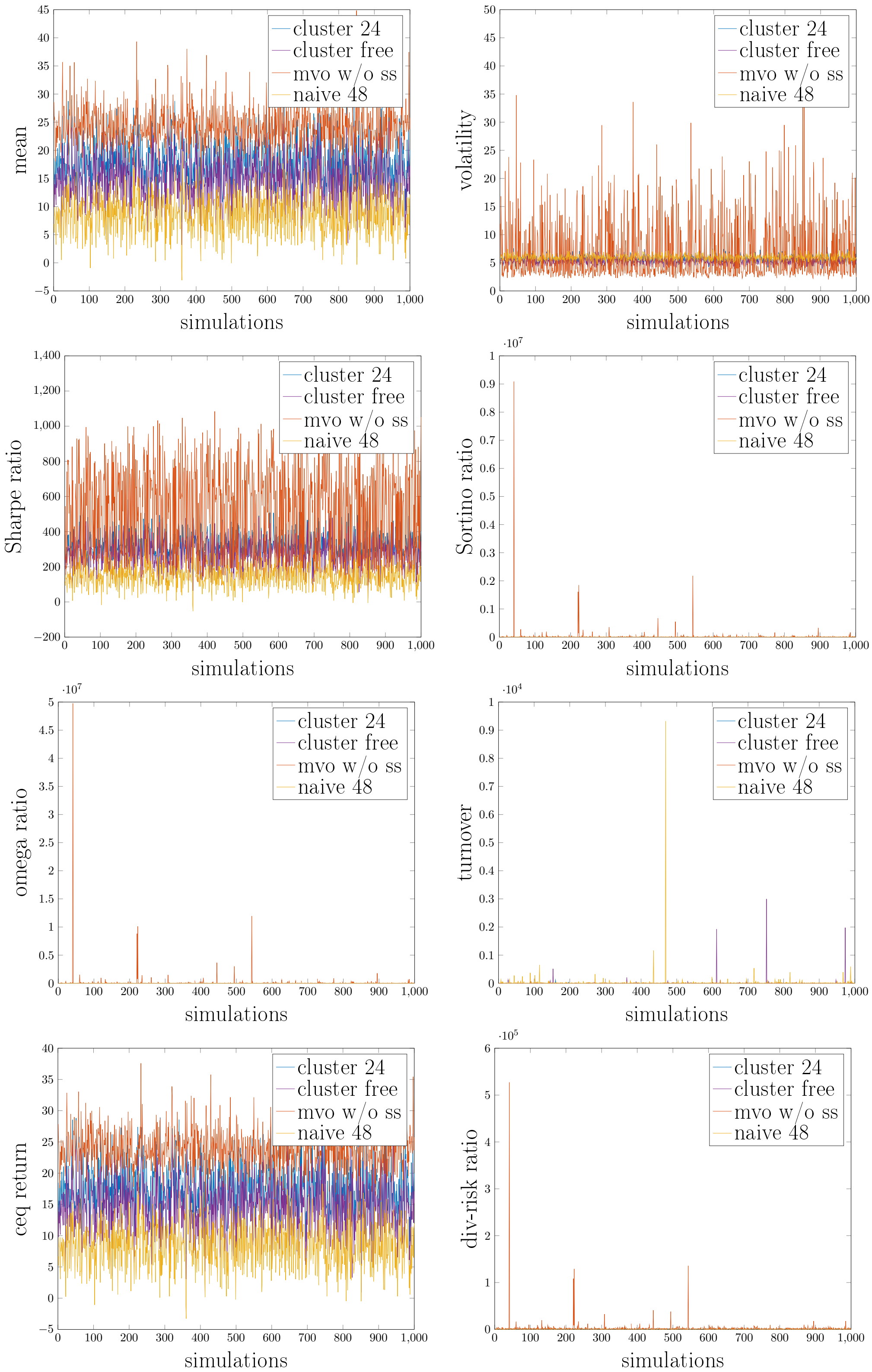


Figure 5.6: Model 4.28: Annual performance (in percent, except turnover). Simulation model 4, 48 assets, 28 correlation groups, 180 months estimation window, 360 months total testing period. 1000 simulation sets. (For more details, the corresponding graphs with only the first 100 simulations can be found in Appendix D)

Estimation window: 60 months, testing window: 12 months, total testing period: 480 months. 1000 simulation sets.

annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	4 (178)	2 (121)	2 (201)	2 (199)	2 (187)	3 (237)	4 (182)	2 (198)	3
Pearson cluster: free	3 (175)	1 (703)	1 (388)	1 (366)	1 (360)	4 (225)	3 (184)	1 (529)	1
naive: 48 assets	2 (290)	4 (22)	4 (175)	4 (167)	4 (163)	2 (270)	2 (283)	4 (62)	4
mvo w/o ss	1 (357)	3 (154)	3 (236)	3 (268)	3 (290)	1 (268)	1 (351)	3 (211)	2

Estimation window: 180 months, testing window: 12 months, total testing period: 360 months. 1000 simulation sets.

annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	3 (187)	3 (19)	3 (87)	3 (94)	3 (85)	2 (237)	4 (184)	3 (50)	3
Pearson cluster: free	4 (184)	2 (116)	2 (155)	2 (154)	2 (139)	4 (219)	3 (188)	2 (134)	2
naive: 48 assets	1 (313)	4 (5)	4 (73)	4 (77)	4 (66)	3 (254)	2 (296)	4 (13)	4
mvo w/o ss	2 (316)	1 (860)	1 (685)	1 (675)	1 (710)	1 (290)	1 (332)	1 (803)	1

Estimation window: 360 months, testing window: 12 months, total testing period: 180 months. 1000 simulation sets.

annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	2 (195)	3 (7)	3 (56)	3 (51)	3 (53)	2 (223)	2 (191)	3 (18)	2
Pearson cluster: free	4 (191)	2 (28)	2 (90)	2 (94)	2 (86)	3 (222)	4 (190)	2 (48)	3
naive: 48 assets	1 (309)	4 (3)	4 (50)	4 (54)	4 (51)	4 (253)	3 (293)	4 (14)	4
mvo w/o ss	3 (305)	1 (962)	1 (804)	1 (801)	1 (810)	1 (302)	1 (326)	1 (920)	1

Table 5.10: Model 6.20: Annual ranking. Comparison of the cluster portfolio with 24 assets, the free cluster portfolio, the full naive portfolio and the constrained mean-variance portfolio.
Settings: Simulation model 6.20 with 48 correlated assets with switching parameters and 20 groups. Estimation window: 60, 180 and 360 months, testing window: 12 months, total testing period: 480, 360 and 180 months.
The number in brackets after the ranking indicates how often the portfolio outperformed both other strategies in terms of this performance criterion.

the free cluster portfolio even outperforms the fixed version in over 60 percent of the cases. In the overall ranking the naive method takes the last place, whereas the mean-variance portfolio has fallen to the second place. The clear winner is the free cluster method, whereas the fixed cluster method takes third place.

For the 180 months estimation period case, we still cannot make general statements. However, we see that the increase in estimation period lets the mean-variance portfolio perform better than before. It is again on the first place in the overall ranking, whereas the free cluster portfolio must take the second place. There is a huge change in the performance of the mean-variance portfolio regarding the Sharpe, Sortino and omega ratio. There, it outperforms the fixed cluster method in nearly 80 percent of the cases, the free cluster portfolio in around 75 percent of the cases and the naive method in over 80 percent of the cases. For the diversification-risk ratio this increases to over 90 percent for the naive method, over 80 percent for the free cluster method and to nearly 90 percent for the fixed cluster portfolio. Both cluster portfolios perform quite equally only in terms of the diversification risk ratio the free method outperforms the fixed one in over 70 percent of the cases. For the other ratios this is true in at least 60 percent of the cases. In Figure 5.7 we can observe the superiority of the mean-variance strategy as the performance graph is farther away from the other performance graphs.

We still do not have general statements for the 360 months estimation period case, but the gap of performance between the mean-variance portfolio and the rest has increased slightly. However, the cluster portfolios interchanged their ranking. The equal behavior has now changed to be more in favor of the fixed cluster method. This is especially true for the mean, which also results in a higher CEQ return.

Simulation Model 6 with 24 groups

For 24 groups the cluster portfolios coincide and again we only compare three portfolios. Corre-

Estimation window: 60 months, testing window: 12 months, total testing period: 480 months. 1000 simulation sets.

Annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster	2 (284)	1 (569)	1 (536)	1 (528)	1 (471)	1 (387)	1 (301)	1 (524)	1
naive: 48 assets	1 (361)	3 (0)	3 (5)	3 (4)	3 (5)	3 (297)	3 (332)	3 (0)	3
mvo w/o ss	3 (355)	2 (431)	2 (459)	2 (468)	2 (524)	2 (316)	2 (367)	2 (476)	2

Estimation window: 180 months, testing window: 12 months, total testing period: 360 months. 1000 simulation sets.

Annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster	1 (277)	2 (35)	2 (97)	2 (92)	2 (81)	2 (320)	2 (288)	2 (46)	2
naive: 48 assets	2 (361)	3 (0)	3 (2)	3 (1)	3 (1)	3 (295)	3 (330)	3 (1)	3
mvo w/o ss	3 (362)	1 (965)	1 (901)	1 (907)	1 (918)	1 (385)	1 (382)	1 (953)	1

Estimation window: 360 months, testing window: 12 months, total testing period: 180 months. 1000 simulation sets.

Annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster	1 (302)	2 (3)	2 (35)	2 (47)	2 (39)	2 (315)	1 (302)	2 (18)	2
naive: 48 assets	2 (355)	3 (0)	3 (1)	3 (1)	3 (1)	3 (288)	3 (335)	3 (0)	3
mvo w/o ss	3 (343)	1 (997)	1 (964)	1 (952)	1 (960)	1 (397)	2 (363)	1 (982)	1

Table 5.11: Model 6.24: Annual ranking. Comparison of the cluster portfolio, the full naive portfolio and the constrained mean-variance portfolio.
Settings: Simulation model 6.24 with 48 correlated assets with switching parameters and 24 groups. Estimation window: 60, 180 and 360 months, testing window: 12 months, total testing period: 480, 360 and 180 months.
The number in brackets after the ranking indicates how often the portfolio outperformed both other strategies in terms of this performance criterion.

sponding to the 60 months estimation period case in Table 5.11, we find, that the naive volatility is always greater than the cluster volatility. Moreover, the diversification-risk ratio of the cluster portfolio is always greater than the diversification-risk ratio of the naive portfolio. With shorter estimation periods, the mean-variance portfolio falls again to the second place in the overall ranking. Now, the cluster and the mean-variance portfolio perform equally, in favor of the cluster portfolio for the mean, the volatility, the Sharpe and Sortino ratio as well as the CEQ return and the diversification-risk ratio. There is a clear superiority (about and over 90 percent of the cases) of both strategies over the naive portfolio in terms of the Sharpe, Sortino and omega ratio, as well as the diversification-risk ratio.

For the 180 months estimation period case, we have the general statement, that the volatility of the naive portfolio is always greater than the volatility of the other strategies. However, the mean-variance portfolio again takes first place in the overall ranking. Although the parameters change over time, it is likely to get the optimal portfolio. It is mainly superior in terms of the Sharpe, Sortino and omega ratio, as well as the diversification-risk ratio, in over 90 percent of the cases. This is where also both strategies are superior to the naive portfolio with over 90 percent. Now there are gaps visible between the performance graphs, see Figure 5.8. As before, the cluster method is squeezed in between the performances of the naive and the mean-variance portfolio.

Considering the 360 months estimation period case, we see that the naive volatility is always

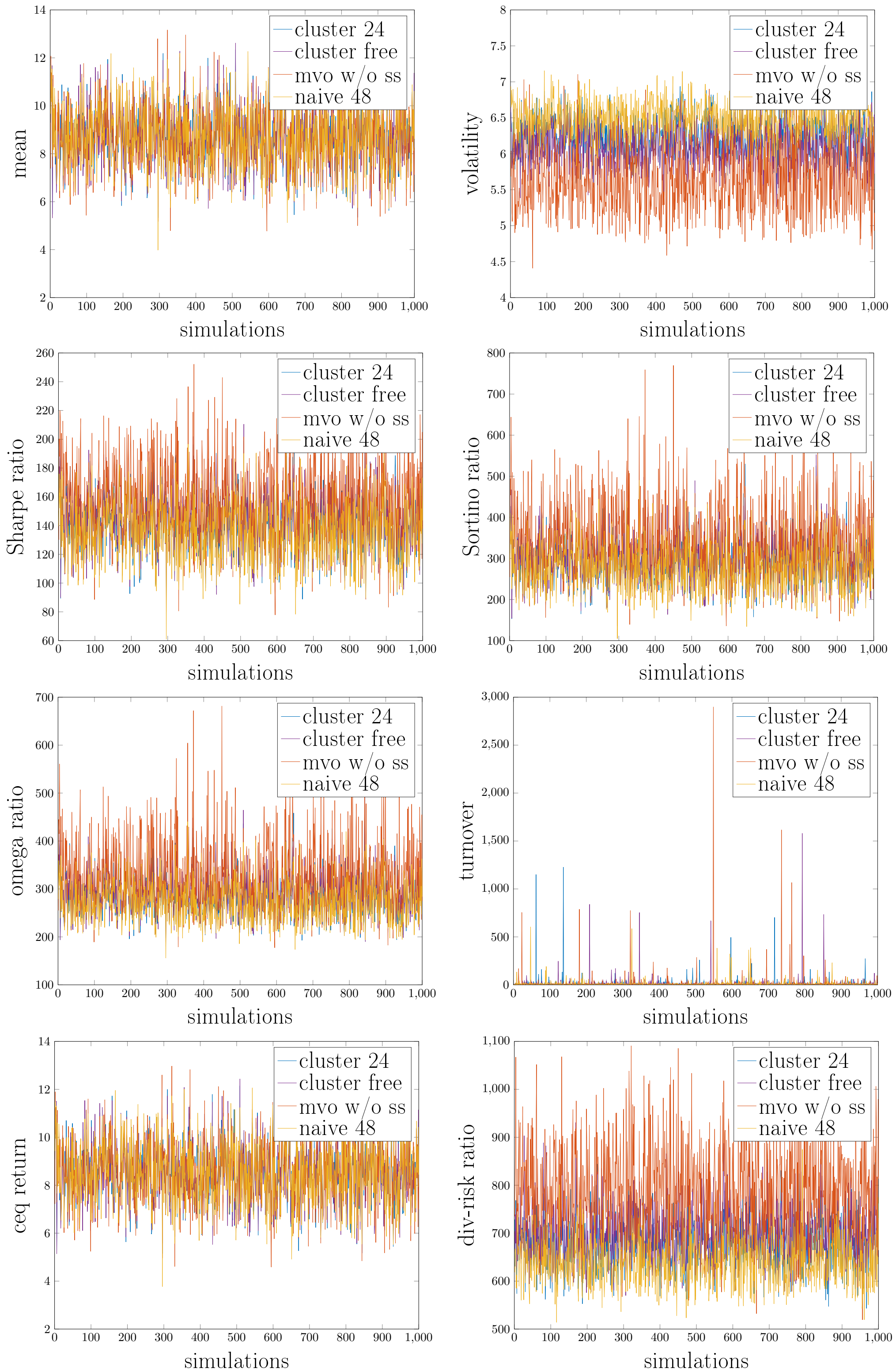


Figure 5.7: Model 6.20: Annual performance (in percent, except turnover). Simulation model 6, 48 assets, 20 correlation groups, 180 months estimation window, 360 months total testing period. 1000 simulation sets. (For more details, the corresponding graphs with only the first 100 simulations can be found in Appendix D)

greater than the mean-variance volatility. Like in the graphs for the 180 months case, there is not much difference in performance between the three methods concerning the mean, the turnover or the CEQ return. However, the gaps in the other cases rise slightly and support the superiority of the mean-variance portfolio over the cluster and the cluster over the naive portfolio.

Simulation Model 6 with 28 groups

For 28 groups see Table 5.12 and the graphs in Figure 5.9.

Estimation window: 60 months, testing window: 12 months, total testing period: 480 months. 1000 simulation sets.

annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	3 (198)	3 (3)	3 (26)	3 (34)	3 (23)	4 (215)	4 (191)	3 (29)	4
Pearson cluster: free	2 (181)	2 (335)	1 (371)	1 (371)	2 (332)	2 (235)	1 (188)	2 (327)	1
naive: 48 assets	1 (291)	4 (0)	4 (19)	4 (22)	4 (17)	1 (265)	3 (279)	4 (0)	3
mvo w/o ss	4 (330)	1 (662)	2 (584)	2 (573)	1 (628)	3 (285)	2 (342)	1 (644)	2

Estimation window: 180 months, testing window: 12 months, total testing period: 360 months. 1000 simulation sets.

annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	1 (212)	3 (1)	3 (5)	3 (7)	3 (2)	4 (192)	3 (199)	3 (1)	3
Pearson cluster: free	3 (174)	2 (10)	2 (62)	2 (61)	2 (52)	2 (213)	2 (179)	2 (31)	2
naive: 48 assets	2 (290)	4 (0)	4 (4)	4 (5)	4 (4)	3 (253)	4 (275)	4 (0)	4
mvo w/o ss	4 (324)	1 (989)	1 (929)	1 (927)	1 (942)	1 (342)	1 (347)	1 (968)	1

Estimation window: 360 months, testing window: 12 months, total testing period: 180 months. 1000 simulation sets.

annual ranking	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	3 (203)	3 (1)	3 (5)	3 (4)	3 (3)	4 (190)	3 (187)	3 (0)	3
Pearson cluster: free	1 (177)	2 (1)	2 (26)	2 (27)	2 (23)	2 (226)	2 (184)	2 (13)	2
naive: 48 assets	4 (296)	4 (0)	4 (9)	4 (13)	4 (10)	3 (223)	4 (279)	4 (1)	4
mvo w/o ss	2 (324)	1 (998)	1 (960)	1 (956)	1 (964)	1 (361)	1 (350)	1 (986)	1

Table 5.12: Model 6.28: Annual ranking. Comparison of the cluster portfolio with 24 assets, the free cluster portfolio, the full naive portfolio and the constrained mean-variance portfolio.
Settings: Simulation model 6.28 with 48 correlated assets with switching parameters and 28 groups. Estimation window: 60, 180 and 360 months, testing window: 12 months, total testing period: 480, 360 and 180 months.
The number in brackets after the ranking indicates how often the portfolio outperformed both other strategies in terms of this performance criterion.

For the 60 months estimation period we have the following: The naive volatility is always greater than the cluster free volatility. Again, the cluster performances are in the middle of the corresponding performances for the naive and the mean-variance portfolio. This is obvious for the volatility, the Sharpe, Sortino and omega ratio, as well as for the diversification risk ratio. The free cluster method is slightly superior to the fixed cluster method in terms of the mean and the CEQ return. However, in around 90 percent of the cases the free method leads to better performance concerning the volatility, the Sharpe, Sortino and omega ratio, as well as the diversification-risk ratio. For the mean, the turnover and the CEQ return, no method is definitely better than the other. Still, we have an overall ranking which puts the free cluster method in first place, the mean-variance strategy in second place and the naive method slightly before the fixed cluster method in third place.

Corresponding to the 180 months estimation period, we have no general statements. With the larger estimation windows the mean-variance strategy regains its strength and outperforms

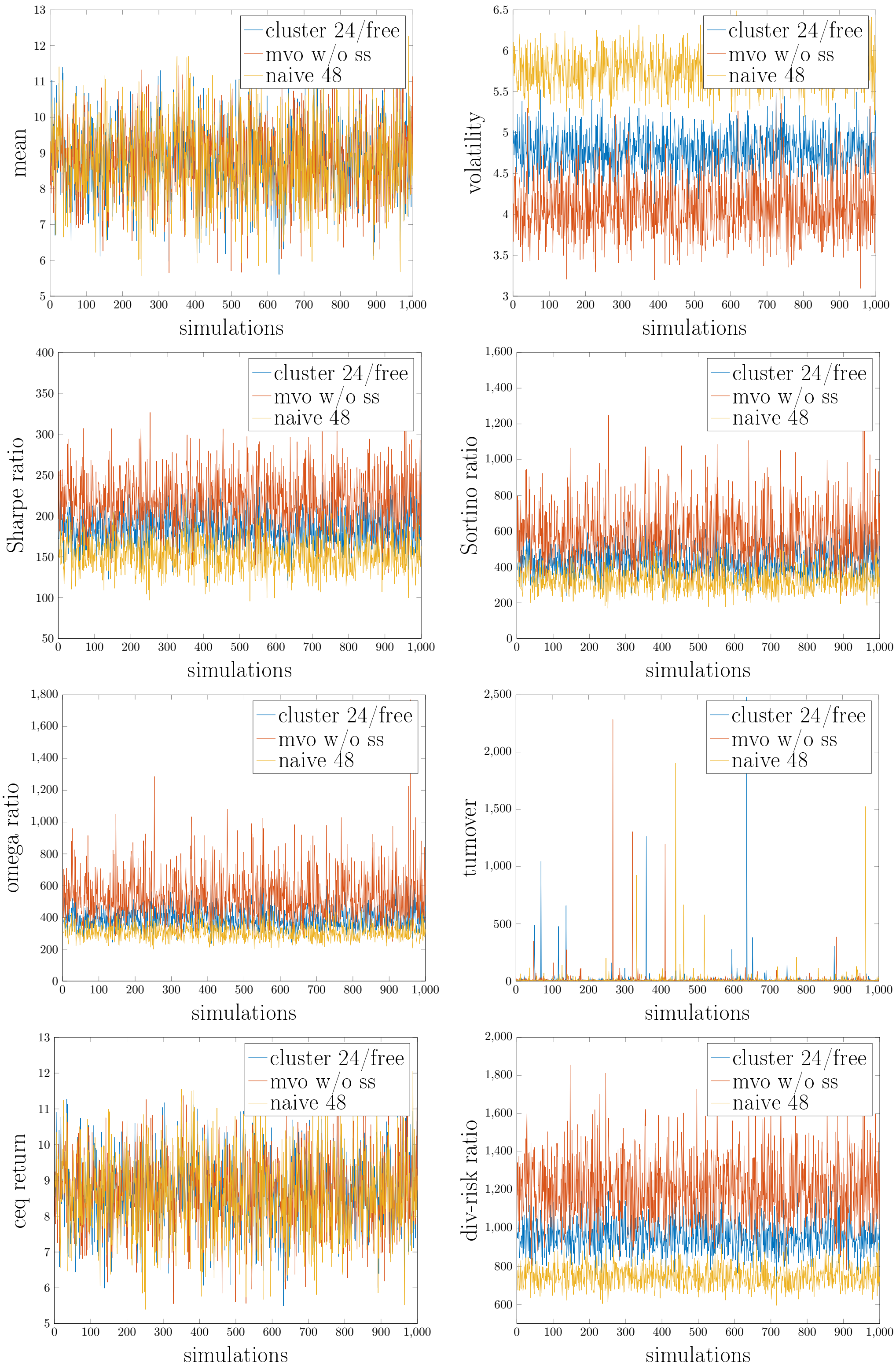


Figure 5.8: Model 6.24: Annual performance (in percent, except turnover). Simulation model 6, 48 assets, 24 correlation groups, 180 months estimation window, 360 months total testing period. 1000 simulation sets. (For more details, the corresponding graphs with only the first 100 simulations can be found in Appendix D)

the other strategies in terms of the volatility, the Sharpe, Sortino and omega ratio, as well as the diversification risk ratio. Thus, it reaches the first place in the overall ranking, followed by the free and the fixed cluster method and the naive portfolio again takes last place. The difference between the cluster portfolios is evident for the Sharpe, Sortino, omega and diversification-risk ratio, where the free method outperforms the fixed one in over 80 percent of the cases. However, if the true number of clusters lies around half the number of assets then the free cluster method leads to better performance, sometimes even better than the mean-variance portfolio.

There we see, like in the setting with independent assets, that the length of estimation period plays a great role. Even when the performance of the cluster portfolios is superior for the 60 months case, this is lost for longer estimation periods. As already discussed before, long estimation periods are non-realistic. Either they are not available or there is too much change in the assets model to have a chance to estimate the right moments. Even if we tried to get realistic settings with Markov parameters, the state space would still be finite and could be detected by the mean-variance method.

In the case of 360 months estimation periods, we have: The naive volatility is always greater than the mean-variance volatility. However, the general structure stays the same, only the mean-variance portfolio can further improve its performance. An example for model 6.28 performance values over all simulations we see in Figure 5.9 with all performance criteria for the 180 months estimation period.

5.2.4 Concluding Remark

For the real data set, we observe that the fixed cluster method is among the best performing portfolios, on the first or second place. The cluster portfolio with free number of clusters stays close behind the fixed version. As the performance of all considered portfolios is good, it is often ranked on the last place. However, both cluster portfolios improve the performance of a naive portfolio with fewer assets compared to the naive portfolio with 24 assets from Chapter 4.

Within Simulation Model 4, with fixed parameters over time, the cluster methods clearly outperform the naive strategy. As this setting favors the mean-variance optimization, it is not possible to consistently outperform it. When the true number of correlation groups is smaller than the number of clusters in the fixed cluster portfolio, then the free version performs better. In the case with 28 groups the fixed cluster portfolio with 24 assets performs better than the free version.

In Model 6, we consider parameter uncertainty with switching parameters. Here, the free cluster method performs better than the fixed version, because it is able to find the relevant clusters. However, with increasing estimation period it falls from the first place to the second, right after the mean-variance portfolio. Only for the 360 months case in the simulation model with 20 groups it falls behind the fixed cluster portfolio. This may be due to losing diversification effects, when only choosing representatives from the clusters, whose number is smaller than half the number of assets.

As in the setting with independent assets, we see that the length of estimation period plays an important role. Even, when the performance of the cluster portfolios is superior for the 60 months case, this is lost for longer estimation periods. As already discussed before, long estimation periods are non-realistic. Either they are not available or there is too much change in the assets model to have a chance to estimate the right moments. Even if we tried to get realistic settings with Markov parameters, the state space is still finite and can be detected by the mean-variance method. In addition, the mean-variance optimization uses the whole covariance matrix

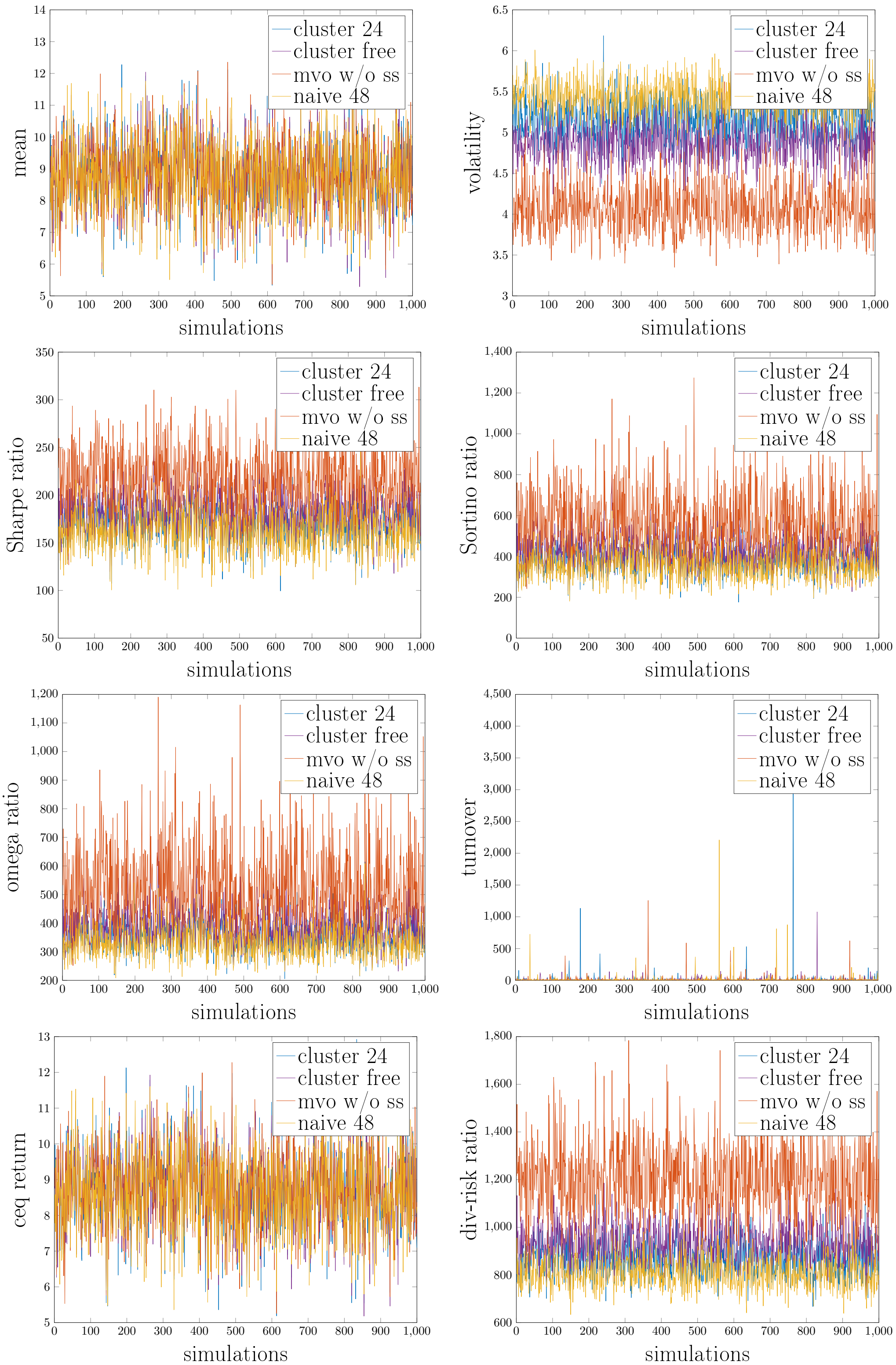


Figure 5.9: Model 6.28: Annual performance (in percent, except turnover). Simulation model 6, 48 assets, 28 correlation groups, 180 months estimation window, 360 months total testing period. 1000 simulation sets. (For more details, the corresponding graphs with only the first 100 simulations can be found in Appendix D)

and thus the portfolio also benefits from the dependence structure. In the end, we see that the cluster portfolio may be a way to improve the performance of a naive portfolio with fewer assets.

Surely, it depends on the specific setting, which cluster method should be chosen and how it performs in the end. However, the simulation results show that in all cases at least one cluster portfolio performs better than the worst performing benchmark, either the naive or the mean-variance portfolio. Thus, the clustering method is more robust. The choice of the cluster method seems to be dependent on the total number of available assets and the percentage of groups, which can be found in the data. If there are too few groups, then it is better to invest in a portfolio with a fixed number of clusters. Thus, the benefit of diversification from the full naive portfolio can be enhanced. If the detected number of clusters is close to the total number of available assets, then again it is better to choose the fixed version to enhance the benefits from selection.

In Section 5.3 we aim to explain why the cluster methods work, at least compared to the naive portfolio, as the constrained mean-variance portfolio has no theoretical closed form solution.

5.3 Comonotonicity

As an extreme case we can assume a market, where the assets are either comonotonic, which reduces to perfect positive correlation in our case of normally distributed returns, or independent. We will introduce the concept of comonotonicity in Section 5.3.1 and show some results for the diversification measure in Section 5.3.2. Then, we examine two cases for the distribution of the underlying returns. First, the standard multivariate normal setting (Section 5.3.3) and second, a mixed normal setting, where a Markov chain drives the mean of the assets (Section 5.3.4). Finally, we will see data experiments for the comonotonic setting in Section 5.4.

5.3.1 Concept of Comonotonicity

In finance one is often interested in the distribution of the sum of random variables. If independence between these random variables can be assumed then the calculation of the distribution of the sum is unproblematic. However, the assumption of independence is in most of the cases unrealistic. Including a dependence structure often makes it impossible or at least hard to find a solution. Additionally, in high dimensions (more than two assets), there is no good measure for correlation. This is where comonotonicity is used to get approximations, e.g. for analyzing bounds for the sum of random variables.

A derivation of such bounds can be found in Dhaene, Denuit, Goovaerts, Kaas, and Vyncke (2002). Some applications in finance are presented in Deelstra, Dhaene, and Vanmaele (2011).

Now, we will present a short introduction to the concept of comonotonicity following the paper by Dhaene, Denuit, Goovaerts, Kaas, and Vyncke (2002). To define a comonotonic random vector we first need the definition of a comonotonic support:

Definition 5.1 (Comonotonic Set). *The set $A \subseteq \mathbb{R}^n$ is said to be comonotonic if for any x and y in A, either $x \leq y$ or $y \leq x$ holds. Where $x \leq y$ is the componentwise order, meaning that $x_i \leq y_i$ for all $i = 1, 2, \dots n$.*

Then the definition of a comonotonic random vector is as follows.

Definition 5.2 (Comonotonic Vector). *A random vector $X = (X_1, \dots, X_n)$ is said to be comonotonic if it has a comonotonic support.*

This defines a very strong positive dependence structure on the components of the vector. Thus, two outcomes can be ordered by only comparing one component, as the others must follow the same ordering. So, if one component goes up the other components must go up as well and the other way round. This explains the wording comonotonic, as the components are commonly monotonic. Then, the definition of a comonotonic vector is equivalent to pairwise comonotonicity.

Theorem 5.3. *A random vector X is comonotonic if and only if (X_i, X_j) is comonotonic for all $i \neq j \in \{1, 2, \ldots, n\}$.*

Proof. See proof of Theorem 4 in Dhaene, Denuit, Goovaerts, Kaas, and Vyncke (2002). □

If we stick to the world of same location-scale family, the definition of comonotonic random vector reduces to pairwise perfect positive correlation.

Definition 5.4. *The random vector X has marginal cummulative distribution functions (cdf) F_{X_i}, that belong to the same location-scale family of distributions, if there exists a random variable Y, positive real constants a_i and real constants b_i. such that the relation*

$$X_i \stackrel{d}{=} a_i Y + b_i$$

holds for $i = 1, 2, \ldots, n$. Where $\stackrel{d}{=}$ denotes equality in distribution.

Theorem 5.5. *A random vector X with marginal cdf's F_{X_i} belonging to the same location-scale family, is comonotonic if and only if $\rho(X_i, X_j) = 1$ for all $i, j \in \{1, 2, \ldots, n\}$.*

Proof. See Theorem 5 in Dhaene, Denuit, Goovaerts, Kaas, and Vyncke (2002). □

Therefore, if we assume that the asset returns are all distributed within the same location-scale family as in our case the normal distribution, perfect positive correlation will be sufficient to introduce comonotonicity.

5.3.2 Diversification

Consider the diversification measure with correlation $H_D^N(w) = \sum_{i,j=1}^{N}(1 - \rho_{i,j})w_i w_j$ from Section 3.3, where we now use an upper index N or L to denote whether we work with N assets or with a strategy on L clusters. We can shown that in this setting diversification is larger in a smaller portfolio.

Proposition 5.6. *Given N assets in L comonotonic groups of size k_l for $l = 1, \ldots, L$ and considering the full naive portfolio with weight vector w_{ew} with $w_{ew,i} = \frac{1}{N}$, for $i = 1, \ldots, N$ and the cluster portfolio choosing one asset from each group and weight vector w_{cp} with $w_{cp,l} = \frac{1}{L}$, for $l = 1 \ldots L$, then*

$$H_D^L(w_{cp}) \geq H_D^N(w_{ew}).$$

Proof. Here the number of summands in $H_D^N(w_{ew})$ only depends on the $(N \times N)$ dissimilarity matrix.

$$D = (1 - \rho_{i,j})_{i,j} = \begin{cases} 0, & \rho_{i,j} = 1, \\ 1, & \rho_{i,j} = 0. \end{cases}$$

And D can be written in bloc matrix form:

$$D_{ew} = \begin{pmatrix} D_1 & & 1 \\ & \ddots & \\ 1 & & D_L \end{pmatrix},$$

where D_l for $l = 1, \dots, L$ are square matrices of size k_l and all entries equal zero. The rest of the matrix entries are one. The number of non-zero terms in H_D^N equals the number of non-zero entries in D which is:

$$N^2 - \sum_{l=1}^{L} k_l^2.$$

Now each non-zero term in $H_D^N(w_{ew})$ equals $w_i w_j = \frac{1}{N^2}$. The dissimilarity matrix for the cluster portfolio is just the $L \times L$ matrix with zero entries on the diagonal and the rest are ones. Thus, the number of non-zero terms in $H_D^L(w_{cp})$ is:

$$L^2 - L.$$

Each non-zero term in $H_D^L(w_{cp})$ is $w_i w_j = \frac{1}{L^2}$, thus we get:

$$H_D^L(w_{cp}) = (L^2 - L)\frac{1}{L^2} = 1 - \frac{1}{L}.$$

We split up the remainder of the proof in two cases:
Case 1: groups of equal size
With equal-sized groups we have with $k_l = k$ for $l = 1, \dots, L$ the equality $N = Lk$. Thus, we get:

$$\begin{aligned} H_D^N(w_{ew}) &= (N^2 - \sum_{l=1}^{L} k_l^2)\frac{1}{N^2} = (N^2 - Lk^2)\frac{1}{N^2} = 1 - \frac{Lk^2}{N^2} \\ &= 1 - \frac{k}{N} = 1 - \frac{1}{L} \\ &= H_D^L(w_{cp}). \end{aligned}$$

Case 2: groups of different sizes
For different sized groups we get:

$$H_D^N(w_{ew}) = (N^2 - \sum_{l=1}^{L} k_l^2)\frac{1}{N^2} = 1 - \sum_{l=1}^{L} \frac{k_l^2}{N^2}.$$

Now, we want to show that $H_D^L(w_{cp}) \geq H_D^N(w_{ew})$. As all variables are greater zero, this is equivalent to showing

$$\frac{1}{L} \leq \sum_{l=1}^{L} \frac{k_l^2}{N^2}.$$

This inequality holds for all $L \geq 1$, we can see this by using the Cauchy-Schwarz inequality $\sum x_i y_i \leq (\sum x_i^2)(\sum y_i^2)(***)$.

$$N^2 = \left(\sum_{l=1}^{L} k_l\right)^2 = \left(\sum_{l=1}^{L} 1 \cdot k_l\right)^2 \overset{(***)}{\leq} \left(\sum_{l=1}^{L} 1\right)\left(\sum_{l=1}^{L} k_l^2\right) = L\left(\sum_{l=1}^{L} k_l^2\right). \tag{5.1}$$

Equality only holds for $L = 1$, which is the case where all assets are comonotonic and both diversification measures are zero. Dividing (5.1) by $L > 0$ and by $N^2 > 0$ this leads to

$$\frac{1}{L} \leq \sum_{l=1}^{L} \frac{k_l^2}{N^2}.$$

□

5.3.3 Multivariate Normal Returns

We now analyze different settings, which can provide, despite of the comonotonic setting, an explanation for real market performance. First, we use a simple one period model with N multivariate normally distributed asset returns $X \sim \mathcal{N}(\mu, \Sigma)$. The set of assets can be divided into L subsets of size k_l for $l = 1, 2, ...L$. The assets in each group are comonotonic, i.e. $\text{corr}(X_i, X_j) = 1$ for all $i, j \in I_l$, where I_l is the index set of group l because the distributions of the assets belong to the same location and scale family. Assets from different groups are independent of each other, i.e. $\text{corr}(X_i, X_j) = 0$ for all $i \in I_l$ and $j \in I \backslash I_l$, where I is the index set of all assets ($I = \bigcup\limits_{l=1}^{L} I_l$). Further, we assume for simplicity that the variances of the assets are equal: $X_i \sim \mathcal{N}(\mu_i, \sigma^2)$, $i = 1, 2, ..., N$.

Now, we consider two portfolios, the equal weight portfolio Y_{ew} and the cluster portfolio Y_{cp}, where only one asset of each cluster is chosen with equal weight for the portfolio. The expected return of the portfolios is just the mean of the selected assets mean returns.

The variance of the $\frac{1}{N}$ portfolio can be derived by the variance of the sum of normal distributions. So, if $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, then $\sum_{i=1}^{N} X_i \sim \mathcal{N}(\sum_{i=1}^{N} \mu_i, \sum_{i=1}^{N} \sigma_i^2 + 2\sum_{i<j} \text{Cov}(X_i, X_j))$. Thus, we get:

$$\text{Var}_{ew} = \text{Var}\left[\frac{1}{N}\sum_{i=1}^{N} X_i\right] = \frac{1}{N^2}\left(\sum_{i=1}^{N} \sigma^2 + \sum_{l=1}^{L} 2 \sum_{i<j\in I_l} \text{Cov}(X_i, X_j)\right).$$

with I_l the index set of group l and $\text{Cov}(X_i, X_j) = \rho\sqrt{\text{Var}X_i}\sqrt{\text{Var}X_j} = \sigma^2$ for all assets within a group and zero otherwise. Then we need to know how many covariances are nonzero. Therefore, we calculate the number of nonzero covariances within a group. This is exactly the number of possible combinations of the elements in a group without ordering. Then, starting with the first element, we get $k_l - 1$ combinations with the other elements, the next element has only $k_l - 2$ possible combinations left, and so on. In total this leads to the number of $\sum_{i=1}^{k_l-1} k_l - i = (k_l - 1)k_l - \sum_{i=1}^{k_l-1} i = (k_l - 1)k_l - \frac{(k_l-1)k_l}{2} = \frac{(k_l-1)k_l}{2}$ nonzero covariances resulting from group l. Inserting this into the variance formula above we get:

$$\text{Var}_{ew} = \frac{1}{N^2}\left(N\sigma^2 + \sum_{l=1}^{L} (k_l - 1)k_l\sigma^2\right) = \frac{1}{N^2}\left(N\sigma^2 + \sigma^2\sum_{l=1}^{L} k_l^2 - \sigma^2\sum_{l=1}^{L} k_l\right).$$

Besides, we know that the sum of the group sizes equals the total number of assets: $N = \sum_{l=1}^{L} k_l$ Thus, we have:

$$\text{Var}_{ew} = \frac{\sigma^2}{N^2}\sum_{l=1}^{L} k_l^2. \tag{5.2}$$

However, in the cluster portfolio all assets are independent, i.e. the covariances are all equal to zero. Thus, we get:

$$\mathrm{Var}_{cp} = \mathrm{Var}\left[\frac{1}{L}\sum_{i\in I_{cp}} X_i\right] = \frac{1}{L^2}\left(\sum_{l=1}^{L}\sigma^2\right) = \frac{1}{L}\sigma^2, \tag{5.3}$$

where I_{cp} denotes the index set of the chosen assets of the cluster portfolio. We then consider two scenarios, one where the size of the groups is equal and one where at least one group has a different size. This allows to analyze how the variances of the two portfolios are related.

Proposition 5.7. *(Variance in Scenario 1: Groups of equal sizes)*
Given N multivariate normally distributed asset returns $X \sim \mathcal{N}(\mu, \Sigma)$ with equal variance $\sigma^2 \neq 0$, consisting of L comonotonic groups of equal size k, the variance of the full naive portfolio is equal to the variance of the cluster portfolio.

Proof. Let Var_{ew} be the variance of the full naive portfolio and Var_{cp} the variance of the cluster portfolio. In the setting of normally distributed assets we have that the sum is again normally distributed. Remember, then $\sum_{i=1}^{N} X_i \sim \mathcal{N}(\sum_{i=1}^{N}\mu_i, N\sigma^2 + 2\sum_{i<j}\mathrm{Cov}(X_i, X_j))$. From the variances in (5.2) and (5.3) and with the condition of equal-sized groups $Lk = N$ $(*)$ we get:

$$\mathrm{Var}_{ew} = \left(\frac{\sigma^2}{N^2}\sum_{l=1}^{L} k_l^2\right) = \left(\frac{\sigma^2}{N^2}Lk^2\right) \overset{(*)}{=} \frac{\sigma^2}{N^2}Nk \overset{(*)}{=} \frac{1}{L}\sigma^2 = \mathrm{Var}_{cp}.$$

□

Proposition 5.8. *(Variance in Scenario 2: Groups of different sizes)*
Given N multivariate normally distributed asset returns $X \sim \mathcal{N}(\mu, \Sigma)$ with equal variance $\sigma^2 \neq 0$, consisting of $L \geq 2$ comonotonic groups of sizes k_l for $l = 1, 2, ...L$, where $k_i \neq k_j$ for at least one pair $i \neq j$, the variance of the cluster portfolio is smaller than the variance of the full naive portfolio.

Proof. In this setting we get from equations (5.2) and (5.3) together with the condition for the group size $N = \sum_{l=1}^{L} k_l (**)$ the following:

$$\begin{aligned}
\mathrm{Var}_{cp} < \mathrm{Var}_{ew} &\iff \frac{1}{L}\sigma^2 < \left(\frac{1}{N^2}\sum_{l=1}^{L} k_l^2\right)\sigma^2, \\
&\overset{\sigma, L, N\neq 0}{\iff} N^2 < L\sum_{l=1}^{L} k_l^2, \\
&\overset{(**)}{\iff} \left(\sum_{l=1}^{L} k_l\right)^2 < L\sum_{l=1}^{L} k_l^2.
\end{aligned}$$

This inequality holds for all $L > 1$, we can see this by using the Cauchy-Schwarz inequality $\sum x_i y_i \leq (\sum x_i^2)(\sum y_i^2)(***)$,

$$\left(\sum_{l=1}^{L} k_l\right)^2 = \left(\sum_{l=1}^{L} 1\cdot k_l\right)^2 \overset{(***)}{\leq} \left(\sum_{l=1}^{L} 1\right)\left(\sum_{l=1}^{L} k_l^2\right) = L\left(\sum_{l=1}^{L} k_l^2\right), \tag{5.4}$$

and equality only holds for $L = 1$. Thus, $\text{Var}_{cp} < \text{Var}_{ew}$ holds for all $L > 1$. In consequence the variance of the cluster portfolio is smaller than the variance of the full naive portfolio if and only if the number of comonotonic groups is larger than one and they are not all of the same size. □

In the end, the variance of the cluster portfolio is always smaller or equal than the variance of the full naive portfolio. If $L = 1$ we are given only one comonotonic group and again the variances of both portfolios are equal. Here, the total number of assets is the same as the number of elements in the group, $N = k$. Thus, $\text{Var}_{ew} = \frac{1}{N^2}k^2\sigma^2 = \sigma^2 = \text{Var}_{cp}$.

Now we want to compare the two portfolio strategies by their Sharpe ratio. We can show under which circumstances the Sharpe ratio of the cluster portfolio is greater than the Sharpe ratio of the full naive portfolio.

Theorem 5.9. *(Sharpe ratio in Scenario 1: Groups of equal sizes)*
Given N multivariate normally distributed asset returns $X \sim \mathcal{N}(\mu, \Sigma)$ with equal variance $\sigma^2 \neq 0$, consisting of L comonotonic groups of equal size k, the Sharpe ratio of the full naive portfolio is smaller than the Sharpe ratio of the cluster portfolio if the mean return of the cluster portfolio is greater than the mean return of the full naive portfolio.

Proof. Let $R_{ew/cp} = \frac{\mu_{ew/cp}}{\sigma_{ew/cp}}$ denote the Sharpe ratios of the full naive portfolio and the cluster portfolio, respectively:

$$R_{ew} = \frac{1}{N\sigma_{ew}}\sum_{i=1}^{N}\mu_i,$$
$$R_{cp} = \frac{1}{L\sigma_{cp}}\sum_{i=1}^{N}\mu_i\mathbf{1}_{\{i\in I_{cp}\}}.$$

We know from Proposition 5.7 that in this case the variances of the portfolios σ_{ew}^2 and σ_{ew}^2 are the same, denoted by σ_P^2. Again, I_{cp} is the index set of the chosen assets in the cluster portfolio. To decide when the Sharpe ratio of the cluster portfolio is higher, we examine the difference $R_{cp} - R_{ew}$,

$$R_{cp} - R_{ew} = \frac{1}{\sigma_P}\left(\frac{1}{L}\sum_{i=1}^{N}\mu_i\mathbf{1}_{\{i\in I_{cp}\}} - \frac{1}{N}\sum_{i=1}^{N}\mu_n\right).$$

The term in brackets is just the difference of the mean returns of the two portfolios. Under the condition that the mean return of the cluster portfolio is greater than the mean return of the full naive portfolio this term is greater zero. Thus, the Sharpe ratio of the cluster portfolio is greater than the Sharpe ratio of the full naive portfolio. □

Theorem 5.10. *(Sharpe ratio in Scenario 2: Groups of different sizes)*
Given N multivariate normally distributed asset returns $X \sim \mathcal{N}(\mu, \Sigma)$ with equal variance $\sigma^2 \neq 0$, consisting of $L > 1$ comonotonic groups of sizes k_l for $l = 1, 2, ...L$, where $k_i \neq k_j$ for at least one pair $i \neq j$. The Sharpe ratio of the cluster portfolio is greater than the Sharpe ratio of the full naive portfolio if the mean return of the cluster portfolio is greater than the mean return of the full naive portfolio.

Proof. Let $R_{ew/cp} = \frac{\mu_{ew/cp}}{\sigma_{ew/cp}}$ denote the Sharpe ratios of the full naive portfolio and the cluster portfolio, respectively. From equations 5.2 and 5.3 we already know the variances and we get the following Sharpe ratios:

$$R_{ew} = \frac{\frac{1}{N}\sum_{i=1}^{N}\mu_i}{\frac{\sigma}{N}\sqrt{\sum_{l=1}^{L}k_l^2}},$$
$$R_{cp} = \frac{\frac{1}{L}\sum_{i=1}^{N}\mu_i\mathbf{1}_{\{i\in I_{cp}\}}}{\frac{\sigma}{\sqrt{L}}}.$$

Then, we can calculate the difference,

$$\begin{aligned} R_{cp} - R_{ew} &= \frac{\frac{1}{L}\sum_{i=1}^{N}\mu_i\mathbf{1}_{\{i\in I_{cp}\}}}{\frac{\sigma}{\sqrt{L}}} - \frac{\frac{1}{N}\sum_{i=1}^{N}\mu_i}{\frac{\sigma}{N}\sqrt{\sum_{l=1}^{L}k_l^2}} \\ &= \frac{\frac{1}{LN}(\sum_{i=1}^{N}\mu_i\mathbf{1}_{\{i\in I_{cp}\}})(\sqrt{\sum_{l=1}^{L}k_l^2}) - \frac{1}{N\sqrt{L}}\sum_{i=1}^{N}\mu_i}{\frac{\sigma}{N\sqrt{L}}\sqrt{\sum_{l=1}^{L}k_l^2}}. \end{aligned}$$

From the proof of Proposition 5.8 we derive that $\frac{1}{N}\sqrt{\sum_{l=1}^{L}k_l^2} \overset{(5.4)}{\geq} \frac{1}{N}\frac{\sum_{l=1}^{L}k_l}{\sqrt{L}} \overset{(**)}{=} \frac{1}{\sqrt{L}}$. Using this in the equation above we get

$$\begin{aligned} R_{cp} - R_{ew} &\geq \frac{\frac{1}{\sqrt{L}}\left(\frac{1}{L}\sum_{i=1}^{N}\mu_i\mathbf{1}_{\{i\in I_{cp}\}} - \frac{1}{N}\sum_{i=1}^{N}\mu_i\right)}{\frac{\sigma}{N\sqrt{L}}\sqrt{\sum_{l=1}^{L}k_l^2}} \\ &= \frac{\frac{1}{L}\sum_{i=1}^{N}\mu_i\mathbf{1}_{\{i\in I_{cp}\}} - \frac{1}{N}\sum_{i=1}^{N}\mu_i}{\frac{\sigma}{N}\sqrt{\sum_{l=1}^{L}k_l^2}}. \end{aligned}$$

Then the difference of the Sharpe ratios is greater zero if and only if the numerator is greater zero. This is the case if the mean return of the cluster portfolio is greater than the mean return of the full naive portfolio: $\frac{1}{L}\sum_{i=1}^{N}\mu_i\mathbf{1}_{\{i\in I_{cp}\}} > \frac{1}{N}\sum_{i=1}^{N}\mu_n$. □

For the diversification-risk ratio we can make the following statements:

Theorem 5.11. *(Diversification-risk ratio in Scenario 1: Groups of equal sizes)*
Given N multivariate normally distributed asset returns $X \sim \mathcal{N}(\mu, \Sigma)$ with equal variance $\sigma^2 \neq 0$, consisting of L comonotonic groups of equal size k, the diversification-risk ratio of the full naive portfolio is smaller than the diversification-risk ratio of the cluster portfolio if the mean return of the cluster portfolio is greater than the mean return of the full naive portfolio.

Proof. We know from Proposition 5.7, that the variances of both portfolios coincide in this case. Thus, the corresponding cumulative distribution functions are the same except for a shift on the x-axis, thus there is no crossing of the graphs of the functions. Then it only depends on the difference in the mean return, which portfolio has the higher risk. Then we have three cases.

1. $\mu_{ew} > \mu_{cp}$
   Here, the cumulative distribution function $F_{ew}(x)$ of the equal weights portfolio is just a

right shift of the cumulative distribution function $F_{cp}(x)$ of the cluster portfolio. Therefore, $F_{ew}(x)$ is pointwise smaller than $F_{cp}(x)$ for all $x \in \mathbb{R}$. Thus,

$$\Omega\text{-risk}(P_{ew}) = \int_{-\infty}^{t_r} F_{ew}(x)\mathrm{d}x < \int_{-\infty}^{t_r} F_{cp}(x)\mathrm{d}x = \Omega\text{-risk}(P_{cp})$$

for all $t_r \in \mathbb{R}$.

2. $\mu_{ew} < \mu_{cp}$
Here, the cumulative distribution function $F_{ew}(x)$ of the equal weights portfolio is just a left shift of the cumulative distribution function $F_{cp}(x)$ of the cluster portfolio. Therefore, $F_{ew}(x)$ is pointwise greater than $F_{cp}(x)$ for all $x \in \mathbb{R}$. Thus,

$$\Omega\text{-risk}(P_{ew}) = \int_{-\infty}^{t_r} F_{ew}(x)\mathrm{d}x > \int_{-\infty}^{t_r} F_{cp}(x)\mathrm{d}x = \Omega\text{-risk}(P_{cp})$$

for all $t_r \in \mathbb{R}$.

3. $\mu_{ew} = \mu_{cp}$
Then, both distribution functions coincide and thus $\Omega\text{-risk}(P_{ew}) = \Omega\text{-risk}(P_{cp})$.

From Proposition 5.6 we know that the diversification measure of the equal weights portfolio is always smaller or equal than the diversification measure of the cluster portfolio. We can conclude that the diversification-risk measure of the equal weights portfolio is always smaller than the diversification-risk measure of the cluster portfolio if and only if the mean of the equal weights portfolio is smaller than the mean of the cluster portfolio,

$$DivRiskRatio(P_{ew}) = \frac{H_D^N(w_{ew})}{\Omega\text{-risk}(P_{ew})} < \frac{H_D^L(w_{cp})}{\Omega\text{-risk}(P_{cp})} = DivRiskRatio(P_{cp}),$$

since in the denominator $\Omega\text{-risk}(P_{ew}) > \Omega\text{-risk}(P_{cp})$, even if the diversification measures of both portfolios are equal. □

Theorem 5.12. *(Diversification-risk ratio in Scenario 2: Groups of different sizes)*
Given N multivariate normally distributed asset returns $X \sim \mathcal{N}(\mu, \Sigma)$ with equal variance $\sigma^2 \neq 0$, consisting of $L > 1$ comonotonic groups of sizes k_l for $l = 1, 2, ...L$, where $k_i \neq k_j$ for at least one pair $i \neq j$. Then the diversification-risk ratio of the full naive portfolio is smaller than the diversification-risk ratio of the cluster portfolio if the threshold is smaller than the value x_{inters}, for which both cumulative distribution functions intersect.

Proof. In this scenario, we already know from Proposition 5.8 that the variance of the cluster portfolio is smaller than the variance of the equal weights portfolio. Thus, the distribution functions intersect at

$$x = x_{inters} = \frac{\mu_{ew}\sigma_{cp} - \mu_{cp}\sigma_{ew}}{\sigma_{cp} - \sigma_{ew}},$$

where $\sigma_{ew/cp}$ denotes the standard deviation of the portfolio. The formula follows from the following equivalences, starting with the condition that both cumulative distribution functions have the same value:

$$F_{ew}(x) = F_{cp}(x),$$

$$\iff \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{x - \mu_{ew}}{\sigma_{ew}\sqrt{2}}\right)\right) = \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{x - \mu_{cp}}{\sigma_{cp}\sqrt{2}}\right)\right),$$

where $\operatorname{erf}(x) = \frac{1}{\sqrt{\pi}} \int_{-x}^{x} e^{-t^2} \mathrm{d}t$ is the Gauss error function. Then, this is equivalent to

$$\begin{aligned} \operatorname{erf}\left(\frac{x-\mu_{ew}}{\sigma_{ew}\sqrt{2}}\right) &= \operatorname{erf}\left(\frac{x-\mu_{cp}}{\sigma_{cp}\sqrt{2}}\right), \\ \iff \frac{x-\mu_{ew}}{\sigma_{ew}\sqrt{2}} &= \frac{x-\mu_{cp}}{\sigma_{cp}\sqrt{2}}, \\ \iff (x-\mu_{ew})\sigma_{cp}\sqrt{2} &= (x-\mu_{cp})\sigma_{ew}\sqrt{2}, \\ \iff x(\sigma_{cp}-\sigma_{ew}) &= (\mu_{ew}\sigma_{cp}-\mu_{cp}\sigma_{ew}). \end{aligned}$$

Then, dividing by the non-zero term $\sigma_{cp} - \sigma_{ew} < 0$ yields the formula for x_{inters}. After that, we see that the Ω-risk of the naive portfolio is greater than the Ω-risk of the cluster portfolio on the left side of the intersection. This follows from the fact that in this range the cdf of the naive portfolio is always above the cdf of the cluster portfolio, independent of the mean. We can use the same transformations as in the calculation of the intersection point, but this time for the inequality. There is no transformation changing the direction of the inequality sign until the last transformation, where we divide by the negative but nonzero term $\sigma_{cp} - \sigma_{ew} < 0$.

$$\begin{aligned} F_{ew}(x) &> F_{cp}(x), \\ &\vdots \\ \iff x(\sigma_{cp}-\sigma_{ew}) &> (\mu_{ew}\sigma_{cp}-\mu_{cp}\sigma_{ew}), \\ \iff x &< \frac{\mu_{ew}\sigma_{cp}-\mu_{cp}\sigma_{ew}}{\sigma_{cp}-\sigma_{ew}} = x_{inters}. \end{aligned}$$

Thus,

$$\Omega\text{-risk}(P_{ew}) = \int_{-\infty}^{t_r} F_{ew}(x)\mathrm{d}x > \int_{-\infty}^{t_r} F_{cp}(x)\mathrm{d}x = \Omega\text{-risk}(P_{cp})$$

for all $t_r < x_{inters}$. Together with the results from Proposition 5.8, this yields:

$$DivRiskRatio(P_{ew}, t_r) = \frac{H_D^N(w_{ew})}{\Omega\text{-risk}(P_{ew}, t_r)} < \frac{H_D^L(w_{cp})}{\Omega\text{-risk}(P_{cp}, t_r)} = DivRiskRatio(P_{cp}, t_r),$$

for all $t_r < x_{inters}$. □

Note that this statement can be extended. Until the intersection point the ordering of risk is clear, but it does not change instantly after this point. The relation holds until the integrals over both cdfs are equal, which is definitely on the right of the intersection point. Unfortunately, there is no theoretical closed form solution for this point. It can only be approximated, using estimators for the expected values such that $\mathrm{E}[(R_{ew} - x_\Omega)^-] = \mathrm{E}[(R_{ew} - x_\Omega)^-]$ holds for some $x_\Omega \in \mathbb{R}$, denoting the x-value for which the Ω-risks are equal.

5.3.4 Mixed Normal Returns

Now, we model parameter uncertainty about the mean return of an asset by using a discrete, irreducible, homogeneous Markov chain, which admits a stationary distribution v to select the mean in the current period. Assume we have S states for the mean represented by a vector $\mu = (\mu_1, \ldots, \mu_S)'$. Then the Markov chain is the process $s(t)$ defining which is the current mean. Thus, the returns of an asset in group l are modeled as follows: $X_i^l(t) = \mu_{s(t)} + \sigma\varepsilon_l$ for $t = 1, \ldots, T$ and $\varepsilon_l \sim \mathcal{N}(0,1)$ for $l = 1, \ldots, L$, and $\mathrm{E}[\mu_{s(t)}] = (A^t v_0)'\mu$ with transition matrix A, where each column sums up to one, and initial distribution vector $v_0 = (v_1^0, \ldots, v_S^0)'$. We assume $v_0 = v$, which has the interpretation that the chain runs for a long time, since for any v_0 we have $A^t v_0 \to v, (t \to \infty)$. For $v_0 = v$ we indeed have $A^t v_0 = v$ for all t. As before we assume equal variances. To evaluate the Sharpe ratios, we first need to calculate the portfolio variances in this setting.

Proposition 5.13. *(Variance)*
With the assumptions about the model above, the variance of the full naive and the cluster portfolio are as follows:

$$\mathrm{Var}_{ew}(t) = \frac{1}{N^2}\left(\sum_{s=1}^{S} v_s \left((\mu_s - \mu_X)^2 + \sigma^2\right)\right)\sum_{l=1}^{L} k_l^2, \tag{5.5}$$

$$\mathrm{Var}_{cp}(t) = \frac{1}{L}\left(\sum_{s=1}^{S} v_s \left((\mu_s - \mu_X)^2 + \sigma^2\right)\right), \tag{5.6}$$

where v_s is the s-th component of the distribution vector v, $\mu_X = \mathrm{E}[X_i^l] = \sum_{s=1}^{S} v_s\mu_s$ is the expected mean of one asset and k_l is the size of group l.

Proof. First, we calculate the variance of one asset at time t. Therefore, we use Proposition 1 by Timmermann (2000) for centered moments of Markov switching processes, which are started from the limit distribution,

$$\mathrm{Var}_X(t) = \mathrm{Var}[X_i(t)] = \mathrm{E}[(X_i(t) - \mu_X)^2] = \sum_{s=1}^{S} v_s \sum_{j=1}^{2} {}_2C_j\, \sigma^j\, b_j(\mu_s - \mu_X)^{2-j},$$

where ${}_iC_j = \frac{i!}{(i-j)j!}$ and $b_j = \prod_{h=1}^{\frac{j}{2}}(2h-1)$, for j even, and 0 else. Remark that $b_0 = 1$. Thus, we get

$$\mathrm{Var}_X(t) = \sum_{s=1}^{S} v_s \left((\mu_s - \mu_X)^2 + \sigma^2\right).$$

Then using equation 5.2 we get the variance for the full naive portfolio:

$$\mathrm{Var}_{ew}(t) = \frac{1}{N^2}\, Var_X(t)\, \sum_{l=1}^{L} k_l^2 = \frac{1}{N^2}\left(\sum_{s=1}^{S} v_s \left((\mu_s - \mu_X)^2 + \sigma^2\right)\right)\sum_{l=1}^{L} k_l^2.$$

Using equation 5.3 leads to the variance of the cluster portfolio:

$$\mathrm{Var}_{cp}(t) = \frac{1}{L}\, \mathrm{Var}_X(t) = \frac{1}{L}\left(\sum_{s=1}^{S} v_s \left((\mu_s - \mu_X)^2 + \sigma^2\right)\right).$$

□

Then, we can compute the Sharpe ratios of the portfolios.

Theorem 5.14. *(Sharpe ratio)*
Given N normally distributed asset returns $X_i^l(t) = \mu_{s(t)} + \sigma\varepsilon_l,\ i = 1, 2, ..., N;\ t = 1, 2, .., T$, consisting of $L > 1$ comonotonic groups of sizes k_l for $l = 1, 2, ...L$. Further, assume that the mean of the normal distribution is driven by a discrete, irreducible, homogeneous Markov chain $s(t)$, which admits a stationary distribution v with S equally likely states $\mu_s(t) \in \{\mu_1, \cdots, \mu_S\}$. Then A is the transition matrix, where all columns are equal to the initial distribution vector $v_0 = (v_1, \cdots, v_S)'$. As before, we assume $v_0 = v$. Then the Sharpe ratio of the cluster portfolio is greater than the Sharpe ratio of the full naive portfolio.

Proof. The variances of the portfolios are already given in (5.5) and (5.6). Therefore, it remains to calculate the expected returns for the portfolios after t jumps. Let $Y_{ew/cp}(t)$ denote the portfolio process at time t of the full naive portfolio and the cluster portfolio, respectively. The distribution of the possible mean returns for each asset after the t-th jump is the same as the initial distribution, thus the expected returns are independent from the time period. This leads to

$$\mathrm{E}[Y_{ew}(t)] = \frac{1}{N}\sum_{i=1}^{N}\mu_i(t) = \frac{1}{N}\sum_{i=1}^{N}\sum_{s=1}^{S} v_s\mu_s = \frac{1}{N}N\sum_{s=1}^{S} v_s\mu_s = \sum_{s=1}^{S} v_s\mu_s$$

for the naive portfolio, and we get the same result for the cluster portfolio, i.e.

$$\mathrm{E}[Y_{cp}(t)] = \frac{1}{L}\sum_{i=1}^{N}\mu_i(t)\mathbf{1}_{\{i\in I_{cp}\}} = \frac{1}{L}\sum_{i=1}^{N}(\sum_{s=1}^{S} v_s\mu_s)\mathbf{1}_{\{i\in I_{cp}\}} = \frac{1}{L}L\sum_{s=1}^{S} v_s\mu_s = \sum_{s=1}^{S} v_s\mu_s.$$

This yields the following Sharpe ratios:

$$R_{ew}(t) = \frac{\sum_{s=1}^{S} v_s\mu_s}{\frac{1}{N}\sqrt{\sum_{s=1}^{S} v_s\left((\mu_s-\mu_X)^2+\sigma^2\right)}\sqrt{\sum_{l=1}^{L} k_l^2}},$$
$$R_{cp}(t) = \frac{\sum_{s=1}^{S} v_s\mu_s}{\frac{1}{\sqrt{L}}\sqrt{\sum_{s=1}^{S} v_s\left((\mu_s-\mu_X)^2+\sigma^2\right)}},$$

in short, this is just

$$R_{ew}(t) = \frac{\sum_{s=1}^{S} v_s\mu_s}{\frac{1}{N}\sqrt{Var_X(t)\sum_{l=1}^{L} k_l^2}},$$
$$R_{cp}(t) = \frac{\sum_{s=1}^{S} v_s\mu_s}{\frac{1}{\sqrt{L}}\sqrt{Var_X(t)}}.$$

As the numerators are equal, one can decide which one is the greater Sharpe ratio by examining the denominators. From inequality (5.4) we know

$$\frac{1}{\sqrt{L}}\sqrt{\sum_{l=1}^{L} k_l^2} \geq \sum_{l=1}^{L} k_l = N.$$

Again, equality only holds for $L = 1$. Thus, we get for $N > 0, L > 1$:

$$\frac{1}{N}\sqrt{\sum_{l=1}^{L} k_l^2} \geq \sqrt{L} > \frac{1}{\sqrt{L}}.$$

This yields that the denominator of the Sharpe ratio of the equal weight portfolio is greater than the denominator of the Sharpe ratio of the cluster portfolio. Thus, the Sharpe ratio of the cluster portfolio is greater than the Sharpe ratio of the full naive portfolio for all t. □

In the special case of a uniform stationary distribution, where the initial distribution vector is $v_0 = (\frac{1}{S}, \cdots, \frac{1}{S})'$, the formulas for the Sharpe ratios simplify to:

$$R_{ew}(t) = \frac{\frac{1}{S}\sum_{s=1}^{S}\mu_s}{\frac{\sqrt{Var_X(t)}}{N}\sqrt{\sum_{l=1}^{L}k_l^2}},$$
$$R_{cp}(t) = \frac{\frac{1}{S}\sum_{s=1}^{S}\mu_s}{\frac{\sqrt{Var_X(t)}}{\sqrt{L}}}.$$

5.3.5 Interpretation of the Theoretical Results

Regarding the diversification measure, we could show that in case of comonotonic groups, the diversification measure of the cluster portfolio is larger than the diversification measure of the full naive portfolio.

In the case of multivariate normally distributed returns in comonotonic groups with static parameters and equal variances, we proved that the variance of the cluster portfolio is smaller or equal than the variance of the full naive portfolio. This leads to the result that we only need to achieve a higher mean of the cluster portfolio to outperform the naive portfolio with respect to the Sharpe ratio and the diversification risk ratio.

If in addition, the mean parameters are switching, then we were able to show that the Sharpe ratio of the cluster portfolio is greater than the Sharpe ratio of the naive portfolio.

Thus, we have shown when we can benefit from clustering in extreme comonotonic settings. Moreover, real data shows that these results can also explain the observations, when the setting is not that extreme. For example, we expect that choosing fewer assets by clustering can outperform the full naive portfolio, if the clusters are of different size and the variances are comparable.

5.4 Data Experiments

We use simulations with a comonotonic dependence structure to support the theoretical results. Although we needed to follow several conditions in the proofs, we will see that the results can be transferred to more general models.

5.4.1 Simulation Model 5: Comonotonicity and Static Parameters

First, we look at the performance of general simulations of the comonotonic model without changing parameters, using model 5, see Appendix B.5 for a detailed description. (Note, that in some simulations the cluster algorithm fails due to numeric reasons, therefore these cases are excluded.)

Simulation Model 5 with 20 groups
For the 20 groups case see Table 5.13 and the graphs in Figure 5.10. For the 60 months estimation period case, we have the following general statements: The mean-variance performance is better

Estimation window: 60 months, testing window: 12 months, total testing period: 480 months. 996 simulation sets.

ranking (996)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3
Pearson cluster: free	2 (2)	1 (472)	2 (0)	2 (0)	2 (0)	2 (3)	2 (2)	2 (211)	2
naive: 28 assets	4 (0)	4 (2)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4
mvo w/o ss	1 (994)	2 (522)	1 (996)	1 (996)	1 (996)	1 (993)	1 (994)	1 (785)	1

Estimation window: 180 months, testing window: 12 months, total testing period: 360 months. 998 simulation sets.

ranking (998)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	3 (0)	3 (3)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3 (2)	3
Pearson cluster: free	2 (3)	2 (365)	2 (0)	2 (1)	2 (0)	2 (0)	2 (3)	2 (273)	2
naive: 28 assets	4 (0)	4 (2)	4 (0)	4 (0)	4 (0)	4 (1)	4 (0)	4 (0)	4
mvo w/o ss	1 (995)	1 (628)	1 (998)	1 (997)	1 (998)	1 (997)	1 (995)	1 (723)	1

Estimation window: 360 months, testing window: 12 months, total testing period: 180 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	3 (0)	3 (2)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3 (1)	3
Pearson cluster: free	2 (8)	2 (360)	2 (0)	2 (1)	2 (1)	2 (1)	2 (7)	2 (290)	2
naive: 28 assets	4 (0)	4 (3)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4
mvo w/o ss	1 (992)	1 (635)	1 (1000)	1 (999)	1 (999)	1 (999)	1 (993)	1 (709)	1

Table 5.13: Model 5.20: Annual ranking. Comparison of the cluster portfolio with 24 assets, the free cluster portfolio, the full naive portfolio and the constrained mean-variance portfolio.
Settings: Simulation model 5.20 with 48 correlated assets with static parameters and 20 comonotonic groups. Estimation window: 60, 180 and 360 months, testing window: 12 months, total testing period: 480, 360 and 180 months.
The number in brackets after the ranking indicates how often the portfolio outperformed both other strategies in terms of this performance criterion.

than the performance of all other portfolios concerning the Sharpe, Sortino and omega ratio. Its performance is better than the naive portfolio and the fixed cluster portfolio with respect to the mean and CEQ return. The volatility of the fixed cluster portfolio is always greater than the volatility of the free cluster method. The turnover of the naive portfolio is always greater than the turnover of the mean-variance strategy. Moreover, the diversification-risk ratio of the free cluster portfolio is always greater than the one of the naive portfolio. Clearly, the naive portfolio performs the worst in this setting, almost always outperformed by other strategies. Whereas the mean-variance portfolio performs the best, outperforming the other strategies in almost all cases. The free cluster portfolio nearly always outperforms the fixed cluster method.

In the 180 months estimation window case, we have nearly the same general statements as in the 60 months case. The mean-variance performance is greater than the performance of all other portfolios concerning the Sharpe and omega ratio. Its performance is better than the naive portfolio and the fixed cluster portfolio with respect to the mean, the Sortino ratio and the CEQ return. Still, the mean-variance portfolio outperforms the other strategies in most of the cases. It only slightly loses compared to the free cluster method. Still the free cluster method is superior to the fixed cluster method, and the naive portfolio performs the worst here, as well.

For the 360 months estimation period case, we observe the following: Now, the mean-variance portfolio is superior to all other strategies only in terms of the Sharpe ratio. For the Sortino and the omega ratio it is still true compared to the naive and the fixed cluster method. Moreover, the CEQ return of the mean-variance strategy is always greater than the CEQ return of the naive portfolio alone. However, the free cluster method is now definitely better than the naive portfolio regarding the Sharpe and the omega ratio. Apart from that it nearly always outperforms the fixed cluster portfolio. Only the turnover is sometimes higher. In terms of the volatility and the diversification-risk ratio the mean-variance portfolio is only superior over the free cluster method in about 70 percent of the cases.

Simulation Model 5 with 24 groups

For the 24 groups case see Table 5.14 and the graphs in Figure 5.11.

Estimation window: 60 months, testing window: 12 months, total testing period: 480 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster	2 (0)	1 (584)	2 (0)	2 (1)	2 (1)	2 (0)	2 (0)	2 (274)	2
naive: 48 assets	3 (0)	3 (7)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3
mvo w/o ss	1 (1000)	2 (409)	1 (1000)	1 (999)	1 (999)	1 (1000)	1 (1000)	1 (726)	1

Estimation window: 180 months, testing window: 12 months, total testing period: 360 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster	2 (0)	2 (454)	2 (0)	2 (0)	2 (0)	2 (0)	2 (0)	2 (231)	2
naive: 48 assets	3 (0)	3 (4)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3
mvo w/o ss	1 (1000)	1 (542)	1 (1000)	1 (1000)	1 (1000)	1 (1000)	1 (1000)	1 (769)	1

Estimation window: 360 months, testing window: 12 months, total testing period: 180 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster	2 (2)	2 (455)	2 (0)	2 (1)	2 (0)	2 (0)	2 (2)	2 (280)	2
naive: 48 assets	3 (0)	3 (5)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3
mvo w/o ss	1 (998)	1 (540)	1 (1000)	1 (999)	1 (1000)	1 (1000)	1 (998)	1 (720)	1

Table 5.14: Model 5.24: Annual ranking. Comparison of the cluster portfolio, the full naive portfolio and the constrained mean-variance portfolio.
Settings: Simulation model 5.24 with 48 correlated assets with static parameters and 24 comonotonic groups. Estimation window: 60, 180 and 360 months, testing window: 12 months, total testing period: 480, 360 and 180 months.
The number in brackets after the ranking indicates how often the portfolio outperformed both other strategies in terms of this performance criterion.

We get the following general statements in the 60 months estimation period case. The mean-variance portfolio outperforms the naive portfolio for all criteria except the volatility and the diversification-risk ratio. However, it generates higher values for over 90 percent of the cases. It also outperforms the cluster portfolio in terms of the mean, the Sharpe ratio, the turnover and the CEQ return. However, it shows better performance for the Sortino and omega ratio in 99 percent of the cases, and for the diversification-risk ratio in over 70 percent of the cases. Moreover, the cluster diversification-risk ratio is always greater than the ratio of the naive portfolio. The mean-variance and the cluster portfolio share the same range of volatility, both outperforming the naive method in more than 90 percent of the cases.

For 180 months estimation periods, the mean variance portfolio can improve its performance even further and it outperforms both other strategies in terms of all criteria except the volatility and the diversification-risk ratio. Still, the diversification-risk ratio of the cluster portfolio is always greater than the one of the naive portfolio. The volatility of the naive portfolio is nearly always smaller than the volatility of the other strategies. Again, the mean-variance and the cluster portfolio share the same range of volatility. The cluster portfolio outperforms the naive method in over 90 percent of the cases for all performance criteria.

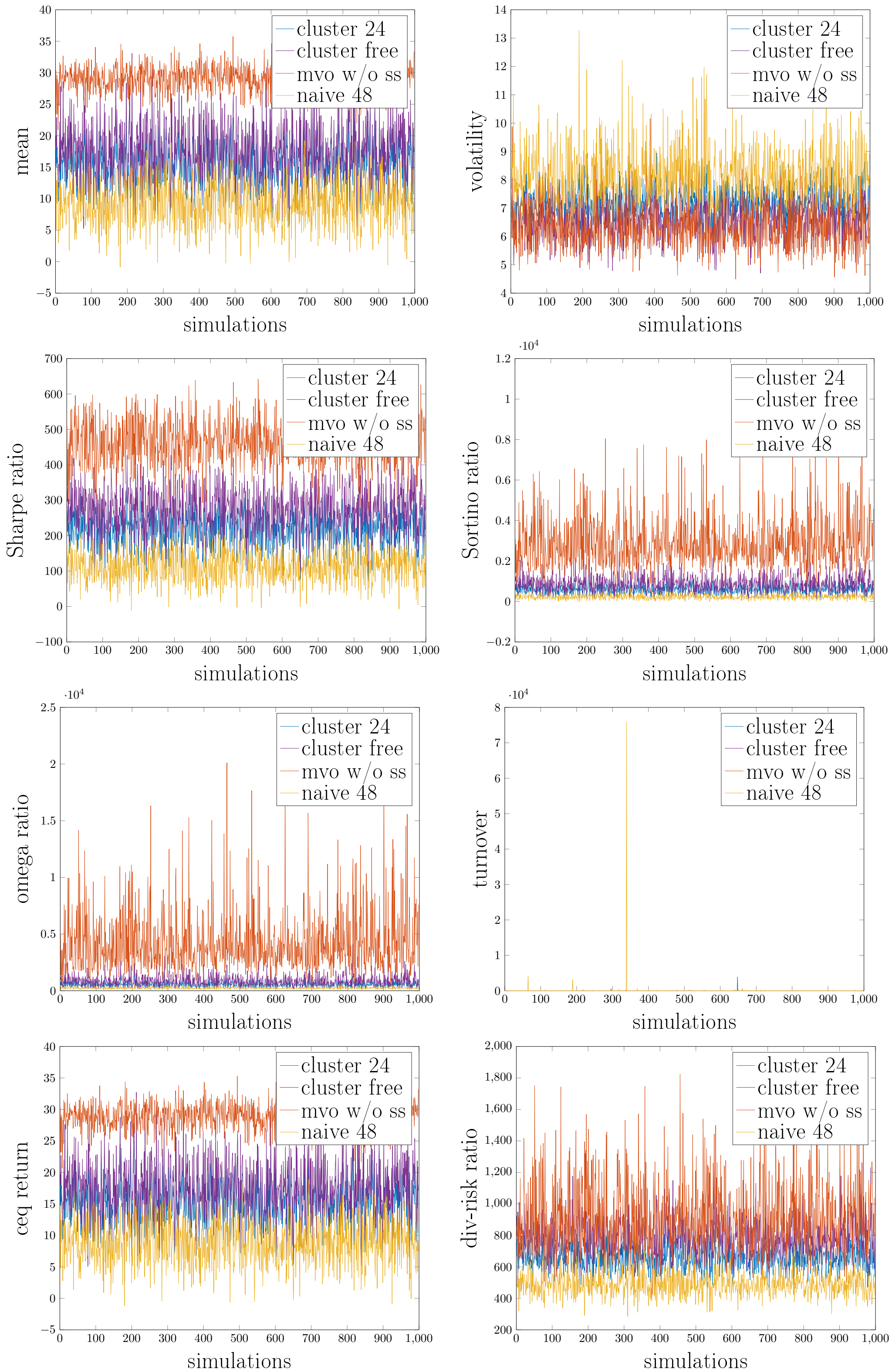


Figure 5.10: Model 5.20: Annual performance (in percent, except turnover). Simulation model 5, 48 assets, 20 correlation groups, 180 months estimation window, 360 months total testing period. 1000 simulation sets. (For more details, the corresponding graphs with only the first 100 simulations can be found in Appendix D)

In case of 360 months estimation period, the general statements of the superiority of the mean-variance portfolio over the naive portfolio stay the same. However, the generally better performance over the cluster portfolio is no longer true for the mean, the Sortino ratio and the CEQ return, but the percentage of outperformance is still over 90 percent. The relation for the standard deviation is also still valid, where the mean-variance and the cluster portfolio share the same range under the naive volatility. The diversification-risk ratio of the cluster portfolio results in higher values than the corresponding values of the naive portfolio, in over 90 percent of the cases. Moreover, the mean-variance portfolio has higher diversification-risk ratios in over 70 percent. Both strategies outperform the naive portfolio.

Simulation Model 5 with 28 groups

For the 28 groups case see Table 5.15 and the graphs in Figure 5.12. Corresponding to the 60

Estimation window: 60 months, testing window: 12 months, total testing period: 480 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	2 (0)	3 (0)	2 (0)	2 (0)	2 (0)	2 (0)	2 (0)	3 (7)	2
Pearson cluster: free	3 (0)	1 (766)	3 (1)	3 (2)	3 (1)	3 (0)	3 (0)	2 (304)	3
naive: 28 assets	4 (0)	4 (14)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4
mvo w/o ss	1 (1000)	2 (220)	1 (999)	1 (998)	1 (999)	1 (1000)	1 (1000)	1 (689)	1

Estimation window: 180 months, testing window: 12 months, total testing period: 360 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	2 (0)	3 (4)	2 (0)	2 (0)	2 (0)	2 (0)	2 (0)	3 (41)	2
Pearson cluster: free	3 (0)	1 (584)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	2 (181)	3
naive: 28 assets	4 (0)	4 (8)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4
mvo w/o ss	1 (1000)	2 (404)	1 (1000)	1 (1000)	1 (1000)	1 (1000)	1 (1000)	1 (778)	1

Estimation window: 360 months, testing window: 12 months, total testing period: 180 months. 999 simulation sets.

ranking (999)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	2 (1)	3 (3)	2 (0)	2 (0)	2 (0)	2 (0)	2 (1)	3 (66)	2
Pearson cluster: free	3 (0)	1 (460)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	2 (111)	3
naive: 28 assets	4 (0)	4 (7)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4
mvo w/o ss	1 (998)	2 (529)	1 (999)	1 (999)	1 (999)	1 (999)	1 (998)	1 (822)	1

Table 5.15: Model 5.28: Annual ranking. Comparison of the cluster portfolio with 24 assets, the free cluster portfolio, the full naive portfolio and the constrained mean-variance portfolio.
Settings: Simulation model 5.28 with 48 correlated assets with static parameters and 28 comonotonic groups. Estimation window: 60, 180 and 360 months, testing window: 12 months, total testing period: 480, 360 and 180 months.
The number in brackets after the ranking indicates how often the portfolio outperformed both other strategies in terms of this performance criterion.

months estimation window case, we have: The mean-variance portfolio always outperforms all other methods in terms of the mean, the turnover and the CEQ return. The fixed cluster portfolio and the naive portfolio are always outperformed by the mean-variance method regarding the Sharpe, Sortino and omega ratio, where the free cluster portfolio is nearly always outperformed. The volatility of the fixed cluster portfolio is always greater than the values of the free method. However, it always has the higher CEQ return, as the mean of the fixed cluster portfolio is always higher than the mean of the free cluster portfolio. It outperforms the free method in about 60 percent of the cases regarding the Sharpe, Sortino and omega ratio, as well as the turnover. The standard deviation of the naive portfolio is usually higher than the one of the other strategies starting from nearly 80 percent for the mean-variance method up to over 90

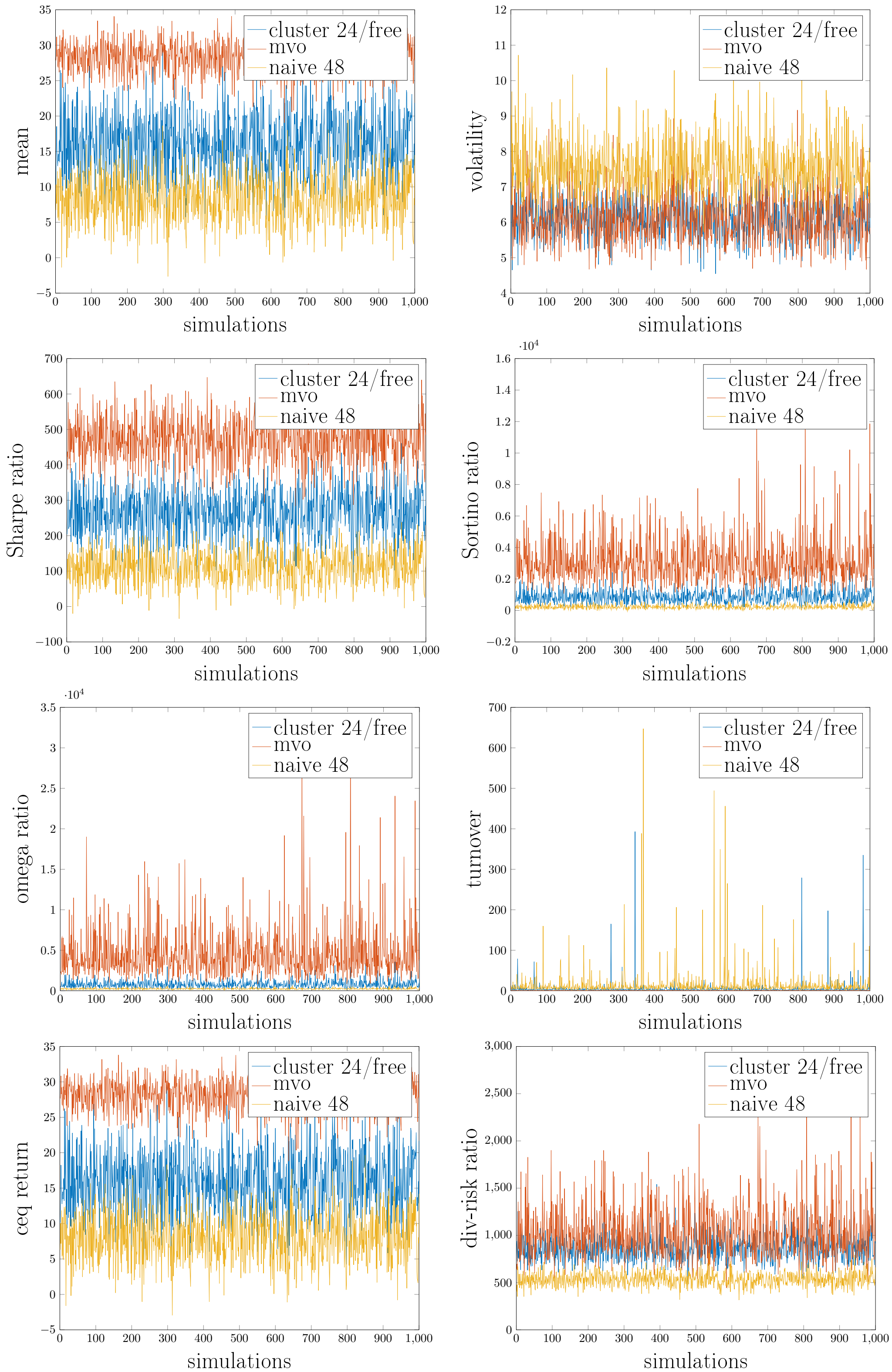


Figure 5.11: Model 5.24: Annual performance (in percent, except turnover). Simulation model 5, 48 assets, 24 correlation groups, 180 months estimation window, 360 months total testing period. 1000 simulation sets. (For more details, the corresponding graphs with only the first 100 simulations can be found in Appendix D)

percent for the free cluster method. The highest diversification-risk ratio is usually obtained by the mean-variance portfolio, only compared to the free cluster method which it outperforms in only nearly 70 percent of the cases. In contrast, the free cluster portfolio outperforms the fixed version and the naive portfolio in over 90 percent of the cases. Moreover, we see that the naive portfolio is nearly always outperformed by the other strategies.

For 180 months estimation periods, we see the following: With larger estimation period the mean-variance strategy dominates all other strategies in all categories, except the volatility and the diversification-risk ratio. For the diversification-risk ratio it usually outperforms the other strategies, in over 90 percent of the cases the naive portfolio, in nearly 80 percent the free cluster portfolio and in over 80 percent the fixed cluster portfolio. The free cluster method usually generates the smallest volatility only overlapping with the mean-variance portfolio. Moreover, the naive portfolio has nearly always the highest volatility. In this setting, the fixed cluster portfolio usually outperforms the free cluster, except for the volatility and the diversification-risk ratio.

Still we have that the mean-variance portfolio always or nearly always outperforms the other strategies except for the standard deviation and the diversification-risk ratio in the 360 months estimation period case. The fixed cluster method is usually superior to the free version, over 90 percent in terms of mean and CEQ return and around 70 percent for the Sharpe, Sortino and omega ratio and the turnover. For the volatility it generates nearly always a higher value than the free version. Moreover, it only outperforms the free version in 40 percent of the cases regarding the diversification-risk ratio. The naive portfolio performs the worst in almost all cases.

5.4.2 Simulation Model 7: Comonotonicity and Switching Parameters

In Simulation Model 7 we additionally have annual changing parameters of the assets, see Appendix B.7 for details. We use this again to examine the effects of model uncertainty.

Simulation Model 7 with 20 groups

Again we start with the variant with 20 groups, see Table 5.16. An example for the performance values in this model over all simulation sets and all performance criteria we see in Figure 5.13 for the 180 months estimation period. For 60 months estimation period, we have that the mean-variance portfolio is always better than the fixed cluster and the naive portfolio and nearly always better than the free cluster method in terms of the mean, the Sharpe, Sortino and omega ratio as well as the CEQ return. The cluster portfolios always outperform the naive portfolio in terms of the Sharpe, Sortino and omega ratio. Moreover, they nearly always outperform it in terms of the mean, the CEQ return and the diversification-risk ratio. The free cluster method nearly always generates better performance values than the fixed version, for all criteria, except for the turnover, where it is better in over 80 percent of the cases.

Corresponding to the 180 months estimation window case, we have that the mean-variance portfolio is always better than the fixed cluster and the naive portfolio and nearly always better than the free cluster method in terms of the mean, the Sharpe, Sortino and omega ratio, the turnover as well as the CEQ return. The free cluster portfolios always outperform the fixed cluster and the naive portfolio in terms of the Sharpe, Sortino and omega ratio, and nearly always for the mean, the volatility, the turnover, the CEQ return and the diversification risk ratio.

For 360 months estimation periods, we see that the mean-variance portfolio is always better

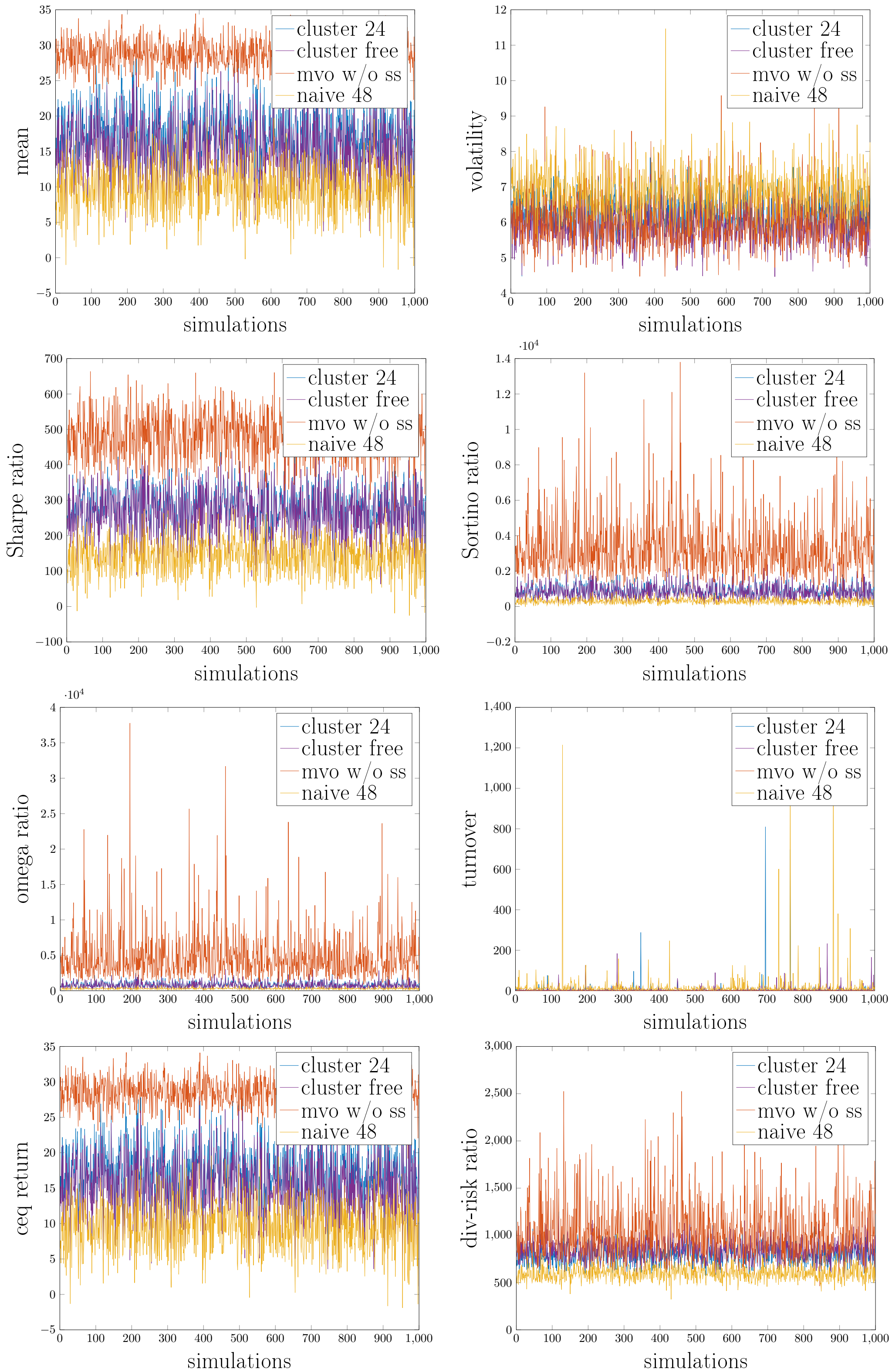


Figure 5.12: Model 5.28: Annual performance (in percent, except turnover). Simulation model 5, 48 assets, 28 correlation groups, 180 months estimation window, 360 months total testing period. 1000 simulation sets. (For more details, the corresponding graphs with only the first 100 simulations can be found in Appendix D)

Estimation window: 60 months, testing window: 12 months, total testing period: 480 months. 999 simulation sets.

ranking (999)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3 (1)	3 (0)	3 (0)	3
Pearson cluster: free	2 (3)	1 (877)	2 (12)	2 (8)	2 (5)	2 (12)	2 (3)	1 (497)	2
naive: 28 assets	4 (0)	4 (1)	4 (0)	4 (0)	4 (0)	4 (1)	4 (0)	4 (0)	4
mvo w/o ss	1 (996)	2 (121)	1 (987)	1 (991)	1 (994)	1 (985)	1 (996)	2 (502)	1

Estimation window: 180 months, testing window: 12 months, total testing period: 360 months. 999 simulation sets.

ranking (999)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	3 (0)	3 (1)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3
Pearson cluster: free	2 (2)	1 (813)	2 (2)	2 (3)	2 (3)	2 (3)	2 (2)	2 (475)	2
naive: 28 assets	4 (0)	4 (2)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4
mvo w/o ss	1 (997)	2 (183)	1 (997)	1 (996)	1 (996)	1 (986)	1 (997)	1 (524)	1

Estimation window: 360 months, testing window: 12 months, total testing period: 180 months. 999 simulation sets.

ranking (999)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	3 (0)	3 (2)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3 (5)	3
Pearson cluster: free	2 (2)	1 (672)	2 (1)	2 (3)	2 (1)	2 (0)	2 (1)	2 (408)	2
naive: 28 assets	4 (0)	4 (2)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4
mvo w/o ss	1 (997)	2 (323)	1 (998)	1 (996)	1 (998)	1 (999)	1 (998)	1 (586)	1

Table 5.16: Model 7.20: Annual ranking. Comparison of the cluster portfolio with 24 assets, the free cluster portfolio, the full naive portfolio and the constrained mean-variance portfolio.
Settings: Simulation model 7.20 with 48 correlated assets with switching parameters and 20 comonotonic groups. Estimation window: 60, 180 and 360 months, testing window: 12 months, total testing period: 480, 360 and 180 months.
The number in brackets after the ranking indicates how often the portfolio outperformed both other strategies in terms of this performance criterion.

than the fixed cluster and the naive portfolio and nearly always better than the free cluster method in terms of the mean, the Sharpe, Sortino and omega ratio as well as the turnover and the CEQ return. The free cluster method nearly always outperforms the fixed version and the naive portfolio for all criteria. Moreover, the naive strategy nearly always performs the worst.

Simulation Model 7 with 24 groups

For 24 groups the cluster portfolios coincide and again we only compare three portfolios, see Table 5.17 and Figure 5.14. Corresponding to the 60 months estimation period case, we have: The mean-variance portfolio always performs better than the naive portfolio, except for the volatility (only in 70 percent of the cases) and the diversification-risk ratio (in over 90 percent of the cases). It always performs better compared to the cluster portfolio in terms of the mean, the turnover and the CEQ return. It nearly always performs better for the Sharpe, Sortino and omega ratio. However, considering the volatility it nearly never outperforms the cluster method. Moreover, the cluster portfolio outperforms the mean-variance portfolio in over 70 percent of the cases for the diversification-risk ratio. The naive portfolio is nearly always overruled by the other strategies, only in terms of the volatility, it performs better than the mean-variance portfolio in 30 percent of the cases.

In the 180 months estimation window case, the mean-variance portfolio is still superior to the naive portfolio for all performance criteria except the volatility (only in nearly 80 percent of the cases) and the diversification-risk ratio (in over 90 percent of the cases). It is always better than the cluster portfolio for the mean, the turnover and the CEQ return, and nearly always better

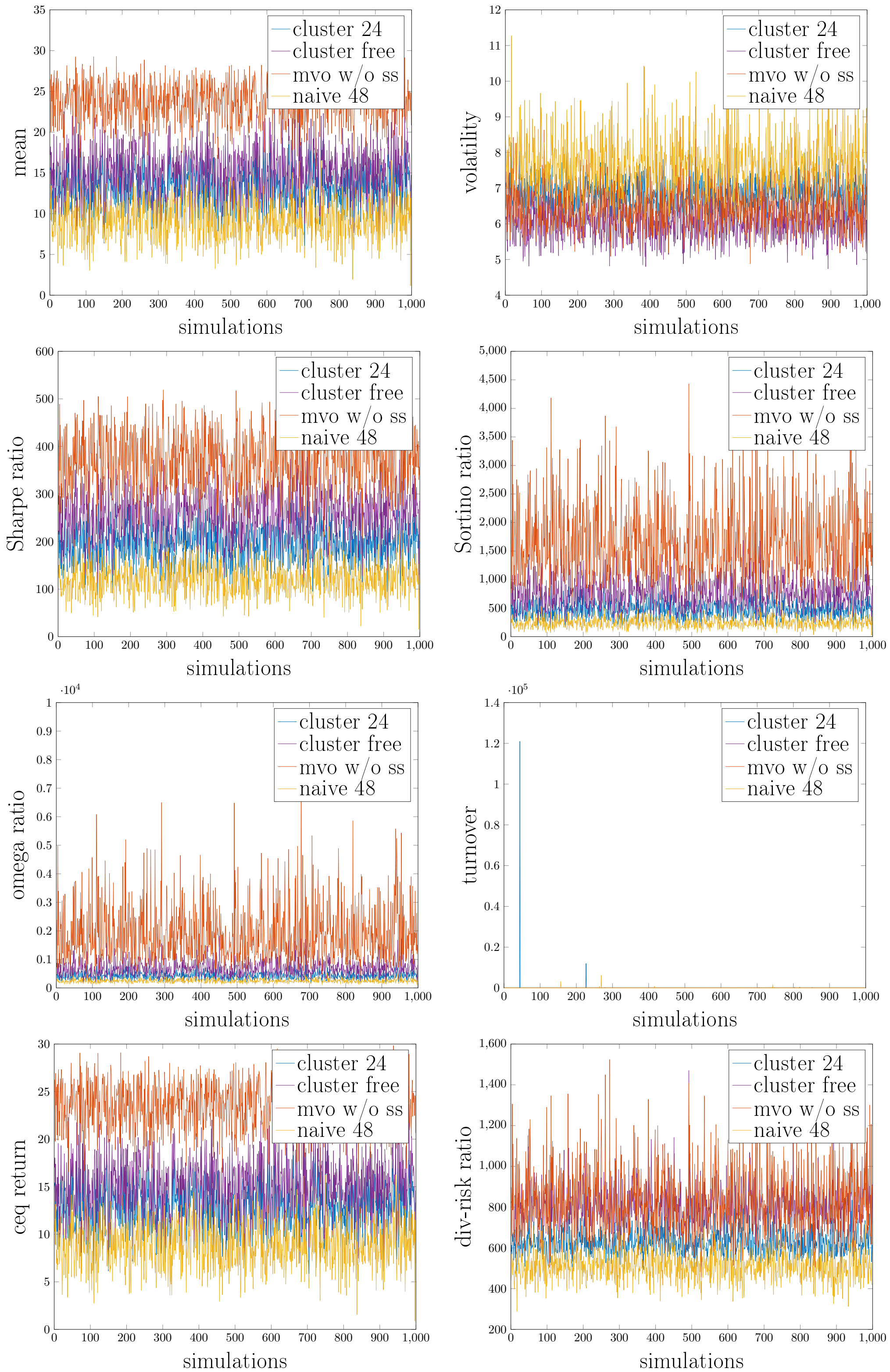


Figure 5.13: Model 7.20: Annual performance (in percent, except turnover). Simulation model 7, 48 assets, 20 correlation groups, 180 months estimation window, 360 months total testing period. 1000 simulation sets. (For more details, the corresponding graphs with only the first 100 simulations can be found in Appendix D)

Estimation window: 60 months, testing window: 12 months, total testing period: 480 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster	2 (0)	1 (988)	2 (5)	2 (4)	2 (4)	2 (0)	2 (0)	1 (760)	2
naive: 48 assets	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3
mvo w/o ss	1 (1000)	2 (12)	1 (995)	1 (996)	1 (996)	1 (1000)	1 (1000)	2 (240)	1

Estimation window: 180 months, testing window: 12 months, total testing period: 360 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster	2 (0)	1 (949)	2 (1)	2 (2)	2 (2)	2 (0)	2 (0)	1 (647)	2
naive: 48 assets	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3
mvo w/o ss	1 (1000)	2 (51)	1 (999)	1 (998)	1 (998)	1 (1000)	1 (1000)	2 (353)	1

Estimation window: 360 months, testing window: 12 months, total testing period: 180 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster	2 (0)	1 (906)	2 (2)	2 (2)	2 (1)	2 (0)	2 (0)	2 (440)	2
naive: 48 assets	3 (0)	3 (3)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	3 (2)	3
mvo w/o ss	1 (1000)	2 (91)	1 (998)	1 (998)	1 (999)	1 (1000)	1 (1000)	1 (558)	1

Table 5.17: Model 7.24: Annual ranking. Comparison of the cluster portfolio, the full naive portfolio and the constrained mean-variance portfolio.
Settings: Simulation model 7.24 with 48 correlated assets with switching parameters and 24 comonotonic groups. Estimation window: 60, 180 and 360 months, testing window: 12 months, total testing period: 480, 360 and 180 months.
The number in brackets after the ranking indicates how often the portfolio outperformed both other strategies in terms of this performance criterion.

for the Sharpe, Sortino and omega ratio. Regarding volatility, it usually has higher values than the cluster portfolio. Moreover, the cluster portfolio outperforms the mean-variance portfolio in over 60 percent of the cases for the diversification-risk ratio. In addition, the cluster portfolio always shows better values as the naive portfolio in terms of the Sharpe and omega ratio as well as the diversification-risk ratio. It nearly always shows a greater mean, Sortino ratio and CEQ return, and in addition a lower turnover in over 90 percent of the cases.

For 360 months estimation periods, the mean-variance portfolio is still superior to the naive portfolio for all performance criteria except the volatility (only in over 80 percent of the cases) and the diversification-risk ratio (in over 90 percent of the cases). It is always better than the cluster portfolio for the mean, the turnover and the CEQ return. Moreover, it is nearly always better for the Sharpe, Sortino and omega ratio. Regarding volatility, it has higher values than the cluster portfolio in over 90 percent of the cases. In addition, the cluster portfolio outperforms the mean-variance portfolio in over 40 percent of the cases for the diversification-risk ratio. The cluster portfolio nearly always shows better values as the naive portfolio in terms of the volatility, the Sharpe, Sortino and omega ratio as well as the turnover and the diversification-risk ratio. Moreover, it performs better in over 90 percent of the cases for the mean and the CEQ return.

Simulation Model 7 with 28 groups

For 28 groups see Table 5.18. Corresponding to the 60 months estimation period case, we see that the mean-variance method outperforms all other methods in terms of the mean and the CEQ return, and nearly always regarding the Sharpe, Sortino and omega ratio as well as the turnover. The mean of the fixed cluster method is always greater than the mean of the free cluster method, but the volatility is also greater. In total the CEQ return of the fixed cluster

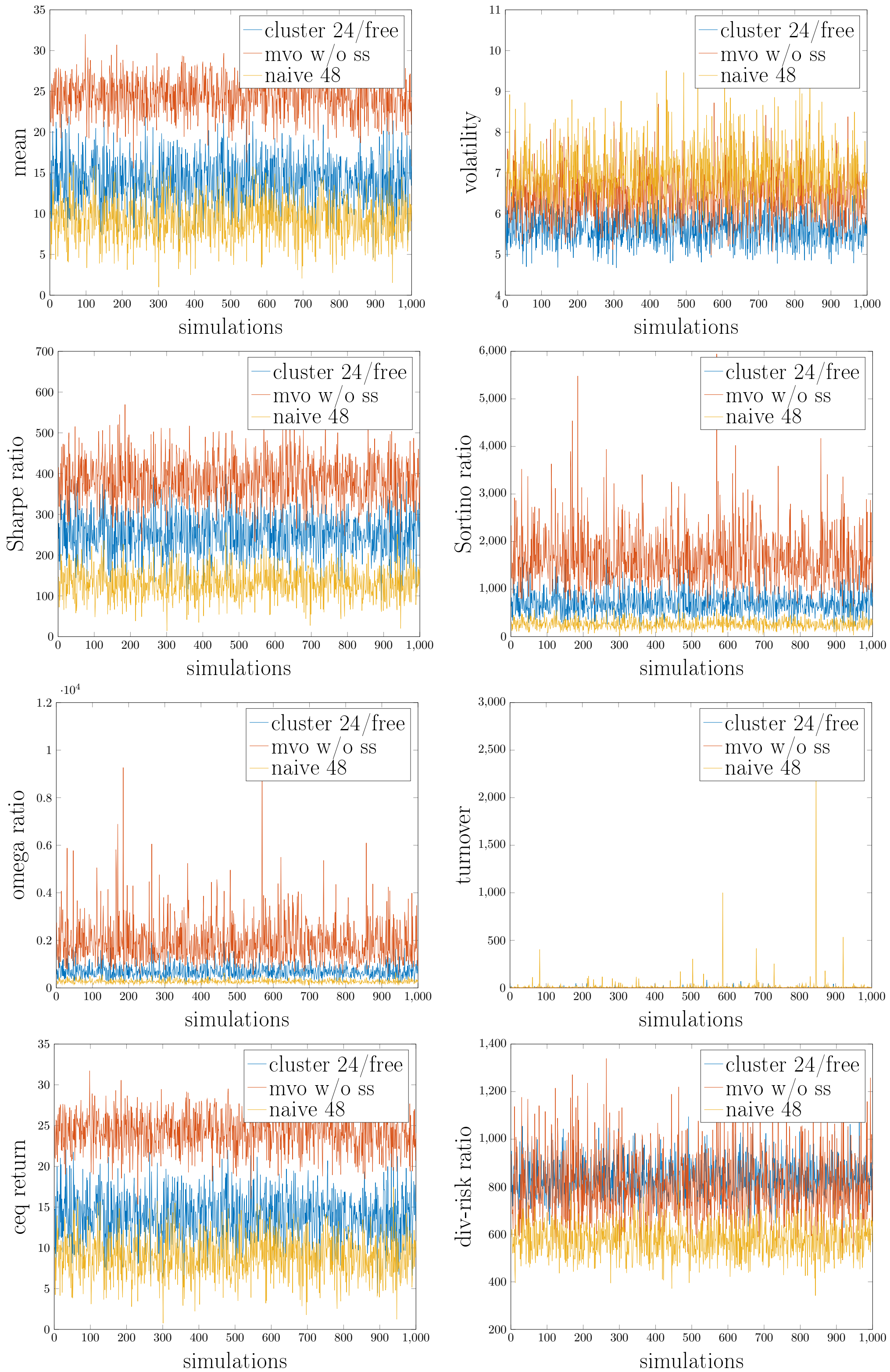


Figure 5.14: Model 7.24: Annual performance (in percent, except turnover). Simulation model 7, 48 assets, 24 correlation groups, 180 months estimation window, 360 months total testing period. 1000 simulation sets. (For more details, the corresponding graphs with only the first 100 simulations can be found in Appendix D)

Estimation window: 60 months, testing window: 12 months, total testing period: 480 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	2 (0)	2 (0)	2 (2)	2 (1)	3 (0)	2 (10)	2 (0)	2 (45)	2
Pearson cluster: free	3 (0)	1 (992)	3 (15)	3 (22)	2 (26)	3 (5)	3 (0)	1 (880)	3
naive: 28 assets	4 (0)	3 (3)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4
mvo w/o ss	1 (1000)	4 (5)	1 (983)	1 (977)	1 (974)	1 (985)	1 (1000)	3 (75)	1

Estimation window: 180 months, testing window: 12 months, total testing period: 360 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	2 (0)	2 (3)	2 (0)	2 (0)	2 (0)	2 (0)	2 (0)	2 (272)	2
Pearson cluster: free	3 (0)	1 (963)	3 (0)	3 (0)	3 (0)	3 (1)	3 (0)	1 (512)	3
naive: 28 assets	4 (0)	4 (5)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4
mvo w/o ss	1 (1000)	3 (29)	1 (1000)	1 (1000)	1 (1000)	1 (999)	1 (1000)	3 (216)	1

Estimation window: 360 months, testing window: 12 months, total testing period: 180 months. 1000 simulation sets.

ranking (1000)	mean	std	Sharpe ratio	Sortino ratio	omega ratio	turnover	ceq	div-risk ratio	overall
Pearson cluster: 24	2 (0)	2 (1)	2 (0)	3 (0)	3 (0)	2 (1)	2 (0)	2 (42)	2
Pearson cluster: free	3 (0)	1 (958)	3 (1)	2 (2)	2 (1)	3 (2)	3 (0)	1 (752)	3
naive: 28 assets	4 (0)	4 (4)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4
mvo w/o ss	1 (1000)	3 (37)	1 (999)	1 (998)	1 (999)	1 (997)	1 (1000)	3 (206)	1

Table 5.18: Model 7.28: Annual ranking. Comparison of the cluster portfolio with 24 assets, the free cluster portfolio, the full naive portfolio and the constrained mean-variance portfolio.
Settings: Simulation model 7.28 with 48 correlated assets with switching parameters and 28 comonotonic groups. Estimation window: 60, 180 and 360 months, testing window: 12 months, total testing period: 480, 360 and 180 months.
The number in brackets after the ranking indicates how often the portfolio outperformed both other strategies in terms of this performance criterion.

return is always greater than the CEQ return of the free version. However, the free cluster portfolio always has a higher diversification risk ratio as the fixed cluster method. The mean-variance portfolio always outperforms the naive method in terms of the Sharpe, Sortino and omega ratio. The standard deviation of the mean-variance portfolio is nearly always greater than the values from the cluster portfolios, and still in over 60 percent of the cases higher than the values from the naive portfolio. Also the naive portfolio usually leads to higher volatility than the volatility of the cluster portfolios.

Now, the mean-variance strategy is always superior to all other strategies, except for the standard deviation and the diversification-risk ratio in the 180 months estimation period case. For the volatility it performs worse than the fixed cluster method in nearly 90 percent of the cases, worse than the free cluster version in more than 90 percent of the cases and only worse than the naive portfolio in under 40 percent of the cases. In terms of the diversification-risk ratio it is outperformed by the cluster portfolios in over 70 percent of the cases, but it performs better than the naive portfolio in more than 90 percent of the cases. The naive portfolio is always or nearly always outperformed by both cluster methods. Moreover, the fixed cluster portfolio nearly always outperforms the free cluster portfolio. Only for the volatility (nearly always) and the diversification-risk ratio (more than 60 percent of the cases) the free cluster performs better.

In the case of 360 estimation periods, we see the following: The mean-variance method performs better than all other strategies for the mean and the CEQ return and nearly always for the turnover. For the Sharpe, Sortino and omega ratio it always performs better than the naive and the fixed cluster portfolio, and nearly always better than the free cluster portfolio. For the volatility the naive portfolio performs worse than the fixed cluster method in nearly

80 percent of the cases, worse than the free cluster version in more than 90 percent of the cases and only worse than the naive portfolio in under 40 percent of the cases. In terms of the diversification-risk ratio it is outperformed by the fixed cluster portfolis in over 50 percent of the cases, by the free version in nearly 80 percent of the cases. However, it performs better than the naive portfolio in nearly 90 percent of the cases. The naive portfolio is always or nearly always outperformed by both cluster methods. The fixed cluster portfolio outperforms the free version in over 90 percent of the cases for the mean and the CEQ return. However, the free version nearly always shows lower volatility than the fixed cluster portfolio. For the Sharpe ratio both cluster portfolios perform equally. For the Sortino and the omega ratio the performance of the free version is slightly better, whereas the turnover is slightly better for the fixed cluster method. However, regarding the diversification-risk ratio, the free version is nearly always superior to the fixed cluster method. An example for model 7.28 performance values over all simulations we see in Figure 5.15 with all performance criteria for the 180 months estimation period.

5.4.3 Concluding Remark

Although we did not simulate the comonotonic model in the multivariate normal case (Model 5) with restrictions to the variance, the results show that the cluster portfolios are in total always superior to the naive portfolio. Especially in terms of the Sharpe ratio and the diversification-risk ratio they are superior, which corresponds to the theorems in Section 5.3.3.

The same is true for Model 7, where we include parameter switching. Even, as we were not able to explicitly prove the corresponding results from the multivariate normal setting, regarding the diversification-risk ratio, we observe the same results empirically. In addition, we see in some cases that the cluster methods outperform the mean-variance portfolio regarding the diversification-risk ratio. However, the results from Section 5.3.4 can be observed in the data examples.

5.5 Summary

In this chapter we have seen that including the knowledge about the correlation structure improves the performance of the naive portfolio with fewer assets considerably.

We started with data experiments on real data comparing the cluster portfolios with free and fixed number of clusters with the full naive and the constrained mean-variance portfolio. There, we observed that at least one cluster portfolio performs better than the mean-variance portfolio. Moreover, both cluster portfolios perform better than the naive portfolio with 24 assets from Chapter 4. However, the performance seemed to highly depend on the data period.

Therefore, we examined the real data set regarding correlation structures. We saw that there is high correlation between the assets although they are branche portfolios. We also observed that groups with high internal correlation change over time and depend on the estimation period.

Thus, we used simulation models, where the correlation structure stayed stable over time. To examine the effects of choosing the number of assets according to the observed groups or under- or over-estimate the groups with the fixed cluster portfolio we used three different grouping structures for each simulation model, namely 20, 24 or 28 groups with high internal correlation.

In Model 4 we had static parameters, and the mean-variance portfolio always performed better than the other strategies, whereas the full naive portfolio always took last place. The cluster portfolios shared the ranking places in between. The fixed version performed better when

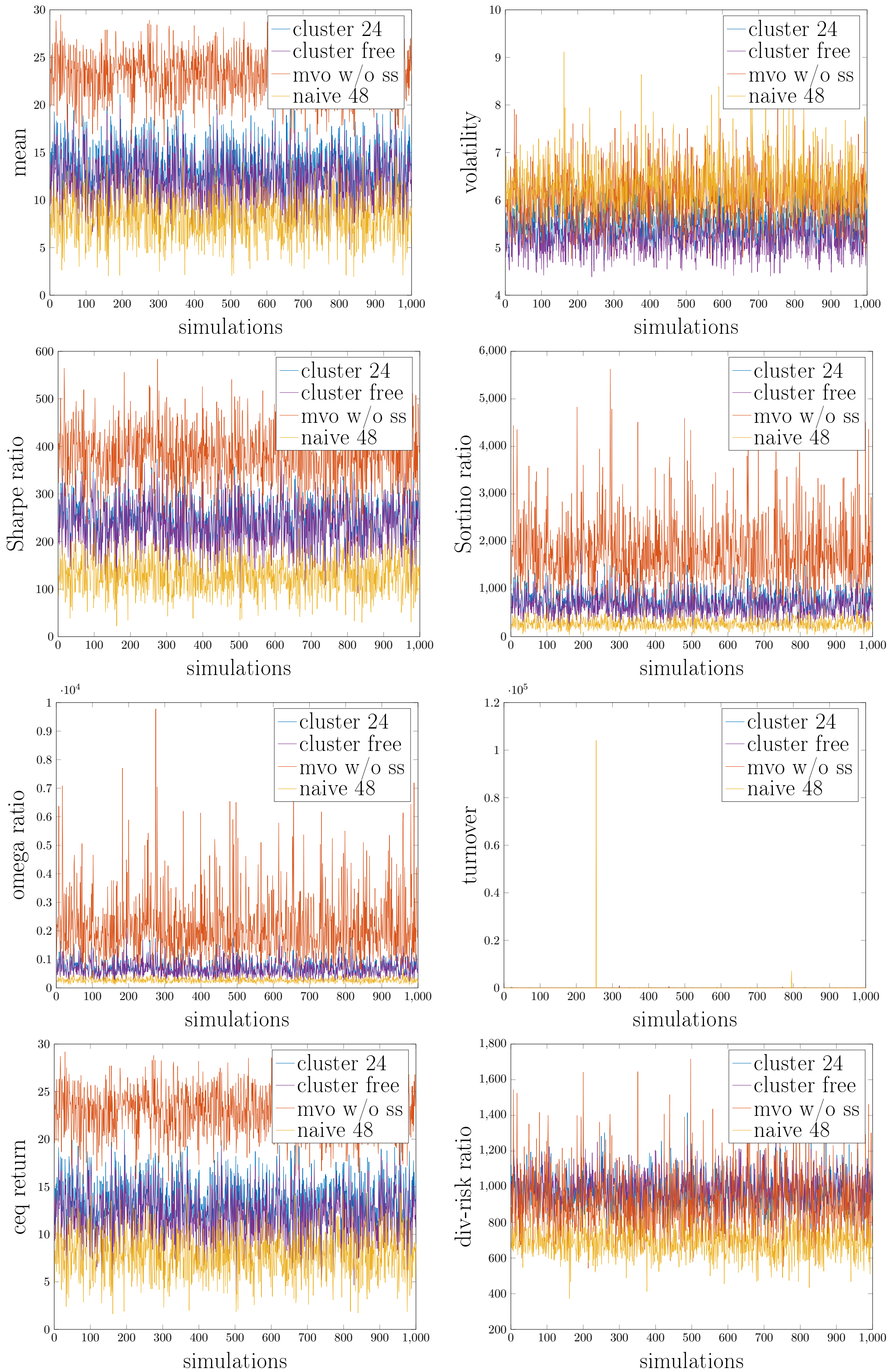


Figure 5.15: Model 7.28: Annual performance (in percent, except turnover). Simulation model 7, 48 assets, 28 correlation groups, 180 months estimation window, 360 months total testing period. 1000 simulation sets. (For more details, the corresponding graphs with only the first 100 simulations can be found in Appendix D)

the true number of groups was actually higher and the free version performed better when it was lower.

In Model 6 we considered model uncertainty by using switching parameters, but the effect on the mean-variance portfolio could only be observed for 60 months estimation period. There the free cluster portfolio could take the first place. The mean-variance portfolio still performed well since the strategy also profits from the covariance structure. In general all strategies led to similar performance, but the cluster portfolios more or less stayed between the very good performance of the mean-variance portfolio and the bad performance of the full naive portfolio.

In order to explain the observed effects of correlation structures, we introduced the concept of comonotonicity. As an extreme case we have no correlation between the clusters and perfect positive correlation between the assets within a cluster. We were able to show that in case of equal variances, we only need to achieve a higher mean-return when selecting assets in order to achieve a higher Sharpe ratio and diversification-risk ratio for a cluster portfolio.

With the Simulation Models 5 and 7 we included the extreme correlation structure, with no correlation between the groups and perfect positive correlation within the groups, in the Models 4 and 6. Although we were not including the equal variance condition we could see that the cluster portfolios are always superior to the full naive portfolio, especially with respect to the Sharpe ratio and the diversification-risk ratio. This was even true for Model 7, where we were not able to show the results for the diversification-risk ratio.

6 Conclusion and Further Research

In this thesis we presented a systematic examination of the naive portfolio with fewer assets. For this, we critically explored the performance of the portfolios from different points of view including empirical and theoretical approaches.

In terms of constant rebalanced portfolios (Cover (1991)), we could prove that the naive portfolio with fewer assets has the same growth rate as the full naive portfolio. Then, we reduced the number of assets in a naive portfolio to half the number of available assets, selecting the ones with the best Sharpe ratio. We also showed statistical advantages regarding the estimation error following the lines of DeMiguel, Garlappi, and Uppal (2009). Alongside we used real data and three simulation models with independent assets, and static and switching parameters to compare both naive portfolios and the unconstrained mean-variance portfolio via several performance criteria.

For this, we suggested a new performance measure, namely the diversification-risk ratio, which uses a diversification measure based on correlation and the risk part of the omega ratio. However, it is restricted to portfolios with no short positions, thus it was only used for the naive portfolios in this part. There we found that the naive portfolio with half the number of assets is always ranked between the mean-variance portfolio and the full naive portfolio. While for simulations with static parameters the mean-variance portfolio is on the first place, for the switching parameters the full naive method is on the first place. Thus, the portfolio with half the number of assets is robust to this kind of model uncertainty. Still, the full naive portfolio outperformed the naive method with fewer assets in terms of the diversification-risk ratio. Often, the results of the unconstrained mean-variance portfolio could not be compared to the other strategies due to unreasonable portfolio returns. Therefore, we examined three mean-variance portfolios and chose the mean-variance portfolio without shortselling as a more challenging benchmark for the remaining simulation studies.

Although the real data set consists of branche portfolios, we still could find correlation between the assets. Thus, we examined the portfolio with fewer assets in the presence of correlation structure. For this, we extended the selection of assets to finding the correlation structure in the market, applying a hierarchical cluster algorithm. We either fixed the number of clusters to one half or let the cluster algorithm specify the number of clusters. We showed data experiments with real data and simulated data including correlation structures with and without parameter switching so that the cluster portfolios perform better than the naive portfolio.

In this setting we used the framework of comonotonicity to show when the cluster portfolio outperforms the full naive portfolio, regarding the Sharpe ratio and the diversification-risk ratio.

We examined the multivariate normal setting without parameter switching. There, we found that the Sharpe ratio of the cluster portfolio is greater than the Sharpe ratio of the naive portfolio in an equal variance case, if the mean of the cluster portfolio is higher than the mean of the naive portfolio. This told us that in case of an enhancement of the mean via selection, the cluster portfolio is always preferable in terms of the Sharpe ratio. In the case of equal-sized groups, the diversification-risk ratio of the cluster portfolio is always greater than the diversification-risk ratio of the naive portfolio in this setting if again the mean of the naive portfolio is smaller than the mean of the cluster portfolio. For different sized groups this depends on the chosen threshold for the Ω-risk. The diversification-risk ratio of the cluster portfolio is greater than the corresponding ratio of the naive portfolio, if the threshold is below the value for which the Ω-risks are equal.

In the mixed normal setting including parameter switching we could show even less restrictive results for the Sharpe ratio. In an equal variance setting, where the mean is driven by a discrete, irreducible, homogeneous Markov chain which admits a stationary distribution, the Sharpe ratio of the cluster portfolio is always greater than the Sharpe ratio of the naive portfolio. To conclude this study, we provided empirical results on real data and comonotonic simulation models with and without parameter switching. They confirm that the theoretical results allow to explain the observed results in real and simulated settings, although the simulated data did not completely follow the settings of the proofs.

In the data experiments, we have examined the portfolios with real and simulated data sets. In the following, we review our observations.

For the real data set we have seen that the selection of assets in a portfolio with half the number of assets plays an important role. The full naive portfolio nearly always outperformed the naive portfolio with 24 assets chosen by the historical best Sharpe ratio. When clustering was used for the selection of assets, then the cluster portfolios nearly always outperform the full naive portfolio. However, it became obvious in the data experiments with simulated data that there is no best cluster portfolio, since it depends on the data.

The sample mean-variance strategy was nearly always outperformed by the naive portfolios, which was usually due to extreme positions. The constrained mean-variance portfolio overcame those drawbacks and was able to outperform the other strategies in some cases. If it did not take the first place in ranking it was outperformed by at least one cluster portfolio. The increase of the estimation period did not consistently improve the performance of the mean-variance portfolios for real data.

However, there was one major difference between the simulation models. The first three models had independent assets. All other models had correlation structures with high positive correlation between some of the assets.

We have seen that, if the models are static, the mean-variance portfolios could not be outperformed by naive methods. This was the case for the first simulation model. The third simulation model led to similar results as the parameters changed only within a small range. In the second model the parameters could change totally and thus the full naive portfolio outperformed all other portfolios. There, the unconstrained mean-variance portfolio was always ranked on the last place. Only the constrained mean-variance portfolio was able to improve its performance with rising estimation period. In all these models the naive portfolio with half the number of assets was ranked on the second place, indifferent to which method was on the first place. This confirmed that the naive portfolio could be made robust via selection.

Imposing correlation structures between the assets showed, that the number of groups in the data plays a role for the cluster portfolio. For data sets with static parameters and fewer groups than half the number of assets the free cluster portfolio performed slightly better than the fixed version. However, it was the other way around, if the number of clusters was higher than half the number of assets. Also for switching parameters, the free version nearly always outperformed the fixed cluster portfolio. Nevertheless, the performance of the cluster portfolios usually lay between the performance of the full naive and the mean-variance portfolio.

The comonotonic simulations confirmed what we had seen in the simulation models with correlation structure. The full naive portfolio was always outperformed by the other strategies. The mean-variance portfolio always took first place and the cluster portfolios were ranked in between. In addition, the models supported that the free cluster method outperforms the fixed version for data, which possesses fewer groups than half the number of assets and the other way around for more groups than half the number of assets. These simulations supported the theoretical results we proved in the comonotonic setting.

During the work on this thesis, we also came across other notions of dependence. We took a look at the cointegration approach, but unfortunately the model uncertainty made it impossible to use this approach for financial assets. Then, we worked with data mining techniques to find similarities between the return series. We think there might be a chance to create a better definition of dependence improving portfolio performance. However, since we only applied some low standard techniques the performance of such a cluster portfolio was always dominated by the correlation based clustering procedure.

Therefore, we recommend more research on the topic of using data mining procedures to decrease problems with estimation errors and to find robust dependence structures, regarding future performance. As Cont, Deguest, and Scandolo (2010) pointed out, the statistical robustness of risk measures should be included in evaluating risk measures and there is still a lot of research to be done. In particular, we recommend the usage of a diversification-risk ratio. Unfortunately, since our version suffers from drawbacks like that it is only applicable to portfolios with no shortselling. The general idea of a diversification-risk ratio should be further examined using different criteria, nevertheless.

A Real Data

The real data set is provided by Kenneth French's website
http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html,
where apart from others, historical data for 48 industry portfolios on the US market are given. The data set we use is from July 1^{st}, 1969 up to July 31^{st}, 2015. When analyzing the portfolios we use the data from July 1^{st}, 1969 up to June 30^{th}, 2014, to have exactly 45 years compatible to the rolling system with 5 years periods, we use to collect the performance of the portfolios.

We use this data set for several reasons. The data sets provided by Kenneth French are often used in literature. There is open access to this data set and therefore results can be compared to results from different papers.

A.1 Industry Portfolios

The combination of the industry portfolios are described by Kenneth French as follows:

> We assign each NYSE, AMEX, and NASDAQ stock to an industry portfolio at the end of June of year t based on its four-digit SIC code at that time. (We use Compustat SIC codes for the fiscal year ending in calendar year t-1. Whenever Compustat SIC codes are not available, we use CRSP SIC codes for June of year t.) We then compute returns from July of t to June of t+1.

The participants of the portfolios are reported by Kenneth French this way:

1. Agriculture
   0100-0199 Agric production - crops
   0200-0299 Agric production - livestock
   0700-0799 Agricultural services
   0910-0919 Commercial fishing
   2048-2048 Prepared feeds for animals

2. Food Products
   2000-2009 Food and kindred products
   2010-2019 Meat products
   2020-2029 Dairy products
   2030-2039 Canned-preserved fruits-vegs
   2040-2046 Flour and other grain mill products
   2050-2059 Bakery products
   2060-2063 Sugar and confectionery products
   2070-2079 Fats and oils
   2090-2092 Misc food preps
   2095-2095 Roasted coffee

2098-2099 Misc food preparations

3. Candy & Soda
   2064-2068 Candy and other confectionery
   2086-2086 Bottled-canned soft drinks
   2087-2087 Flavoring syrup
   2096-2096 Potato chips
   2097-2097 Manufactured ice

4. Beer & Liquor
   2080-2080 Beverages
   2082-2082 Malt beverages
   2083-2083 Malt
   2084-2084 Wine
   2085-2085 Distilled and blended liquors

5. Tobacco Products
   2100-2199 Tobacco products

6. Recreation
   0920-0999 Fishing, hunting & trapping
   3650-3651 Household audio visual equip
   3652-3652 Phonographic records
   3732-3732 Boat building and repair
   3930-3931 Musical instruments
   3940-3949 Toys

7. Entertainment
   7800-7829 Services - motion picture production and distribution
   7830-7833 Services - motion picture theatres
   7840-7841 Services - video rental
   7900-7900 Services - amusement and recreation
   7910-7911 Services - dance studios
   7920-7929 Services - bands, entertainers
   7930-7933 Services - bowling centers
   7940-7949 Services - professional sports
   7980-7980 Amusement and recreation services (?)
   7990-7999 Services - misc entertainment

8. Printing and Publishing
   2700-2709 Printing publishing and allied
   2710-2719 Newspapers: publishing-printing
   2720-2729 Periodicals: publishing-printing

2730-2739 Books: publishing-printing
2740-2749 Misc publishing
2770-2771 Greeting card publishing
2780-2789 Book binding
2790-2799 Service industries for print trade

9. Consumer Goods
2047-2047 Dog and cat food
2391-2392 Curtains, home furnishings
2510-2519 Household furniture
2590-2599 Misc furniture and fixtures
2840-2843 Soap & other detergents
2844-2844 Perfumes cosmetics
3160-3161 Luggage
3170-3171 Handbags and purses
3172-3172 Personal leather goods, except handbags
3190-3199 Leather goods
3229-3229 Pressed and blown glass
3260-3260 Pottery and related products
3262-3263 China and earthenware table articles
3269-3269 Pottery products
3230-3231 Glass products
3630-3639 Household appliances
3750-3751 Motorcycles, bicycles and parts (Harley & Huffy)
3800-3800 Misc inst, photo goods, watches
3860-3861 Photographic equip (Kodak etc, but also Xerox)
3870-3873 Watches clocks and parts
3910-3911 Jewelry-precious metals
3914-3914 Silverware
3915-3915 Jewelers' findings, materials
3960-3962 Costume jewelry and notions
3991-3991 Brooms and brushes
3995-3995 Burial caskets

10. Apparel
2300-2390 Apparel and other finished products
3020-3021 Rubber and plastics footwear
3100-3111 Leather tanning and finishing
3130-3131 Boot, shoe cut stock, findings
3140-3149 Footware except rubber
3150-3151 Leather gloves and mittens
3963-3965 Fasteners, buttons, needles, pins

11. Healthcare
8000-8099 Services - health

12. Medical Equipment
    3693-3693 X-ray, electromedical app
    3840-3849 Surg & med instru
    3850-3851 Ophthalmic goods

13. Pharmaceutical Products
    2830-2830 Drugs
    2831-2831 Biological products
    2833-2833 Medicinal chemicals
    2834-2834 Pharmaceutical preparations
    2835-2835 In vitro, in vivo diagnostics
    2836-2836 Biological products, except diagnostics

14. Chemicals
    2800-2809 Chemicals and allied products
    2810-2819 Industrial inorganic chems
    2820-2829 Plastic material & synthetic resin
    2850-2859 Paints
    2860-2869 Industrial organic chems
    2870-2879 Agriculture chemicals
    2890-2899 Misc chemical products

15. Rubber and Plastic Products
    3031-3031 Reclaimed rubber
    3041-3041 Rubber & plastic hose and belting
    3050-3053 Gaskets, hoses, etc
    3060-3069 Fabricated rubber products
    3070-3079 Misc rubber products (?)
    3080-3089 Misc plastic products
    3090-3099 Misc rubber and plastic products (?)

16. Textiles
    2200-2269 Textile mill products
    2270-2279 Floor covering mills
    2280-2284 Yarn and thread mills
    2290-2295 Misc textile goods
    2297-2297 Nonwoven fabrics
    2298-2298 Cordage and twine
    2299-2299 Misc textile products
    2393-2395 Textile bags, canvas products
    2397-2399 Misc textile products

17. Construction Materials
    0800-0899 Forestry
    2400-2439 Lumber and wood products
    2450-2459 Wood buildings-mobile homes
    2490-2499 Misc wood products
    2660-2661 Building paper and board mills
    2950-2952 Paving & roofing materials
    3200-3200 Stone, clay, glass, concrete etc
    3210-3211 Flat glass
    3240-3241 Cement hydraulic
    3250-3259 Structural clay prods
    3261-3261 Vitreous china plumbing fixtures
    3264-3264 Porcelain electrical supply
    3270-3275 Concrete gypsum & plaster
    3280-3281 Cut stone and stone products
    3290-3293 Abrasive and asbestos products
    3295-3299 Non-metallic mineral products
    3420-3429 Handtools and hardware
    3430-3433 Heating equip & plumbing fix
    3440-3441 Fabricated struct metal products
    3442-3442 Metal doors, frames
    3446-3446 Architectural or ornamental metal work
    3448-3448 Pre-fab metal buildings
    3449-3449 Misc structural metal work
    3450-3451 Screw machine products
    3452-3452 Bolts, nuts screws
    3490-3499 Misc fabricated metal products
    3996-3996 Hard surface floor cover

18. Construction
    1500-1511 Build construction - general contractors
    1520-1529 Gen building contractors - residential
    1530-1539 Operative builders
    1540-1549 Gen building contractors - non-residential
    1600-1699 Heavy Construction - not building contractors
    1700-1799 Construction - special contractors

19. Steel Works Etc
    3300-3300 Primary metal industries
    3310-3317 Blast furnaces & steel works
    3320-3325 Iron & steel foundries
    3330-3339 Prim smelt-refin nonfer metals
    3340-3341 Secondary smelt-refin nonfer metals
    3350-3357 Rolling & drawing nonferrous metals
    3360-3369 Non-ferrous foundries and casting

3370-3379 Steel works etc
3390-3399 Misc primary metal products

20. Fabricated Products
3400-3400 Fabricated metal, except machinery and trans eq
3443-3443 Fabricated plate work
3444-3444 Sheet metal work
3460-3469 Metal forgings and stampings
3470-3479 Coating and engraving

21. Machinery
3510-3519 Engines & turbines
3520-3529 Farm and garden machinery
3530-3530 Constr, mining material handling machinery
3531-3531 Construction machinery
3532-3532 Mining machinery, except oil field
3533-3533 Oil field machinery
3534-3534 Elevators
3535-3535 Conveyors
3536-3536 Cranes, hoists
3538-3538 Machinery
3540-3549 Metalworking machinery
3550-3559 Special industry machinery
3560-3569 General industrial machinery
3580-3580 Refrig & service ind machines
3581-3581 Automatic vending machines
3582-3582 Commercial laundry and drycleaning machines
3585-3585 Air conditioning, heating, refrid eq
3586-3586 Measuring and dispensing pumps
3589-3589 Service industry machinery
3590-3599 Misc industrial and commercial equipment and mach

22. Electrical Equipment
3600-3600 Elec mach eq & supply
3610-3613 Elec transmission
3620-3621 Electrical industrial appar
3623-3629 Electrical industrial appar
3640-3644 Electric lighting, wiring
3645-3645 Residential lighting fixtures
3646-3646 Commercial lighting
3648-3649 Lighting equipment
3660-3660 Communication equip
3690-3690 Miscellaneous electrical machinery and equip
3691-3692 Storage batteries

3699-3699 Electrical machinery and equip

23. Automobiles and Trucks
    2296-2296 Tire cord and fabric
    2396-2396 Auto trim
    3010-3011 Tires and inner tubes
    3537-3537 Trucks, tractors, trailers
    3647-3647 Vehicular lighting
    3694-3694 Elec eq, internal combustion engines
    3700-3700 Transportation equipment
    3710-3710 Motor vehicles and motor vehicle equip
    3711-3711 Motor vehicles & car bodies
    3713-3713 Truck & bus bodies
    3714-3714 Motor vehicle parts
    3715-3715 Truck trailers
    3716-3716 Motor homes
    3792-3792 Travel trailers and campers
    3790-3791 Misc trans equip
    3799-3799 Misc trans equip

24. Aircraft
    3720-3720 Aircraft & parts
    3721-3721 Aircraft
    3723-3724 Aircraft engines, engine parts
    3725-3725 Aircraft parts
    3728-3729 Aircraft parts

25. Shipbuilding, Railroad Equipment
    3730-3731 Ship building and repair
    3740-3743 Railroad Equipment

26. Defense
    3760-3769 Guided missiles and space vehicles
    3795-3795 Tanks and tank components
    3480-3489 Ordnance & accessories

27. Precious Metals
    1040-1049 Gold & silver ores

28. Non-Metallic and Industrial Metal Mining
    1000-1009 Metal mining
    1010-1019 Iron ores
    1020-1029 Copper ores

1030-1039 Lead and zinc ores
1050-1059 Bauxite and other aluminum ores
1060-1069 Ferroalloy ores
1070-1079 Mining
1080-1089 Mining services
1090-1099 Misc metal ores
1100-1119 Anthracite mining
1400-1499 Mining and quarrying non-metallic minerals

29. Coal
    1200-1299 Bituminous coal

30. Petroleum and Natural Gas
    1300-1300 Oil and gas extraction
    1310-1319 Crude petroleum & natural gas
    1320-1329 Natural gas liquids
    1330-1339 Petroleum and natural gas
    1370-1379 Petroleum and natural gas
    1380-1380 Oil and gas field services
    1381-1381 Drilling oil & gas wells
    1382-1382 Oil-gas field exploration
    1389-1389 Oil and gas field services
    2900-2912 Petroleum refining
    2990-2999 Misc petroleum products

31. Utilities
    4900-4900 Electric, gas, sanitary services
    4910-4911 Electric services
    4920-4922 Natural gas transmission
    4923-4923 Natural gas transmission-distr
    4924-4925 Natural gas distribution
    4930-4931 Electric and other services combined
    4932-4932 Gas and other services combined
    4939-4939 Combination utilities
    4940-4942 Water supply

32. Communication
    4800-4800 Communications
    4810-4813 Telephone communications
    4820-4822 Telegraph and other message communication
    4830-4839 Radio-TV Broadcasters
    4840-4841 Cable and other pay TV services
    4880-4889 Communications
    4890-4890 Communication services (Comsat)

4891-4891 Cable TV operators
4892-4892 Telephone interconnect
4899-4899 Communication services

33. Personal Services
7020-7021 Rooming and boarding houses
7030-7033 Camps and recreational vehicle parks
7200-7200 Services - personal
7210-7212 Services - laundry, cleaners
7214-7214 Services - diaper service
7215-7216 Services - coin-op cleaners, dry cleaners
7217-7217 Services - carpet, upholstery cleaning
7219-7219 Services - laundry, cleaners
7220-7221 Services - photo studios, portrait
7230-7231 Services - beauty shops
7240-7241 Services - barber shops
7250-7251 Services - shoe repair
7260-7269 Services - funeral
7270-7290 Services - misc
7291-7291 Services - tax return
7292-7299 Services - misc
7395-7395 Services - photofinishing labs (School pictures)
7500-7500 Services - auto repair, services
7520-7529 Services - automobile parking
7530-7539 Services - auto repair shops
7540-7549 Services - auto services, except repair (car washes)
7600-7600 Services - Misc repair services
7620-7620 Services - Electrical repair shops
7622-7622 Services - Radio and TV repair shops
7623-7623 Services - Refridg and air conditioner repair
7629-7629 Services - Electrical repair shops
7630-7631 Services - Watch, clock and jewelry repair
7640-7641 Services - Reupholster, furniture repair
7690-7699 Services - Misc repair shops
8100-8199 Services - legal
8200-8299 Services - educational
8300-8399 Services - social services
8400-8499 Services - museums, galleries, botanic gardens
8600-8699 Services - membership organizations
8800-8899 Services - private households
7510-7515 Services - truck, auto rental and leasing

34. Business Services
2750-2759 Commercial printing
3993-3993 Signs, advertising specialty

7218-7218 Services - industrial launderers
7300-7300 Services - business services
7310-7319 Services - advertising
7320-7329 Services - credit reporting agencies, collection services
7330-7339 Services - mailing, reproduction, commercial art
7340-7342 Services - services to dwellings, other buildings
7349-7349 Services - cleaning and building maint
7350-7351 Services - misc equip rental and leasing
7352-7352 Services - medical equip rental
7353-7353 Services - heavy construction equip rental
7359-7359 Services - equip rental and leasing
7360-7369 Services - personnel supply services
7370-7372 Services - computer programming and data processing
7374-7374 Services - computer processing, data prep
7375-7375 Services - information retrieval services
7376-7376 Services - computer facilities management service
7377-7377 Services - computer rental and leasing
7378-7378 Services - computer maintenance and repair
7379-7379 Services - computer related services
7380-7380 Services - misc business services
7381-7382 Services - security
7383-7383 Services - news syndicates
7384-7384 Services - photofinishing labs
7385-7385 Services - telephone interconnections
7389-7390 Services - misc business services
7391-7391 Services - R&D labs
7392-7392 Services - management consulting & P.R.
7393-7393 Services - detective and protective (ADT)
7394-7394 Services - equipment rental & leasing
7396-7396 Services - trading stamp services
7397-7397 Services - commercial testing labs
7399-7399 Services - business services
7519-7519 Services - trailer rental and leasing
8700-8700 Services - engineering, accounting, research, management
8710-8713 Services - engineering, accounting, surveying
8720-8721 Services - accounting, auditing, bookkeeping
8730-8734 Services - research, development, testing labs
8740-8748 Services - management, public relations, consulting
8900-8910 Services - misc
8911-8911 Services - engineering & architect
8920-8999 Services - misc
4220-4229 Warehousing and storage

35. Computers
3570-3579 Office computers
3680-3680 Computers

3681-3681 Computers - mini
3682-3682 Computers - mainframe
3683-3683 Computers - terminals
3684-3684 Computers - disk & tape drives
3685-3685 Computers - optical scanners
3686-3686 Computers - graphics
3687-3687 Computers - office automation systems
3688-3688 Computers - peripherals
3689-3689 Computers - equipment
3695-3695 Magnetic and optical recording media
7373-7373 Computer integrated systems design

36. Electronic Equipment
    3622-3622 Industrial controls
    3661-3661 Telephone and telegraph apparatus
    3662-3662 Communications equipment
    3663-3663 Radio TV comm equip & apparatus
    3664-3664 Search, navigation, guidance systems
    3665-3665 Training equipment & simulators
    3666-3666 Alarm & signaling products
    3669-3669 Communication equipment
    3670-3679 Electronic components
    3810-3810 Search, detection, navigation, guidance
    3812-3812 Search, detection, navigation, guidance

37. Measuring and Control Equipment
    3811-3811 Engr lab and research equipment
    3820-3820 Measuring and controlling equipment
    3821-3821 Lab apparatus and furniture
    3822-3822 Automatic controls - Envir and applic
    3823-3823 Industrial measurement instru
    3824-3824 Totalizing fluid meters
    3825-3825 Elec meas & test instr
    3826-3826 Lab analytical instruments
    3827-3827 Optical instr and lenses
    3829-3829 Meas and control devices
    3830-3839 Optical instr and lenses

38. Business Supplies
    2520-2549 Office furniture and fixtures
    2600-2639 Paper and allied products
    2670-2699 Paper and allied products
    2760-2761 Manifold business forms
    3950-3955 Pens pencils and office supplies

39. Shipping Containers
    2440-2449 Wood containers
    2640-2659 Paperboard containers, boxes, drums, tubs
    3220-3221 Glass containers
    3410-3412 Metal cans and shipping containers

40. Transportation
    4000-4013 Railroads-line haul
    4040-4049 Railway express service
    4100-4100 Transit and passenger trans
    4110-4119 Local passenger trans
    4120-4121 Taxicabs
    4130-4131 Intercity bus trans (Greyhound)
    4140-4142 Bus charter
    4150-4151 School buses
    4170-4173 Motor vehicle terminals, service facilities
    4190-4199 Misc transit and passenger transportation
    4200-4200 Motor freight trans, warehousing
    4210-4219 Trucking
    4230-4231 Terminal facilities - motor freight
    4240-4249 Transportation
    4400-4499 Water transport
    4500-4599 Air transportation
    4600-4699 Pipelines, except natural gas
    4700-4700 Transportation services
    4710-4712 Freight forwarding
    4720-4729 Travel agencies, etc
    4730-4739 Arrange trans - freight and cargo
    4740-4749 Rental of railroad cars
    4780-4780 Misc services incidental to trans
    4782-4782 Inspection and weighing services
    4783-4783 Packing and crating
    4784-4784 Fixed facilities for vehicles, not elsewhere classified
    4785-4785 Motor vehicle inspection
    4789-4789 Transportation services

41. Wholesale
    5000-5000 Wholesale - durable goods
    5010-5015 Wholesale - autos and parts
    5020-5023 Wholesale - furniture and home furnishings
    5030-5039 Wholesale - lumber and construction materials
    5040-5042 Wholesale - professional and commercial equipment and supplies
    5043-5043 Wholesale - photographic equipment
    5044-5044 Wholesale - office equipment
    5045-5045 Wholesale - computers

5046-5046 Wholesale - commercial equip
5047-5047 Wholesale - medical, dental equip
5048-5048 Wholesale - ophthalmic goods
5049-5049 Wholesale - professional equip and supplies
5050-5059 Wholesale - metals and minerals
5060-5060 Wholesale - electrical goods
5063-5063 Wholesale - electrical apparatus and equipment
5064-5064 Wholesale - electrical appliance TV and radio
5065-5065 Wholesale - electronic parts
5070-5078 Wholesale - hardware, plumbing, heating equip
5080-5080 Wholesale - machinery and equipment
5081-5081 Wholesale - machinery and equipment (?)
5082-5082 Wholesale - construction and mining equipment
5083-5083 Wholesale - farm and garden machinery
5084-5084 Wholesale - industrial machinery and equipment
5085-5085 Wholesale - industrial supplies
5086-5087 Wholesale - machinery and equipment (?)
5088-5088 Wholesale - trans eq except motor vehicles
5090-5090 Wholesale - misc durable goods
5091-5092 Wholesale - sporting goods, toys
5093-5093 Wholesale - scrap and waste materials
5094-5094 Wholesale - jewelry and watches
5099-5099 Wholesale - durable goods
5100-5100 Wholesale - nondurable goods
5110-5113 Wholesale - paper and paper products
5120-5122 Wholesale - drugs & propietary
5130-5139 Wholesale - apparel
5140-5149 Wholesale - groceries & related prods
5150-5159 Wholesale - farm products
5160-5169 Wholesale - chemicals & allied prods
5170-5172 Wholesale - petroleum and petro prods
5180-5182 Wholesale - beer, wine
5190-5199 Wholesale - non-durable goods

42. Retail
5200-5200 Retail - bldg material, hardware, garden
5210-5219 Retail - lumber & other building mat
5220-5229 Retail
5230-5231 Retail - paint, glass, wallpaper
5250-5251 Retail - hardware stores
5260-5261 Retail - nurseries, lawn, garden stores
5270-5271 Retail - mobile home dealers
5300-5300 Retail - general merchandise stores
5310-5311 Retail - department stores
5320-5320 Retail - general merchandise stores (?)
5330-5331 Retail - variety stores

5334-5334 Retail - catalog showroom
5340-5349 Retail
5390-5399 Retail - Misc general merchandise stores
5400-5400 Retail - food stores
5410-5411 Retail - grocery stores
5412-5412 Retail - convenience stores
5420-5429 Retail - meat, fish mkt
5430-5439 Retail - fruit and vegetable markets
5440-5449 Retail - candy, nut, confectionary stores
5450-5459 Retail - dairy product stores
5460-5469 Retail - bakeries
5490-5499 Retail - miscellaneous food stores
5500-5500 Retail - auto dealers and gas stations
5510-5529 Retail - auto dealers
5530-5539 Retail - auto and home supply stores
5540-5549 Retail - gasoline service stations
5550-5559 Retail - boat dealers
5560-5569 Retail - recreational vehicle dealers
5570-5579 Retail - motorcycle dealers
5590-5599 Retail - automotive dealers
5600-5699 Retail - apparel & access
5700-5700 Retail - home furniture and equipment stores
5710-5719 Retail - home furnishings stores
5720-5722 Retail - household appliance stores
5730-5733 Retail - radio, TV and consumer electronic stores
5734-5734 Retail - computer and computer software stores
5735-5735 Retail - record and tape stores
5736-5736 Retail - musical instrument stores
5750-5799 Retail
5900-5900 Retail - misc
5910-5912 Retail - drug & proprietary stores
5920-5929 Retail - liquor stores
5930-5932 Retail - used merchandise stores
5940-5940 Retail - misc
5941-5941 Retail - sporting goods stores, bike shops
5942-5942 Retail - book stores
5943-5943 Retail - stationery stores
5944-5944 Retail - jewelry stores
5945-5945 Retail - hobby, toy and game shops
5946-5946 Retail - camera and photo shop
5947-5947 Retail - gift, novelty
5948-5948 Retail - luggage
5949-5949 Retail - sewing & needlework stores
5950-5959 Retail
5960-5969 Retail - non-store retailers (catalogs, etc)
5970-5979 Retail

5980-5989 Retail - fuel & ice stores (Penn Central Co)
5990-5990 Retail - retail stores
5992-5992 Retail - florists
5993-5993 Retail - tobacco stores
5994-5994 Retail - newsdealers
5995-5995 Retail - computer stores
5999-5999 Retail stores

43. Restaraunts, Hotels, Motels
    5800-5819 Retail - eating places
    5820-5829 Restaurants, hotels, motels
    5890-5899 Eating and drinking places
    7000-7000 Hotels, other lodging places
    7010-7019 Hotels motels
    7040-7049 Membership hotels and lodging
    7213-7213 Services - linen

44. Banking
    6000-6000 Depository institutions
    6010-6019 Federal reserve banks
    6020-6020 Commercial banks
    6021-6021 National commercial banks
    6022-6022 State banks - Fed Res System
    6023-6024 State banks - not Fed Res System
    6025-6025 National banks - Fed Res System
    6026-6026 National banks - not Fed Res System
    6027-6027 National banks, not FDIC
    6028-6029 Banks
    6030-6036 Savings institutions
    6040-6059 Banks (?)
    6060-6062 Credit unions
    6080-6082 Foreign banks
    6090-6099 Functions related to deposit banking
    6100-6100 Nondepository credit institutions
    6110-6111 Federal credit agencies
    6112-6113 FNMA
    6120-6129 S&Ls
    6130-6139 Agricultural credit institutions
    6140-6149 Personal credit institutions (Beneficial)
    6150-6159 Business credit institutions
    6160-6169 Mortgage bankers
    6170-6179 Finance lessors
    6190-6199 Financial services

45. Insurance
    6300-6300 Insurance
    6310-6319 Life insurance
    6320-6329 Accident and health insurance
    6330-6331 Fire, marine, property-casualty ins
    6350-6351 Surety insurance
    6360-6361 Title insurance
    6370-6379 Pension, health, welfare funds
    6390-6399 Insurance carriers
    6400-6411 Insurance agents

46. Real Estate
    6500-6500 Real estate
    6510-6510 Real estate operators
    6512-6512 Operators - non-resident buildings
    6513-6513 Operators - apartment buildings
    6514-6514 Operators - other than apartment
    6515-6515 Operators - residential mobile home
    6517-6519 Lessors of real property
    6520-6529 Real estate
    6530-6531 Real estate agents and managers
    6532-6532 Real estate dealers
    6540-6541 Title abstract offices
    6550-6553 Real estate developers
    6590-6599 Real estate
    6610-6611 Combined real estate, insurance, etc

47. Trading
    6200-6299 Security and commodity brokers
    6700-6700 Holding, other investment offices
    6710-6719 Holding offices
    6720-6722 Investment offices
    6723-6723 Management investment, closed-end
    6724-6724 Unit investment trusts
    6725-6725 Face-amount certificate offices
    6726-6726 Unit inv trusts, closed-end
    6730-6733 Trusts
    6740-6779 Investment offices
    6790-6791 Miscellaneous investing
    6792-6792 Oil royalty traders
    6793-6793 Commodity traders
    6794-6794 Patent owners & lessors
    6795-6795 Mineral royalty traders
    6798-6798 REIT
    6799-6799 Investors, NEC

48. Almost Nothing
    4950-4959 Sanitary services
    4960-4961 Steam, air conditioning supplies
    4970-4971 Irrigation systems
    4990-4991 Cogeneration - SM power producer

B Simulation Models

In this appendix we introduce the simulation models we use throughout the thesis.

B.1 Model 1: Independence and Static Parameters

The standard model for continuous returns assumes normally distributed returns and therefore log-normally distributed prices for each asset. This is a static model over time, where each asset is simulated independently. The parameter pairs of annual mean and standard deviation are chosen randomly but fixed for each asset from one of the pairs given in Table B.1. The given parameters are calibrated to DAX asset data from 2016.

mean	-0.45	-0.25	-0.2	-0.15	-0.1	-0.05	0.05	0.08	0.1	0.15	0.2	0.21	0.25	0.3	0.45
volatility	0.5	0.2	0.3	0.25	0.3	0.2	0.25	0.23	0.2	0.4	0.2	0.25	0.3	0.15	0.5

Table B.1: annual mean and volatility pairs for the simulation

Thus the daily returns model is as follows:

$$S_t = S_{t-1}\, e^{R_t},$$
$$R_t = \log\left(\frac{S_t}{S_{t-1}}\right) \sim \mathcal{N}(\mu\Delta, \sigma^2\Delta),$$

where μ is the annual mean return, σ the annual standard deviation or volatility of the returns, $\Delta = \frac{1}{250}$ the reciprocal of the number of trading days in a year and $t = 1, \ldots, T$ the points in time. To compare the results to the real data, 48 assets are simulated over 45 years.

B.2 Model 2: Independence and Markov Switching Parameters

The second model is just an extension of the first model. We allow the parameters to change over time. For this we use a Markov chain, which jumps after a specified number of days. The Markov chain simply chooses another parameter pair for each asset. This is done by using a transition matrix, where the moderate pairs are more likely than the extreme cases from the left and right columns of the Table B.1, and these probabilities stay the same over time.

This is achieved with a discrete, irreducible, homogeneous Markov chain, which admits a stationary distribution. Thus the resulting returns are no longer normally distributed, but mixed normally distributed. Here we have 15 states, which are the mean volatility pairs represented by a vector $\theta = (\mu, \sigma) = ((\mu_1, \sigma_1), \ldots, (\mu_{15}, \sigma_{15}))$. Then the Markov chain is the process $s(t)$ defining which pair is the current one. Thus the returns of an asset are modeled as follows:

$$R_t = \mu_{s(t)} + \sigma_{s(t)}\varepsilon,$$

for $t = 1, \ldots, T$ and $\varepsilon \sim \mathcal{N}(0, 1)$, and $\mathrm{E}[\theta_{s(t)}] = ((A^t v_0)'\mu, (A^t v_0)'\sigma)$ with transition matrix A, where each column is the initial distribution vector v_0, given in Table B.2.

v_0	0.01	0.03	0.05	0.07	0.09	0.1	0.1	0.1	0.1	0.1	0.09	0.07	0.05	0.03	0.01

Table B.2: initial distribution vector v_0

B.3 Model 3: Independence and Markov Market Parameters

In this model the Markov chain controls the state of the market. There are three states: bad, normal and good. The normal state coincides with the (μ, σ) pairs from Table B.1. The pairs for the good and bad state of the market you see in Table B.3.

mean_bad	-0.600	-0.330	-0.260	-0.200	-0.130	-0.065	0.035	0.056	0.070	0.105	0.140	0.175	0.210	0.300	0.315
volatility_bad	0.750	0.300	0.450	0.375	0.450	0.300	0.375	0.345	0.300	0.600	0.300	0.375	0.450	0.225	0.750
mean_good	-0.315	-0.175	-0.140	-0.105	-0.070	-0.035	0.065	0.104	0.130	0.195	0.260	0.273	0.325	0.390	0.600
volatility_good	0.350	0.140	0.210	0.175	0.210	0.140	0.175	0.161	0.140	0.280	0.140	0.175	0.210	0.105	0.350

Table B.3: annual mean and volatility pairs in the good and bad state of the market for the simulation

So, we are given 15 return models with parameters $\theta^i = (\mu^i, \sigma^i) = ((\mu_1, \sigma_1), (\mu_2, \sigma_2), (\mu_3, \sigma_3))$ for the good, the normal and the bad state and $i = 1, \ldots, 15$. At first one of these models is assigned to each asset. Then, the initial state of the market is chosen uniform randomly. Now, the Markov chain $s(t)$ defines the current state of the market and the returns of the asset with asset model i is modeled as follows:

$$R_t^i = \mu_{s(t)}^i + \sigma_{s(t)}^i \varepsilon,$$

for $t = 1, \ldots, T$ and $\varepsilon \sim \mathcal{N}(0,1)$, and $\mathrm{E}[\theta_{s(t)}^i] = ((A^t v_0)' \mu^i, (A^t v_0)' \sigma^i)$ with transition matrix A, where each column is the initial distribution vector v_0 preferring the normal state over the other states, given in Table B.4.

v_0	0.25	0.50	0.25

Table B.4: initial market distribution vector v_0

B.4 Model 4: Correlation and Static Parameters

Here, the asset returns are simulated multivariate normally, where we assume high correlation between some of the assets. This is created so that there are groups with highly internal correlation and low correlation towards assets outside the group. As in model 1 in Section B.1, the (μ, σ) pairs for each asset are chosen randomly from the values in Table B.1 and are fixed over time. There are three different settings which result in 20, 24, and 28 groups. The general model of the daily returns is as follows:

$$R \sim \mathcal{N}(\mu, \Sigma),$$

where μ is the vector containing the daily mean return of each asset, and Σ is the variance-covariance matrix with the individual daily variances on the diagonal and the covariances on the off-diagonal elements. To vary the variance-covariance matrix in terms of correlation, the following relation is helpful:

$$\Sigma = \mathrm{diag}(\sigma) \times \mathrm{Corr} \times \mathrm{diag}(\sigma),$$

where Corr is the correlation matrix with ones on the diagonal and the pairwise correlations $\text{corr}(R_i, R_j) \in [-1, 1]$ as off-diagonal elements and $\text{diag}(\sigma)$ is the matrix with the daily standard deviation of each asset on the diagonal and zero elements otherwise.

Then we have the problem that, if we choose a correlation matrix it is hardly ever a true correlation matrix since it may not be positive semi-definite by just choosing the correlations. But applying the method proposed by Higham (2002) we can use the nearest true correlation matrix. A short overview on the method is given in Appendix C. In Table B.5 we see the grouping of the assets in three scenarios with 20, 24 and 28 groups. Again 48 assets are simulated over 45 years.

	6 assets	5 assets	4 assets	3 assets	2 assets	1 asset
20 groups	(17,18,26,28,34,45)	(1,8,10,20,24), (3,4,7,30,48)	(11,15,22,25), (35,36,40,41)	(2,5,47), (6,9,12), (13,27,33)	(14,16), (21,23), (31,44)	(19), (29), (32), (37), (38), (39), (42), (43), (46)
24 groups	(17,18,26,28,34,45)	(1,8,10,20,24)	(4,7,30,48), (11,15,22,25)	(2,5,47), (6,9,12)	(14,16), (21,23), (31,44), (35,40), (36,41)	(3), (13), (19), (27), (29), (32), (33), (37), (38), (39), (42), (43), (46)
28 groups		(1,8,10,20,24)	(4,7,30,48)	(2,5,47), (11,15,22), (17,18,26), (25,36,41)	(14,16), (21,23), (31,44), (34,45), (35,40)	(3), (6), (9), (12), (13), (19), (27), (28), (29), (32), (33), (37), (38), (39), (42), (43), (46)

Table B.5: Grouping of assets for 20, 24 and 28 groups.

This results in the correlation matrices in Figure B.1.

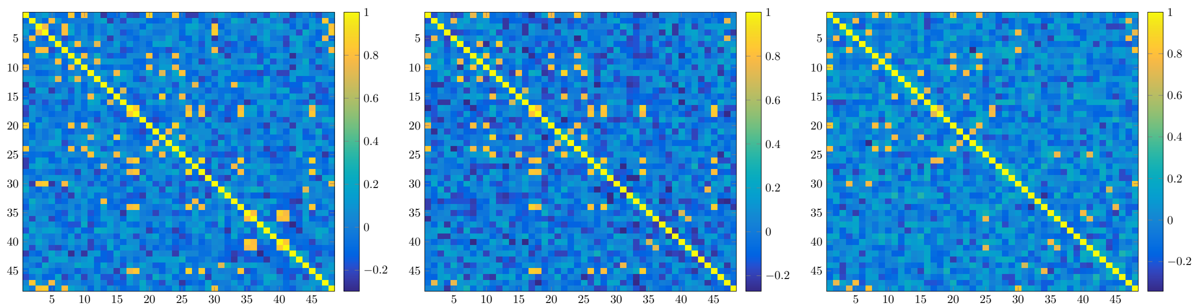


Figure B.1: Correlation matrix with 20, 24 and 28 groups

B.5 Model 5: Comonotonicity and Static Parameters

This is a special case of Model 4, where we assume perfect positive correlation for all assets within one of the L groups and zero correlation between the groups. As the algorithm by Higham only provides the nearest correlation matrix, we use a different model to perform this task. From Definition 5.4 we know that in the case of same location and scale family, every comonotonic asset can be derived by:

$$X_i \stackrel{d}{=} a_i Y + b_i$$

for $i = 1, 2, \ldots, L$. Then L different samples of the standard normal distribution serve as initial Y, together with the randomly but fixed chosen mean and standard deviation from Table B.1 we result in L comonotonic groups with the specified parameters. Within this model it is possible to freely choose the number of groups without specifying the correlation matrix. The size of the groups is chosen randomly. This results in correlation matrices like the one for 20 groups in Figure B.2.

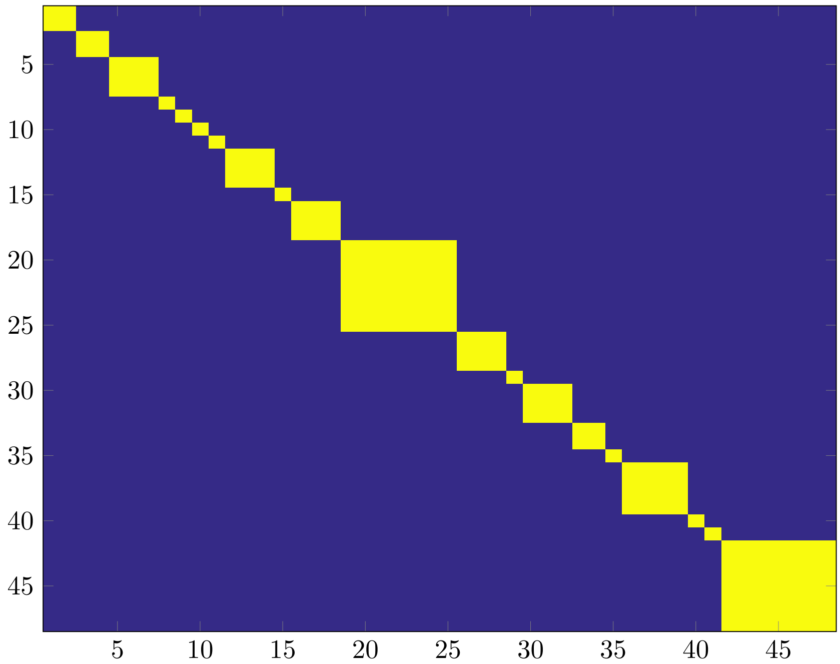


Figure B.2: Comonotonic correlation matrix with 20 groups

B.6 Model 6: Correlation and Markov Switching Parameters

As in Model 2, the parameters of the assets can change over time with a Markov chain, but now we have a correlation structure in addition. That is, in each Markov time step the asset returns are drawn from the multivariate normal distribution with the current parameters but always the same correlation structure:

$$R_t \sim \mathcal{N}(\mu_t, \operatorname{diag}(\sigma_t) \times \mathrm{Corr} \times \operatorname{diag}(\sigma_t)).$$

B.7 Model 7: Comonotonicity and Markov Switching Parameters

The same method as in Model 5 is used to achieve perfect comonotonicity in the Markov parameters setting. The Markov chain uses the same parameters as in Model 3.

C Correlation Matrix Algorithm

Here, we present the main contents of the paper "Computing the nearest correlation matrix - a problem from finance" by Nicholas J. Higham Higham (2002). For the simulation of correlated data it is necessary to use a proper correlation matrix. However, the main problem is to choose a correlation matrix that is positive semidefinite. With the algorithm proposed by Higham this can be achieved by starting with a pseudo correlation matrix, which is symmetric with unit diagonal and all other entries between -1 and $+1$. Then, the algorithm finds the nearest correlation matrix which is also positive semidefinite.

C.1 Theory

In finance it is a common problem that a correlation matrix is estimated from some available data. However, this leads to an approximate correlation matrix C_{appr}, which may not have all properties of a correlation matrix.

Definition C.1. *A* ***correlation matrix*** *C is a symmetric, positive semidefinite matrix with unit diagonal whose entries c_{ij} are equal to the correlation coefficient of the i^{th} and j^{th} element of a random vector. Therefore, all entries have values $-1 \leq c_{ij} \leq +1$.*

To find the nearest correlation matrix, a distance must be specified:

$$\gamma(C_{appr}) = \min\{\|C_{appr} - C\| : \text{C is a correlation matrix}\}, \tag{C.1}$$

where $\|\cdot\|$ is a weighted version of the Frobenius norm ($\|X\|_F^2 = \sum_{i,j} x_{ij}$). The W-norm is easier to work with (compared to a norm, which needs the Hadamard product) in the coding and the calculation of the nearest correlation matrix. It is defined as

$$\|X\|_W = \|W^{1/2} X W^{1/2}\|_F,$$

where W is a symmetric positive definite matrix and in this case can be restricted to a diagonal matrix. The weighting matrix is not able to express confidence to individual elements of C_{appr}, but when the correlation between the i^{th} and j^{th} element is known accurately the elements w_{ii} and w_{jj} can be assigned with high values. This makes sure that the corresponding correlation coefficients in the nearest correlation matrix are close to the original ones. In order to find such a correlation matrix it is obvious that the solution must lie in the intersection of two closed convex sets:

$$S = \{X = X^T \in \mathbb{R}^{n \times n} : X \text{ is positive semidefinite}\},$$

$$U = \{X = X^T \in \mathbb{R}^{n \times n} : x_{ii} = 1 \text{ for all } i\}.$$

In this context it is clear that $C_{appr} \in U$ and $|c_{ij}^{appr}| \leq 1$. However, the following can also be performed with a general symmetric matrix A. As S and U are both closed and convex their intersection is closed and convex, too. Thus, the minimum in (C.1) can be achieved uniquely, which follows from standard results in approximation theory, see e.g. Luenberger (1997). To find the best solution, the alternating projection algorithm for computing the nearest Euclidean distance matrix by Glunt, Hayden, Hong, and Wells (1990) is modified. At first, an inner product on $\mathbb{R}^{n\times n}$ that induces the W-norm must be defined,

$$\langle X, Y\rangle = \text{trace}(X^T W Y W).$$

With this a normal cone of a convex set $K \in \mathbb{R}^{n\times n}$ at $Y \in K$ is

$$\partial K(Y) = \{X = X^T \in \mathbb{R}^{n\times n} : \langle X, Y\rangle = \sup_{Z\in K} \langle X, Z\rangle\}.$$

Then, the optimal solution X^* to the nearest correlation matrix problem can be characterized by $\langle Z - X^*, A - X^*\rangle \leq 0$ for all $Z \in S \cap U$. Thus, $A - X^* \in \partial(S \cap U)(X^*)$, which is the normal cone to $S \cap U$ at X^*. As all positive definite correlation matrices are in the interiors of S and U the characterization can be altered to

$$A - X^* \in \partial S(X^*) + \partial U(X^*). \tag{C.2}$$

To solve the problem, the normal cones of S and U must be specified, after some work we get for $A \in U$

$$\partial U(A) = \{W^{-1}\text{diag}(\theta_i)W^{-1}, \theta_i \in \mathbb{R}\},$$

and using spectral decomposition we have for $A \in S$

$$\begin{aligned}\partial S(A) = \{Y : WYW = -VDV^T, &\text{where } V \in \mathbb{R}^{n\times p} \text{ has orthonormal columns spanning null}(X)\\ &\text{and } D = \text{diag}(d_i) \text{ positive semidefinite}\}.\end{aligned}$$

Together with the characterization of the solution given in (C.2), this leads to the main theorem in Higham (2002):

Theorem C.2. *The correlation matrix X^* solves equation* (C.1) *if and only if*

$$X^* = A + W^{-1}(VDV^T + \text{diag}(\theta_i))W^{-1}, \tag{C.3}$$

where $V \in \mathbb{R}^{n\times p}$ has orthonormal columns spanning null(X^*)*, $D =$* diag(d_i) *positive semidefinite, and arbitrary $\theta_i \in \mathbb{R}$.*

In our case, where we are given an approximate correlation matrix, the theorem can be extended.

Theorem C.3. *Let $C_{appr} = C_{appr}^T$ have diagonal elements $c_{ii}^{appr} \geq 1$ and let W be diagonal. Then, in Theorem C.2, $\theta_i \leq 0$ for all i. Moreover, if C_{appr} has t nonpositive eigenvalues, then the nearest correlation matrix has at least t eigenvalues equal zero.*

C.2 Computations

To actually compute a nearest correlation matrix, it is necessary to project from the symmetric matrices onto the correlation matrices with respect to the W-norm. This can be done in two steps. First, the projection onto the set S and second, onto the set U, where S and U are defined

as in Section C.1.
Projection onto S:

$$P_S(A) = W^{-1/2}((W^{1/2}AW^{1/2})_+)W^{1/2},$$

where for $W^{1/2}AW^{1/2} = QDQ^T$ the spectral decomposition of $W^{1/2}AW^{1/2}$ with $D = \operatorname{diag}(\lambda_i)$ and Q orthogonal we have $(W^{1/2}AW^{1/2})_+ = Q \operatorname{diag}(\max(\lambda_i, 0))\, Q^T$.
Projection onto U:

$$P_U(A) = A - W^{-1} \operatorname{diag}(\theta_i)W^{-1},$$

where $\theta = [\theta_1, \dots, \theta_n]^T$ is the solution of the linear system $(W^{-1} \circ W^{-1})\theta = \operatorname{diag}(A - I)$. For a diagonal W this simplifies to:

$$P_U(A) = \begin{cases} a_{ij}, & i \neq j, \\ 1, & i = j. \end{cases}$$

To find the nearest correlation matrix the alternating projections $A \leftarrow P_U(P_S(A))$ are used together with a Dykstra correction, which guarantees the convergence to an optimal point, see Dykstra (1983). In general the algorithm looks as follows:

Algorithm C.4. *Given a symmetric $A \in \mathbb{R}^{n\times n}$ this algorithm computes the nearest correlation matrix to A in the W-norm.*

$\Delta S_0 = 0, Y_0 = A$
for $k = 1, 2, \dots$
  $R_k = Y_{k-1} - \Delta S_{k-1}$ % ΔS_{k-1} `is Dykstra's correction.`
  $X_k = P_S(R_k)$
  $\Delta S_k = X_k - R_k$
  $Y_k = P_U(X_k)$
end

It can be shown that the algorithm achieves, at best, linear convergence. When programming this code, it is necessary to end the for loop at some appropriate precision level. This means the resulting correlation matrix has the desired properties when it is rounded up to this level. At the end of the for loop in Algorithm C.4 the following condition is used as stopping criterion:

$$\max\left\{\frac{\|X_k - X_{k-1}\|_\infty}{\|X_k\|_\infty}, \frac{\|Y_k - Y_{k-1}\|_\infty}{\|Y_k\|_\infty}, \frac{\|Y_k - X_{k-1}\|_\infty}{\|Y_k\|_\infty}\right\} \leq \text{tol},$$

with the notation of the algorithm, tol a tolerance and $\|\cdot\|_\infty$ the maximum absolute row sum norm.

Matlab Code C.5. *Our realization of the Algorithm C.4:*

```
1  % Input: symmetric matrix A
2  %        tolerance tol
3  % Output: correlation matrix C
4  %         diagonal matrix with eigenvalues D
5  %         matrix with eigenvectors V
6
7  function [C,V,D] = nearcorr(A,tol)
8  [dim,¬] = size(A);
9  % weights for W (diagonal)
10 A_up = triu(A,1);
11 [row,column] = find(A_up>0.7); % find entries with high positive correlation
12 [n_corr,¬] = size(row);
13 % create weight matrix
14 W = eye(dim);
15 for i = 1:n_corr
16     W(row(i),row(i)) = 5;
17     W(column(i),column(i)) = 5;
18 end
19 W_sqr = sqrt(W); % squareroot of W
20 W_sqr_rec = diag(1./diag(W_sqr)); % inverse of W_sqr
21
22 % correlation algorithm
23 dS = zeros(dim); % Dykstra correction
24 C = A; % nearest correlation matrix
25 X = A; % projection onto S
26 max_error = 100; %  error term
27 n = 0;
28 while max_error > tol
29     n = n+1;
30     R = C-dS; % correction by Dykstra
31     [Q,D] = eig((W_sqr*R*W_sqr)); % spectral decomposition of weighted R
32     X_prev = X; % old projection onto S
33     X = W_sqr_rec *(Q * diag(max(diag(D),0))* Q')*W_sqr_rec; %projection onto S
34     dS = X-R; % new Dykstra correction term
35     C_prev = C; % old nearest correlation matrix
36     C = X;
37     C(eye(dim)≠0) = 1; % projection onto U
38     % error terms
39     error1 = norm(X-X_prev,Inf)/norm(X,Inf);
40     error2 = norm(C-C_prev,Inf)/norm(C,Inf);
41     error3 = norm(C-X,Inf)/norm(C,Inf);
42     max_error = max([error1,error2,error3]);
43 end
44
45 [V,D] = eig(C); % spectral decomposition of the nearest correlation matrix
46 end
```

The choice for W in the code above is already adjusted to finding the nearest correlation matrix to a given pseudo correlation matrix. Therefore, all given entries which imply strong correlation (here between 0.7 and 1) are meant to be accurately known. If also strong negative correlation is of interest, the part "weights for W diagonal" after line 9 in the code above should be changed to also choosing high weights for correlation coefficients between -1 and -0.7. To test our code we take the example from the paper, i.e. a positive definite matrix

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}.$$

With tol $= 10^{-8}$ my code needs 17 iterations to stop and the nearest correlation matrix is

$$C = \begin{pmatrix} 1.0000 & -0.8084 & 0.1916 & 0.1068 \\ -0.8084 & 1.0000 & -0.6562 & 0.1916 \\ 0.1916 & -0.6562 & 1.0000 & -0.8084 \\ 0.1068 & 0.1916 & -0.8084 & 1.0000 \end{pmatrix}$$

with $\|A - C\|_F = 2.1337$. The spectral decomposition of C leads to $C = QDQ^T$ with

$$Q = \begin{pmatrix} 0.4011 & -0.6177 & 0.5823 & 0.3442 \\ -0.5823 & 0.3442 & 0.4011 & 0.6177 \\ 0.5823 & 0.3442 & -0.4011 & 0.6177 \\ -0.4011 & -0.6177 & -0.5823 & 0.3442 \end{pmatrix}$$

and

$$D = \begin{pmatrix} 2.3450 & 0 & 0 & 0 \\ 0 & 1.4505 & 0 & 0 \\ 0 & 0 & 0.2044 & 0 \\ 0 & 0 & 0 & -0.0000 \end{pmatrix}.$$

D Additional Figures

In the following, we provide simulation graphs for the first 100 simulations corresponding to the tables from Chapter 4 and Chapter 5.

D.1 Chapter 4: Naive Diversification

Here, we provide results for the Simulation Models 1, 2, and 3 with independent assets. In Model 1 the parameters are static. In Model 2 we have Markov switching parameters, and in Model 3 the parameters change according to the state of the market.

The first figures show the performance of the sample (unconstrained) mean-variance portfolio, the full $\frac{1}{N}$ portfolio and the portfolio choosing the 24 historically best (by Sharpe ratio) assets in an equal weight portfolio, corresponding to Section 4.4.2.

Then, we show a comparison of the three mean-variance methods, the constrained (without shortselling) and unconstrained (sample) mean-variance portfolio and the minimum-variance portfolio, corresponding to Section 4.5.1.

Finally, we provide a comparison of the performance of the constrained (without short-selling) mean-variance portfolio with the full naive portfolio and the one with only 24 assets, corresponding to Section 4.6.2.

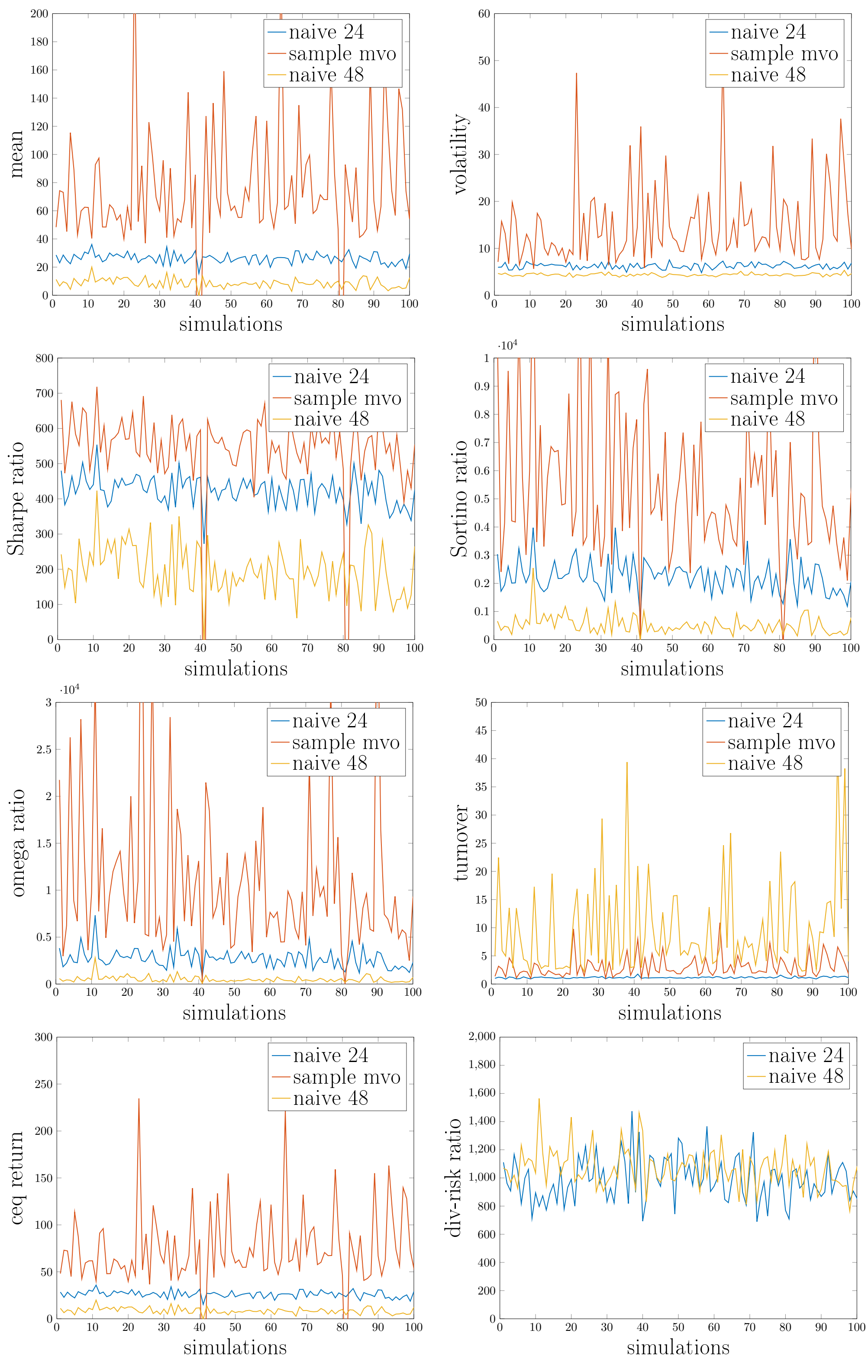


Figure D.1: Model 1: Annual performance (in percent, except turnover). Simulation model 1, 48 assets, 180 months estimation window, 360 months total testing period. First 100 simulation sets.

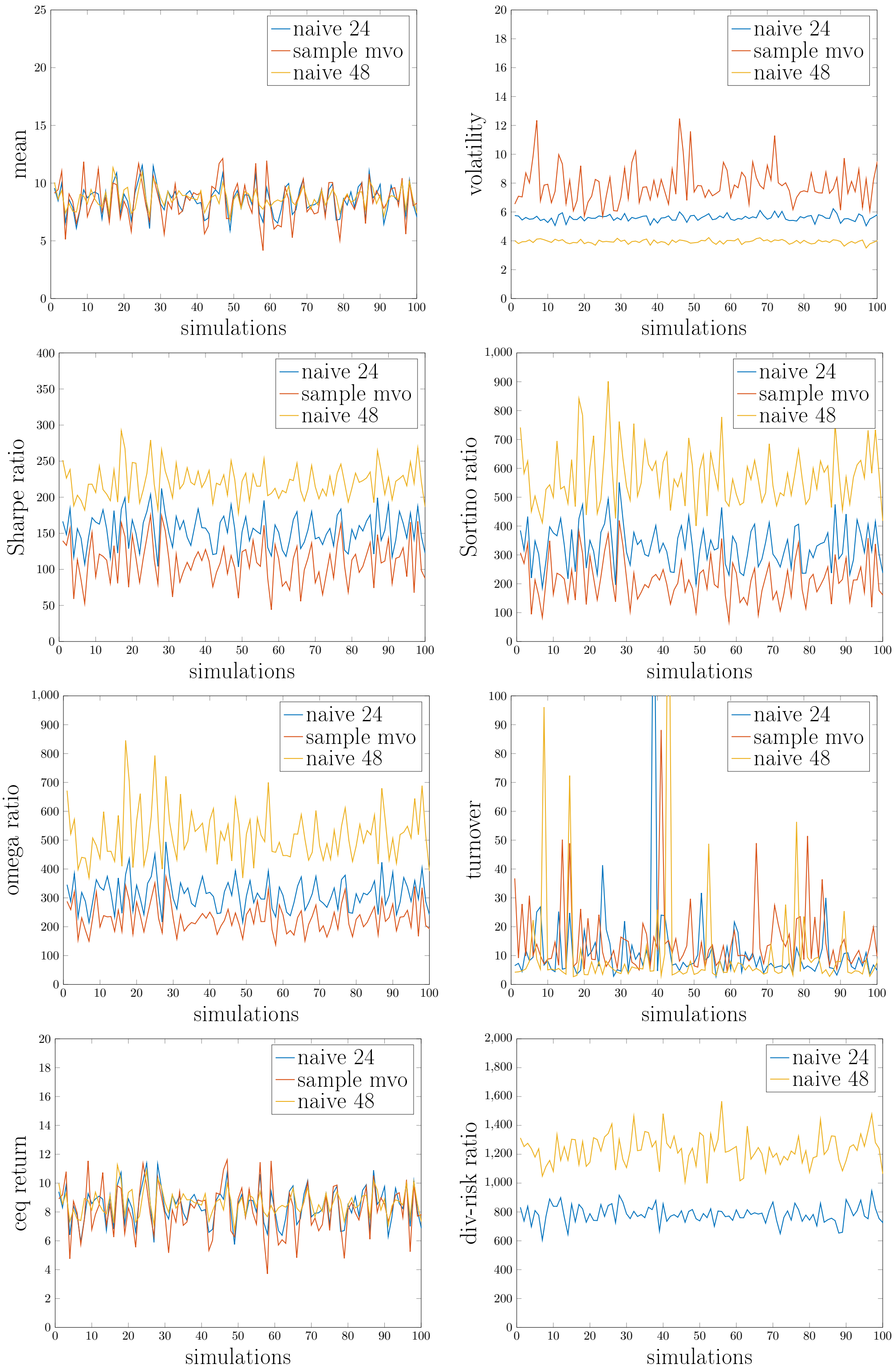


Figure D.2: Model 2: Annual performance (in percent, except turnover). Simulation model 2, 48 assets, 180 months estimation window, 360 months total testing period. First 100 simulation sets.

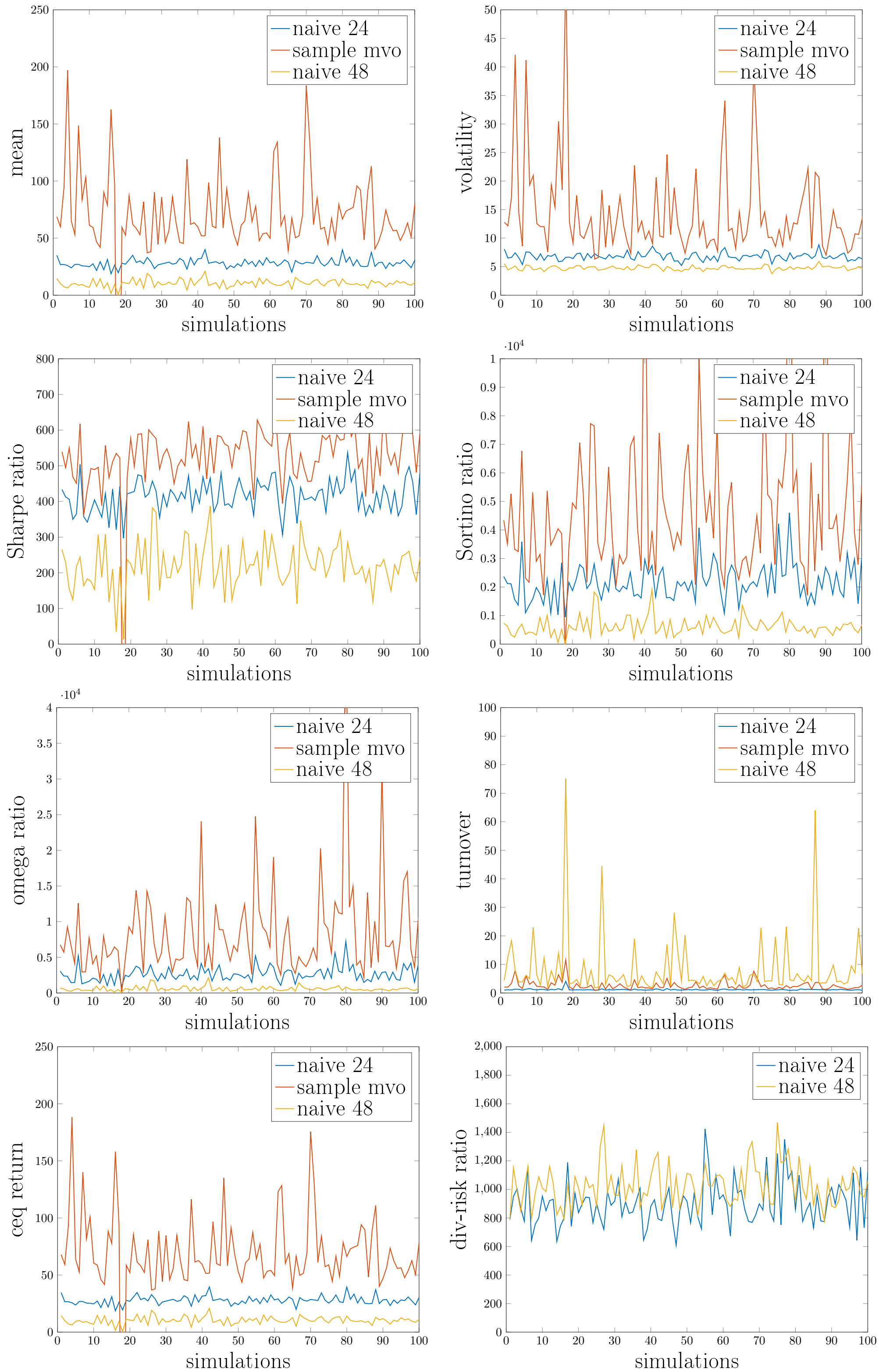


Figure D.3: Model 3: Annual performance (in percent, except turnover). Simulation model 3, 48 assets, 180 months estimation window, 360 months total testing period. 1000 simulation sets.

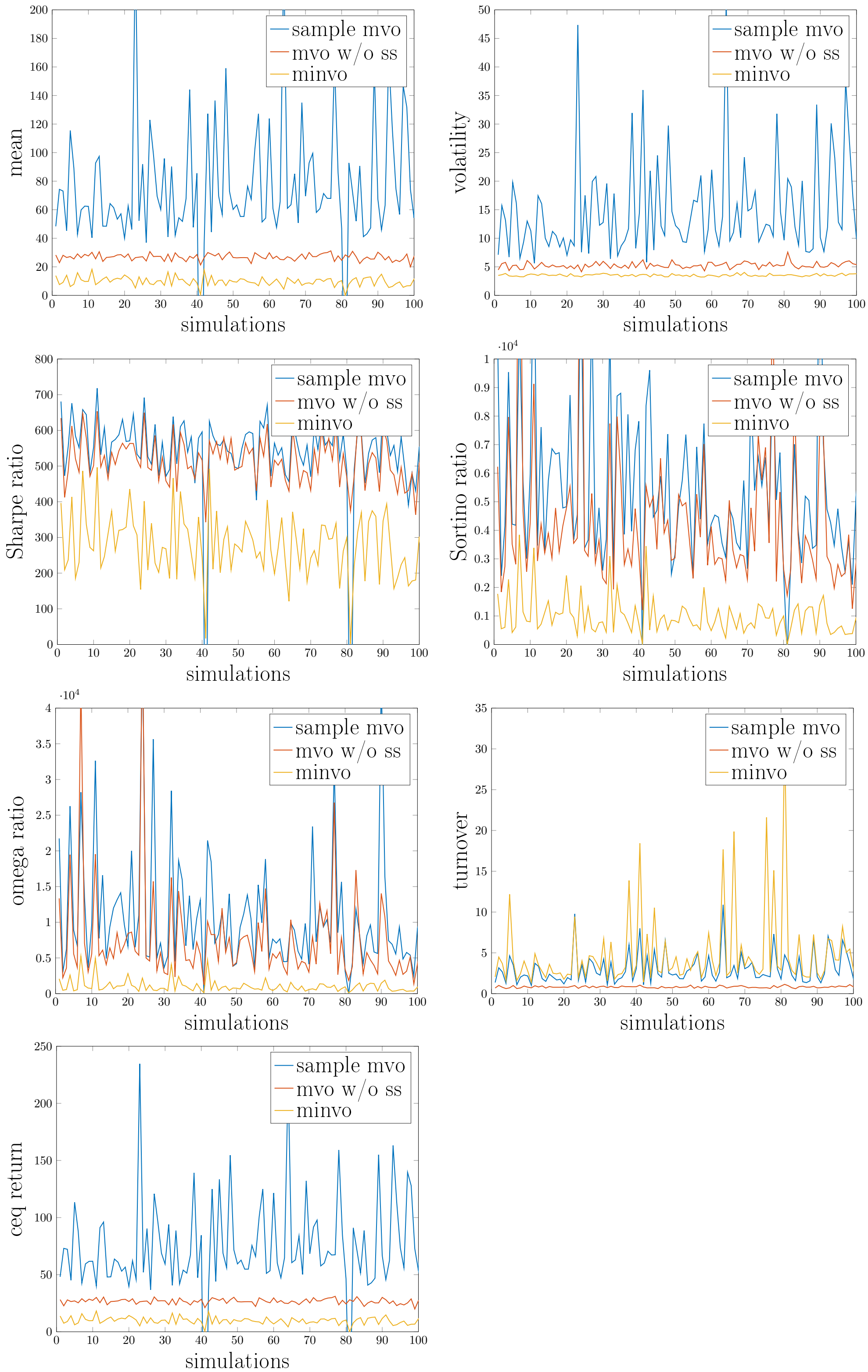


Figure D.4: Model 1: Mean-variance performance (in percent, except turnover) of the sample, no short-selling and minimum variance portfolio. Simulation model 1, 48 assets, 180 months estimation window, 360 months total testing period. First 100 simulation sets.

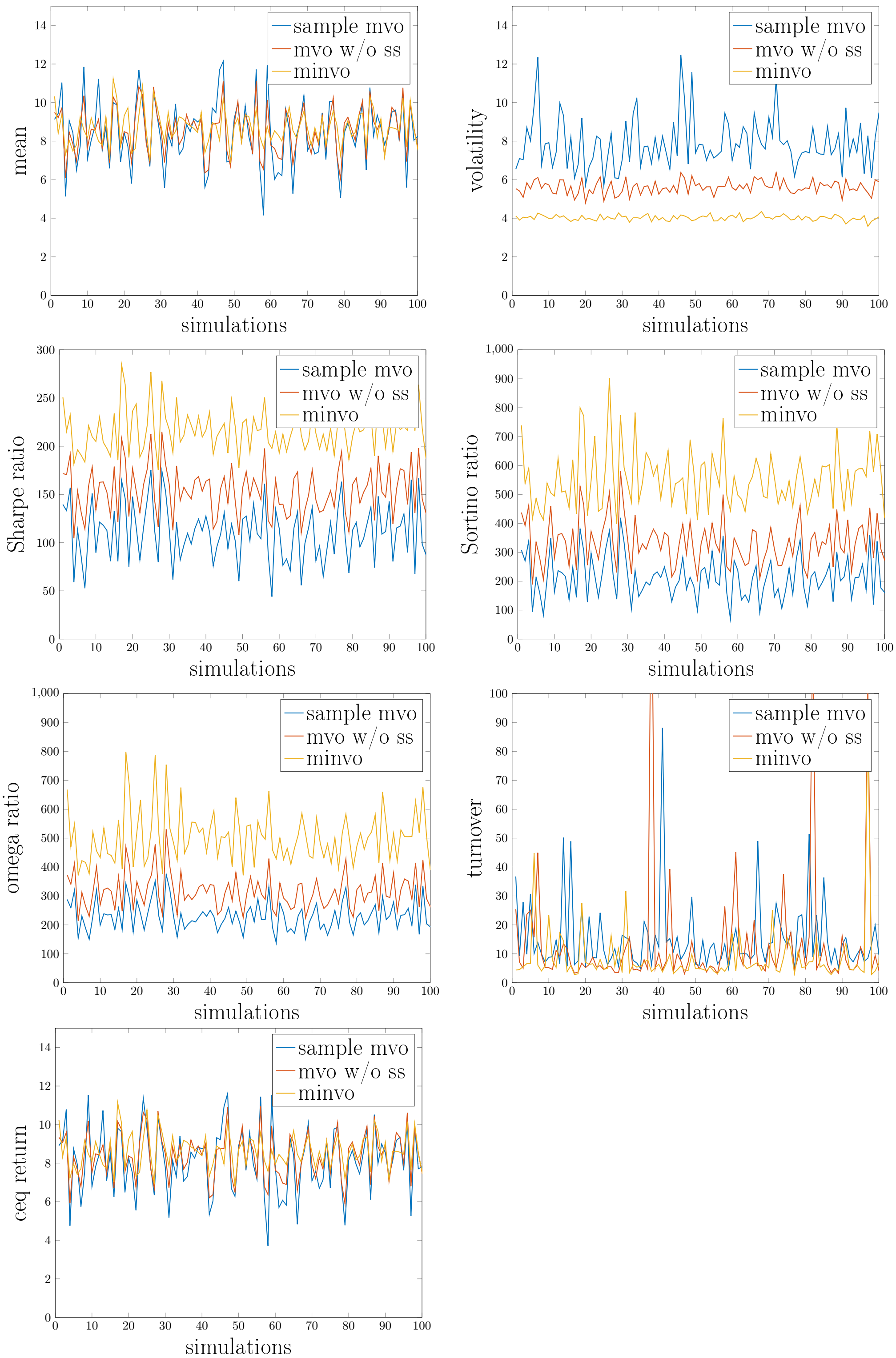


Figure D.5: Model 2: Mean-variance performance (in percent, except turnover) of the sample, no short-selling and minimum variance portfolio. Simulation model 2, 48 assets, 180 months estimation window, 360 months total testing period. First 100 simulation sets.

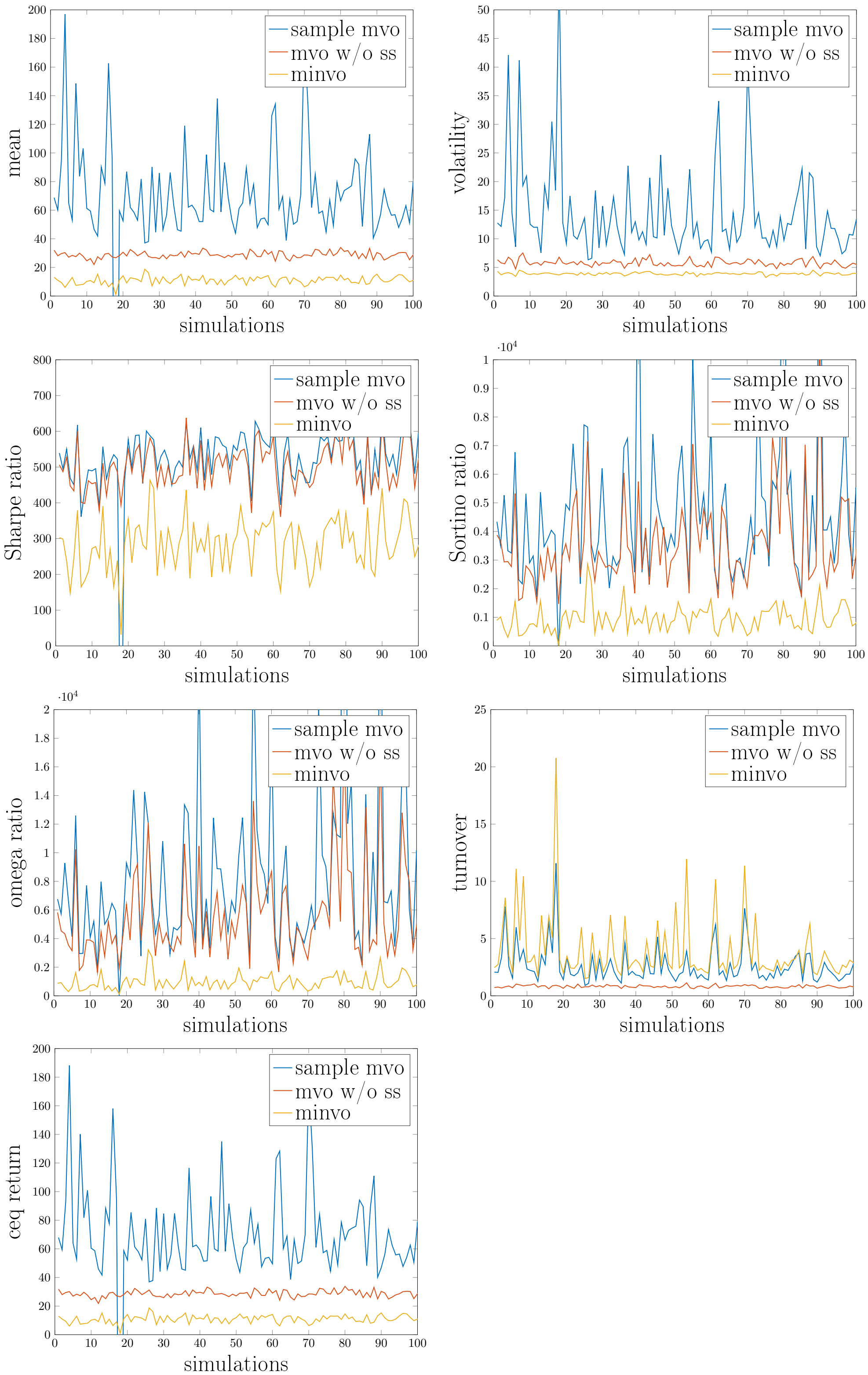


Figure D.6: Model 3: Mean-variance performance (in percent, except turnover) of the sample, no short-selling and minimum variance portfolio. Simulation model 3, 48 assets, 180 months estimation window, 360 months total testing period. First 100 simulation sets.

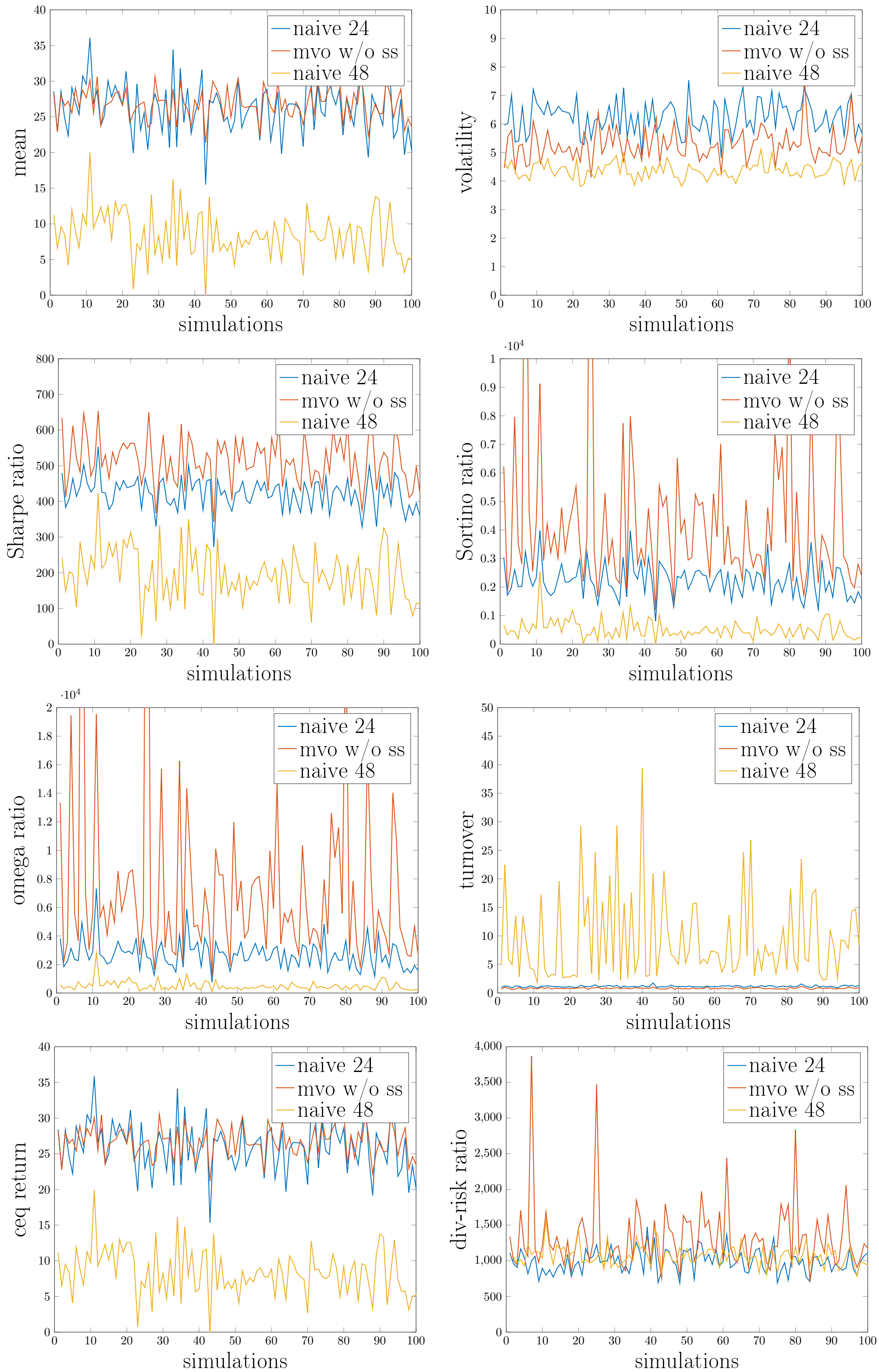


Figure D.7: Model 1: Annual performance (in percent, except turnover). Simulation model 1, 48 assets, 180 months estimation window, 360 months total testing period. First 100 simulation sets.

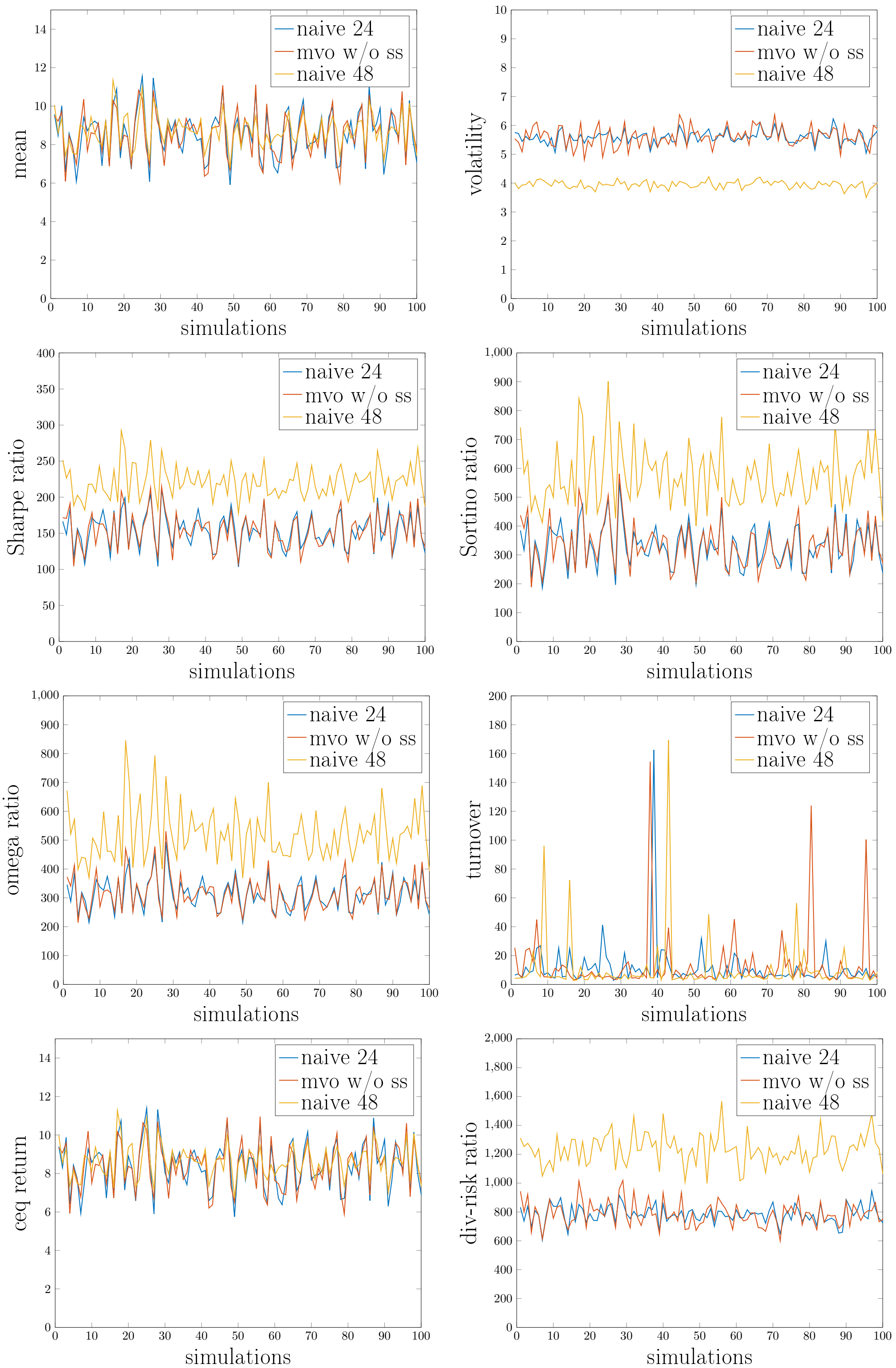


Figure D.8: Model 2: Annual performance (in percent, except turnover). Simulation model 2, 48 assets, 180 months estimation window, 360 months total testing period. First 100 simulation sets.

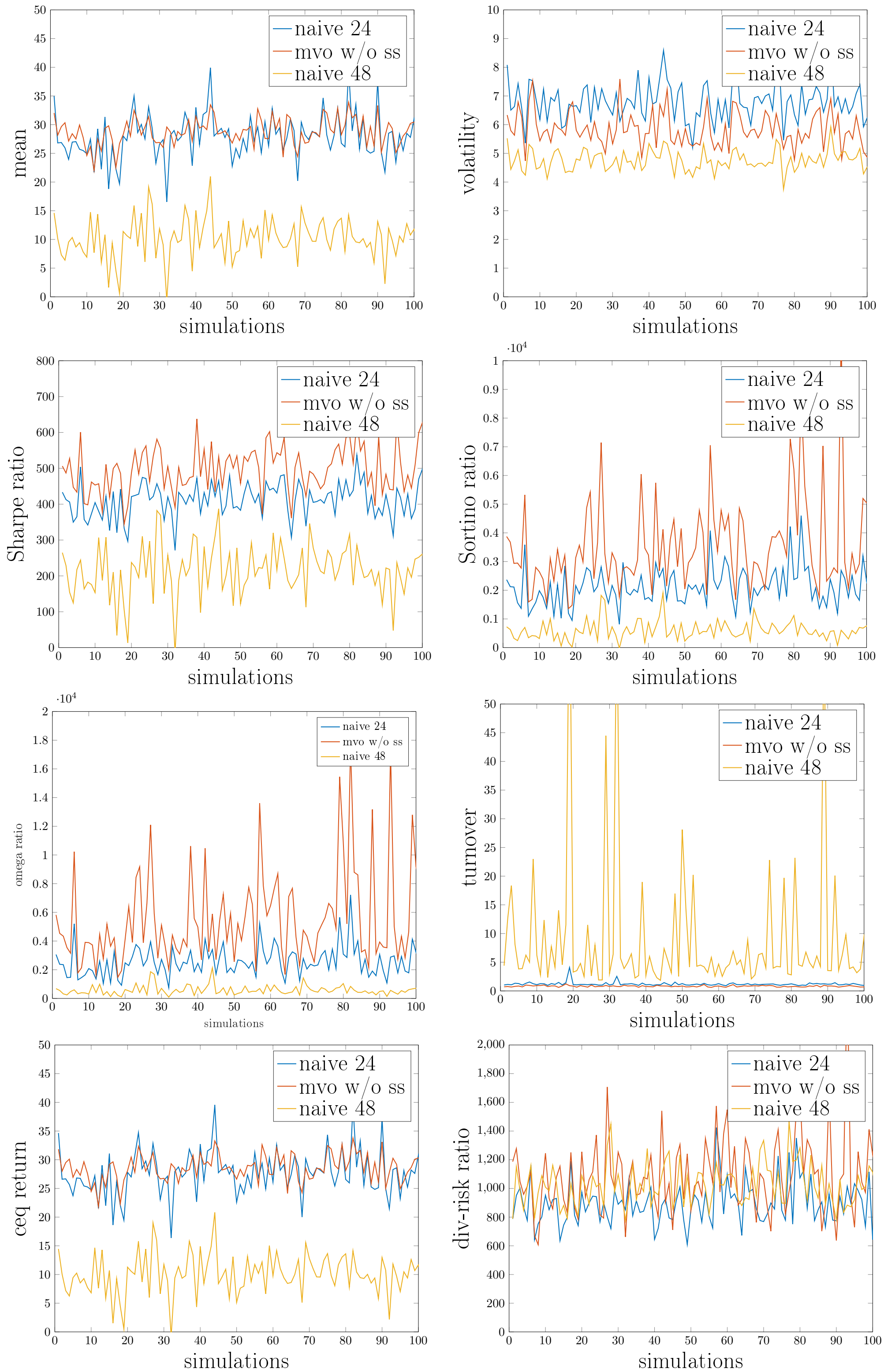


Figure D.9: Model 3: Annual performance (in percent, except turnover). Simulation model 3, 48 assets, 180 months estimation window, 360 months total testing period. First 100 simulation sets.

D.2 Chapter 5: Naive Diversification with Dependence Structure

Here, we provide a comparison of the performance of the constrained (without shortselling) mean-variance portfolio with the full naive portfolio and the fixed and free cluster portfolio.

First, we show the results for simulation model 4 with correlation and static parameters. This is followed by the results for Model 6 with correlation and switching parameters. For each model we have versions for 20, 24 and 28 groups. These figures correspond to the ones in Section 5.2.3.

This is followed by the results for simulations with a comonotonic dependence structure. Model 5 has static and Model 7 has switching parameters. These are cutouts from the figures in Section 5.4.

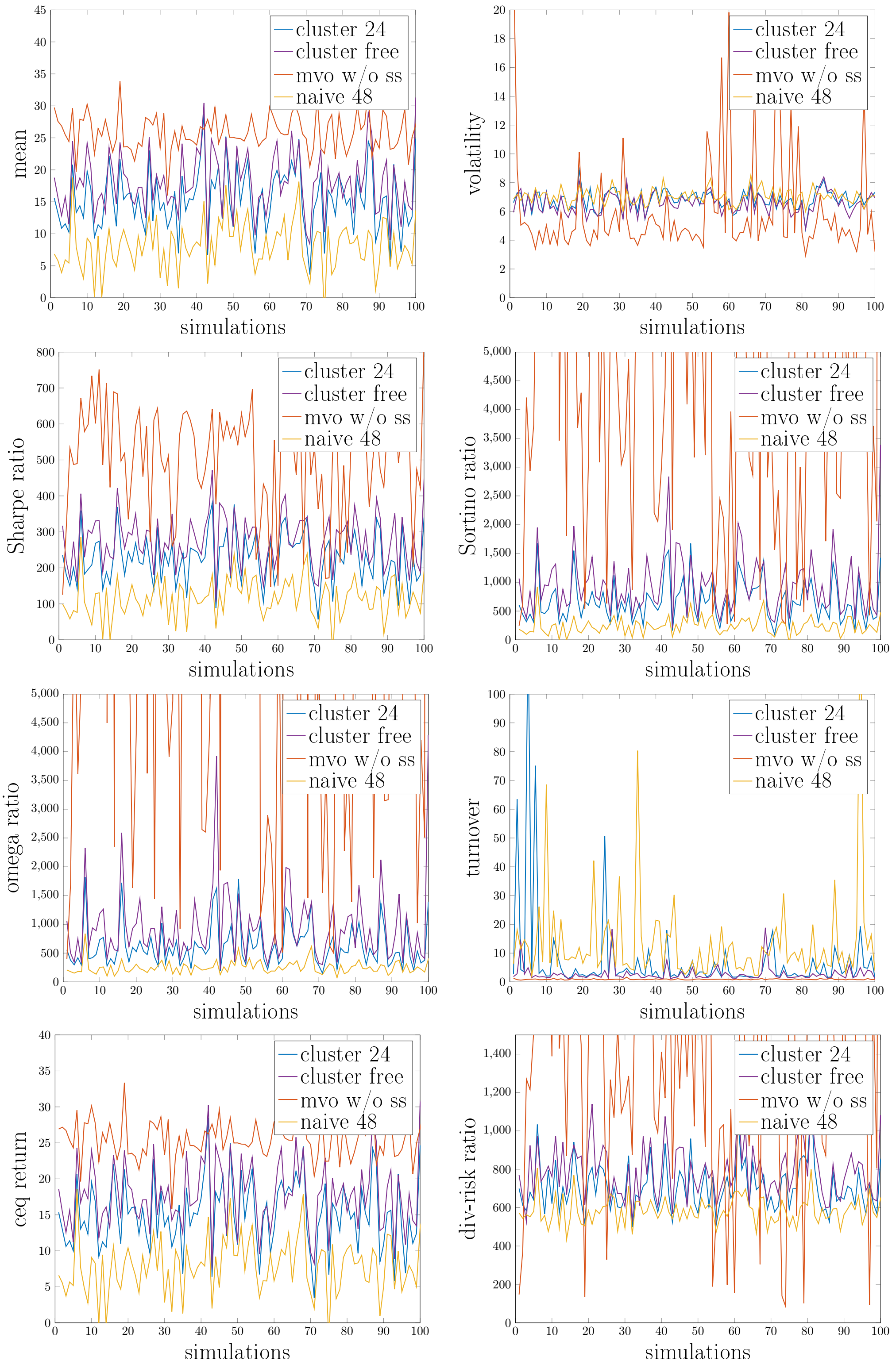


Figure D.10: Model 4.20: Annual performance (in percent, except turnover). Simulation model 4, 48 assets, 20 correlation groups, 180 months estimation window, 360 months total testing period. First 100 simulation sets.

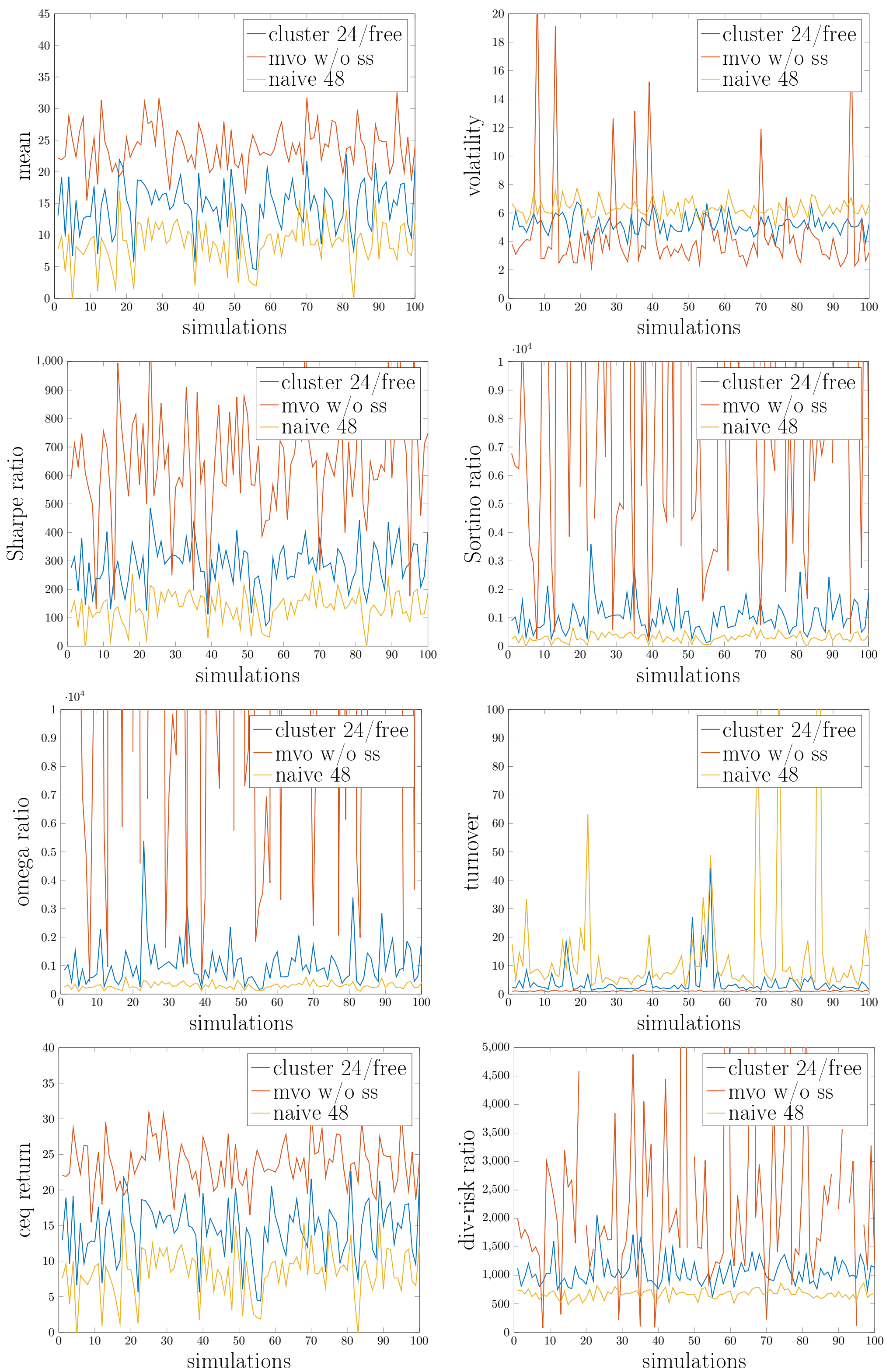


Figure D.11: Model 4.24: Annual performance (in percent, except turnover). Simulation model 4, 48 assets, 24 correlation groups, 180 months estimation window, 360 months total testing period. First 100 simulation sets.

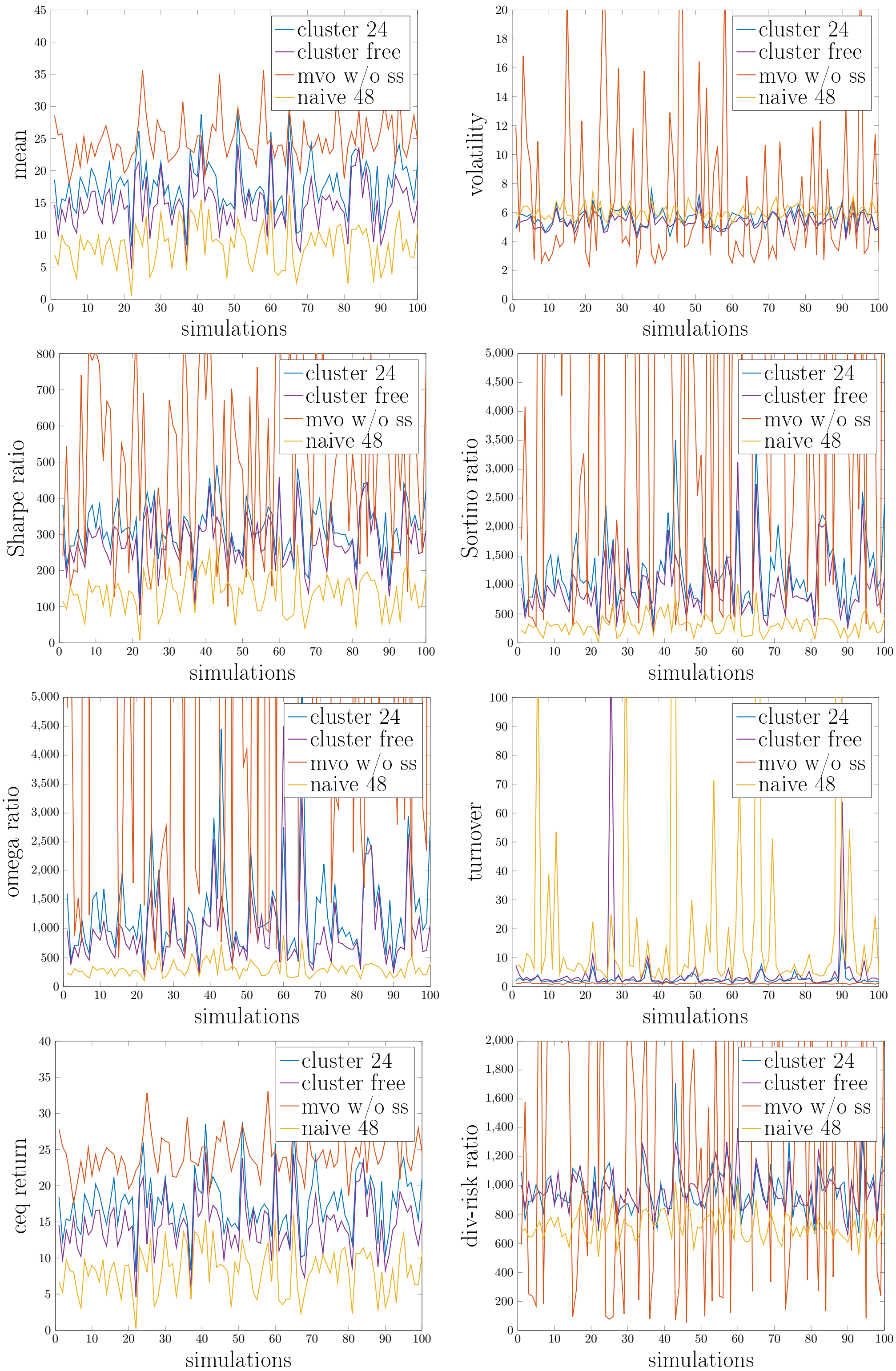


Figure D.12: Model 4.28: Annual performance (in percent, except turnover). Simulation model 4, 48 assets, 28 correlation groups, 180 months estimation window, 360 months total testing period. First 100 simulation sets.

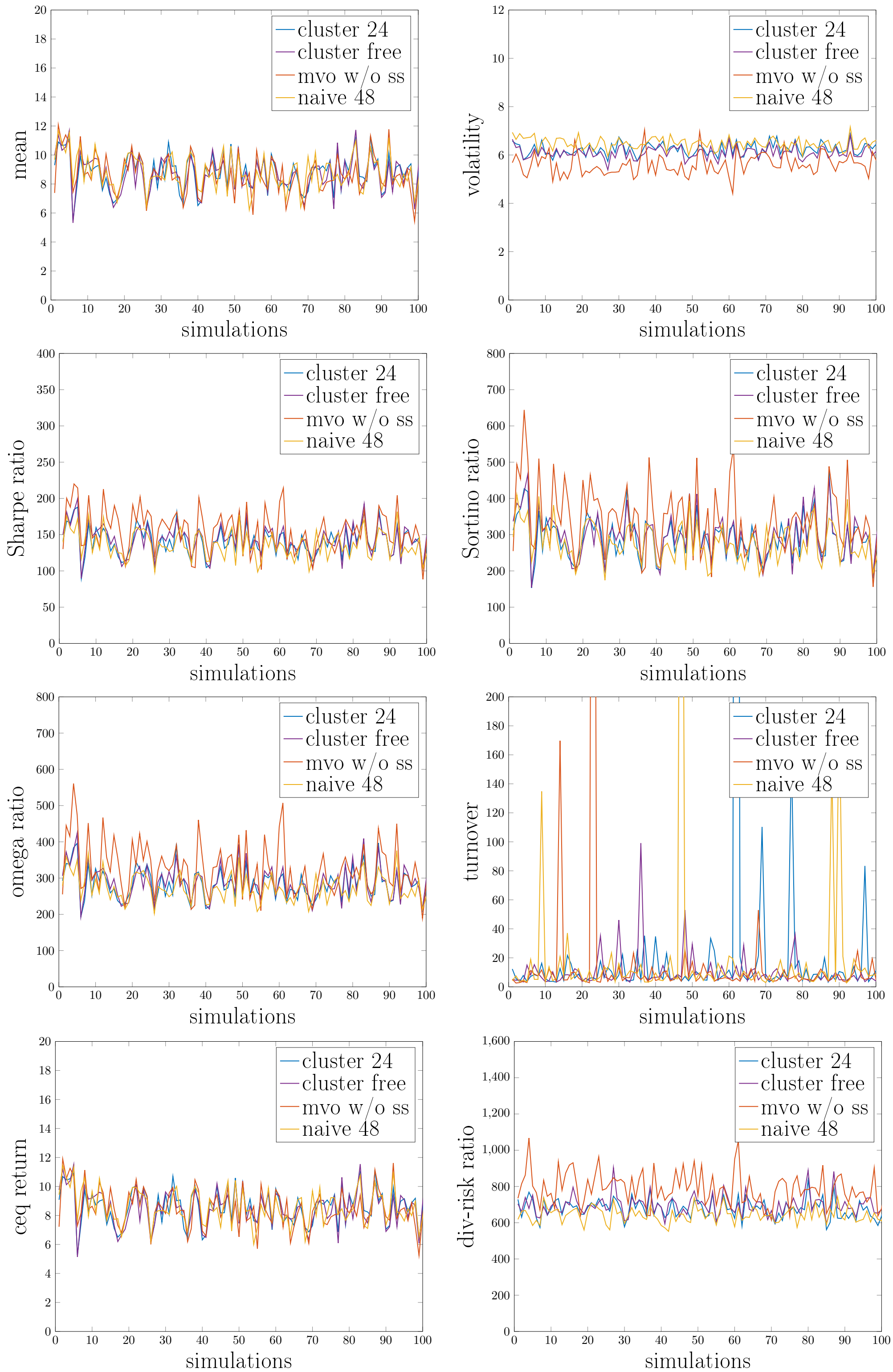


Figure D.13: Model 6.20: Annual performance (in percent, except turnover). Simulation model 6, 48 assets, 20 correlation groups, 180 months estimation window, 360 months total testing period. First 100 simulation sets.

Figure D.14: Model 6.24: Annual performance (in percent, except turnover). Simulation model 6, 48 assets, 24 correlation groups, 180 months estimation window, 360 months total testing period. First 100 simulation sets.

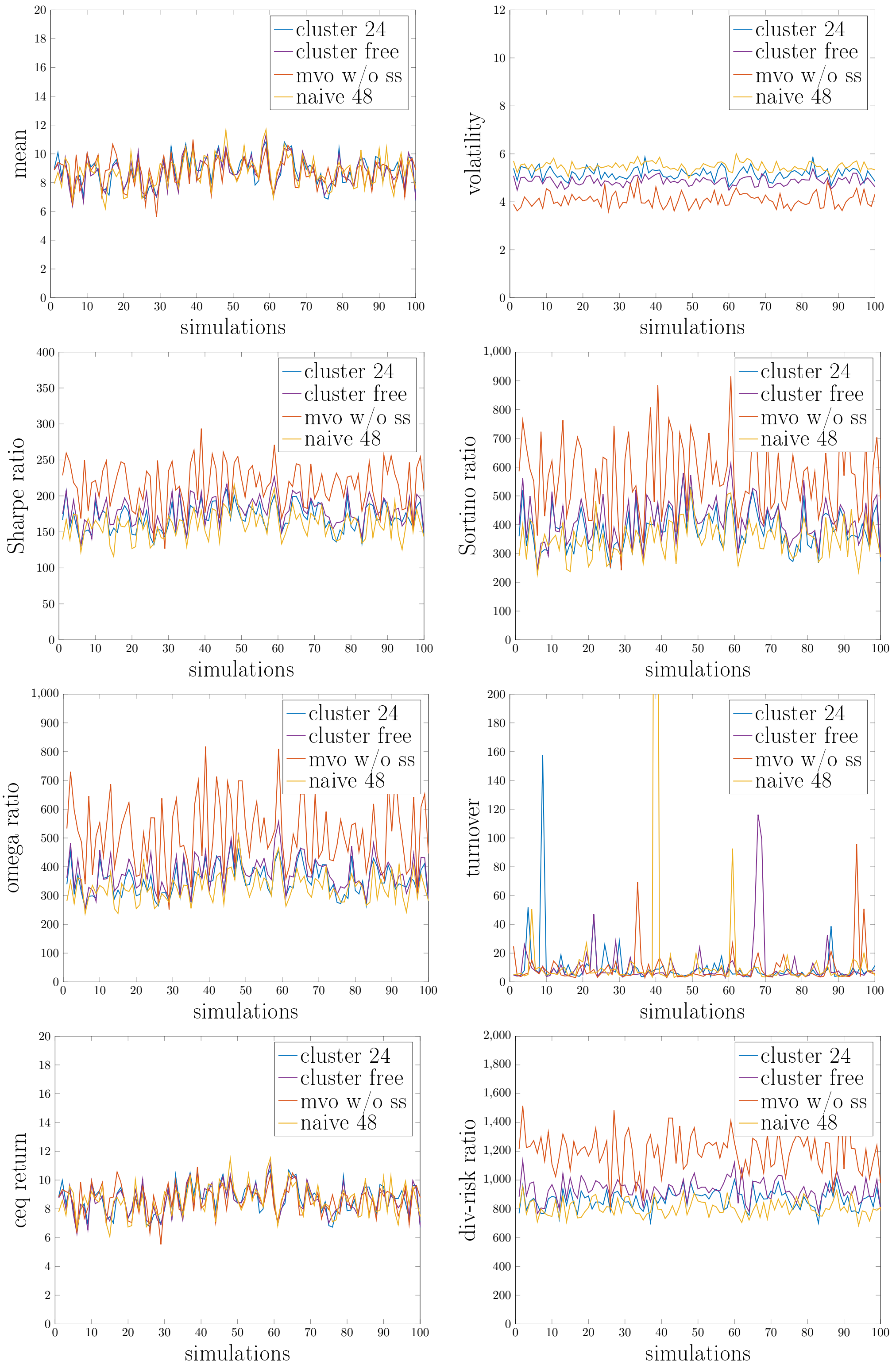


Figure D.15: Model 6.28: Annual performance (in percent, except turnover). Simulation model 6, 48 assets, 28 correlation groups, 180 months estimation window, 360 months total testing period. First 100 simulation sets.

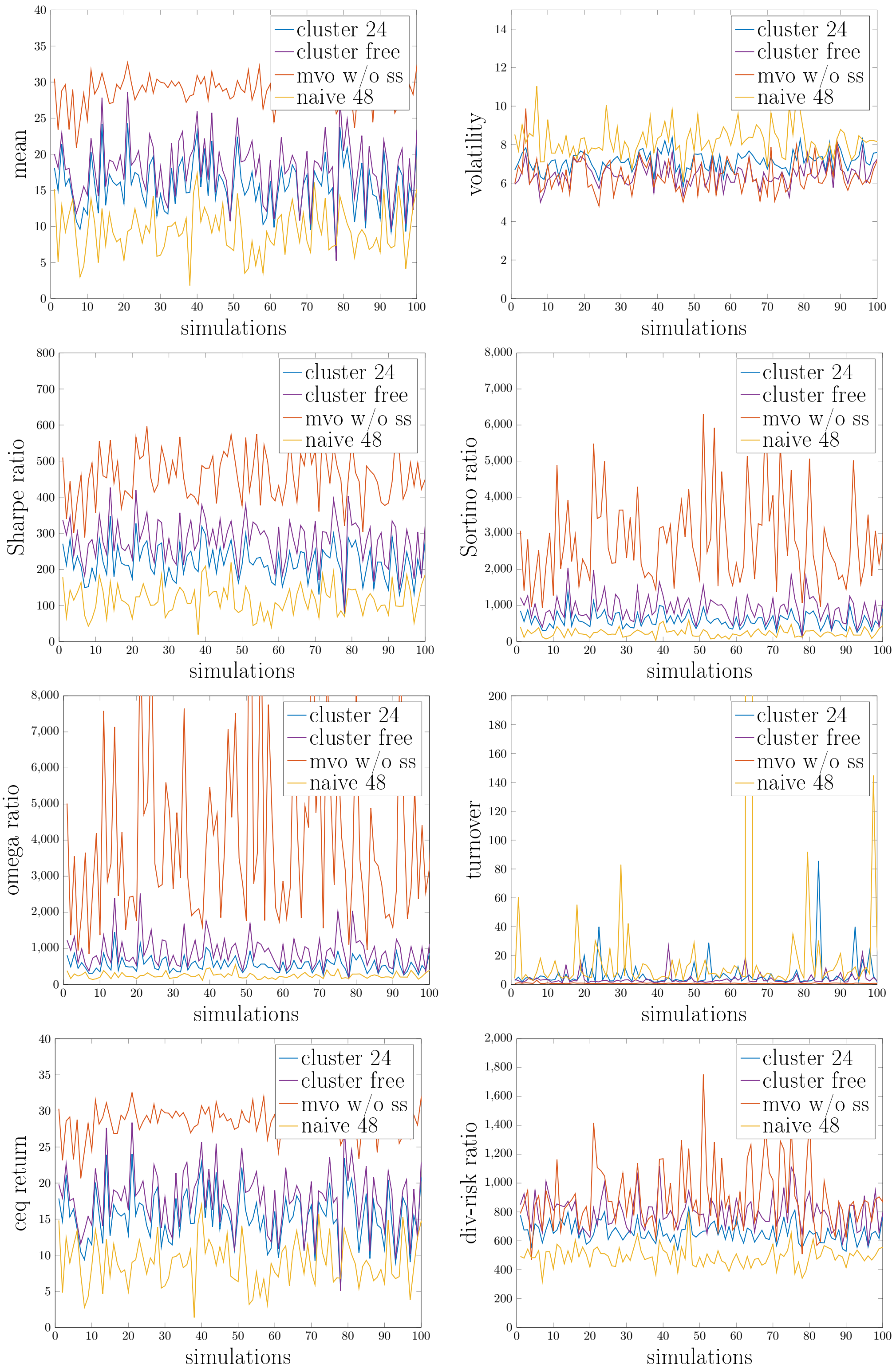


Figure D.16: Model 5.20: Annual performance (in percent, except turnover). Simulation model 5, 48 assets, 20 correlation groups, 180 months estimation window, 360 months total testing period. First 100 simulation sets.

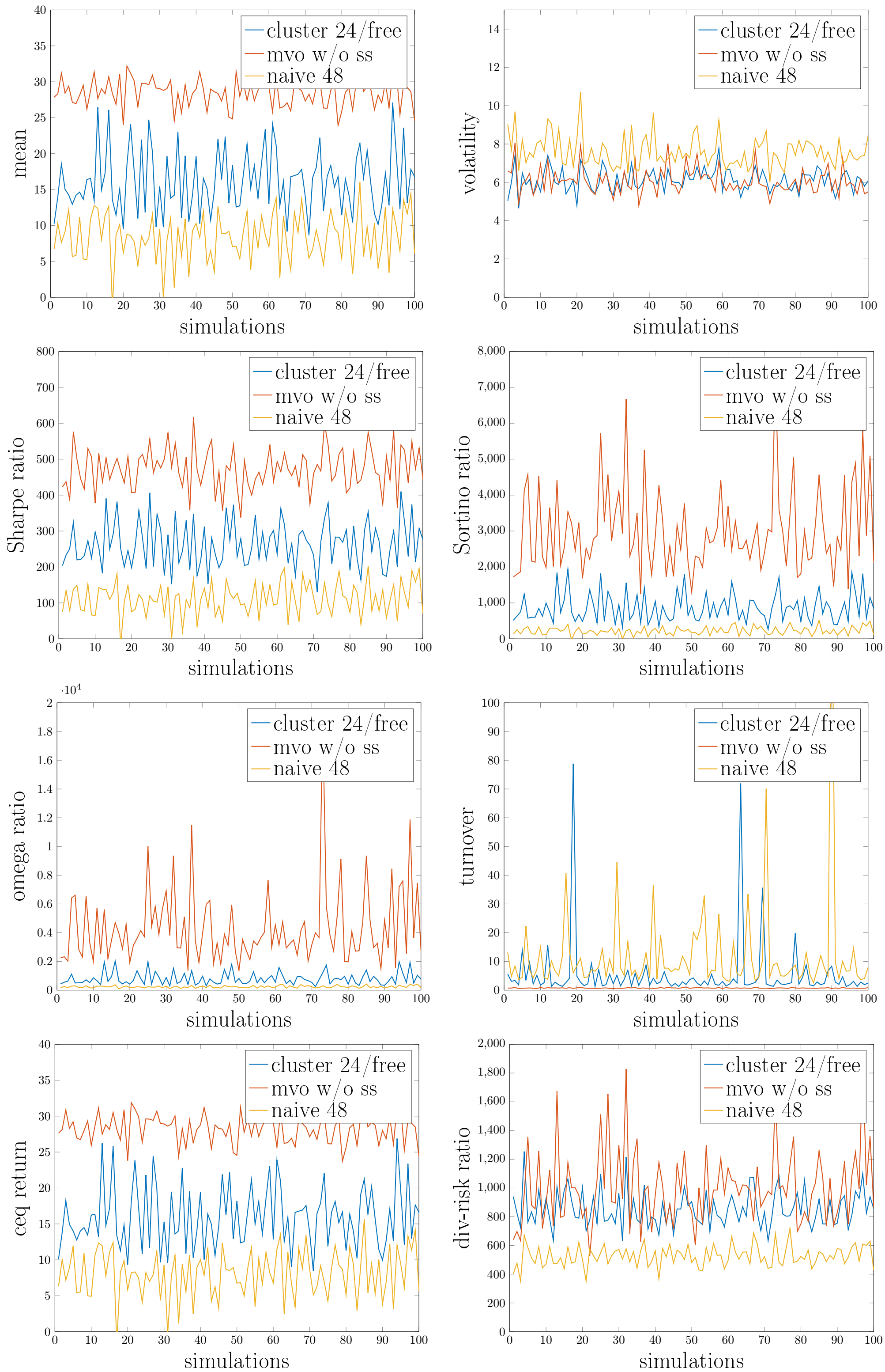


Figure D.17: Model 5.24: Annual performance (in percent, except turnover). Simulation model 5, 48 assets, 24 correlation groups, 180 months estimation window, 360 months total testing period. First 100 simulation sets.

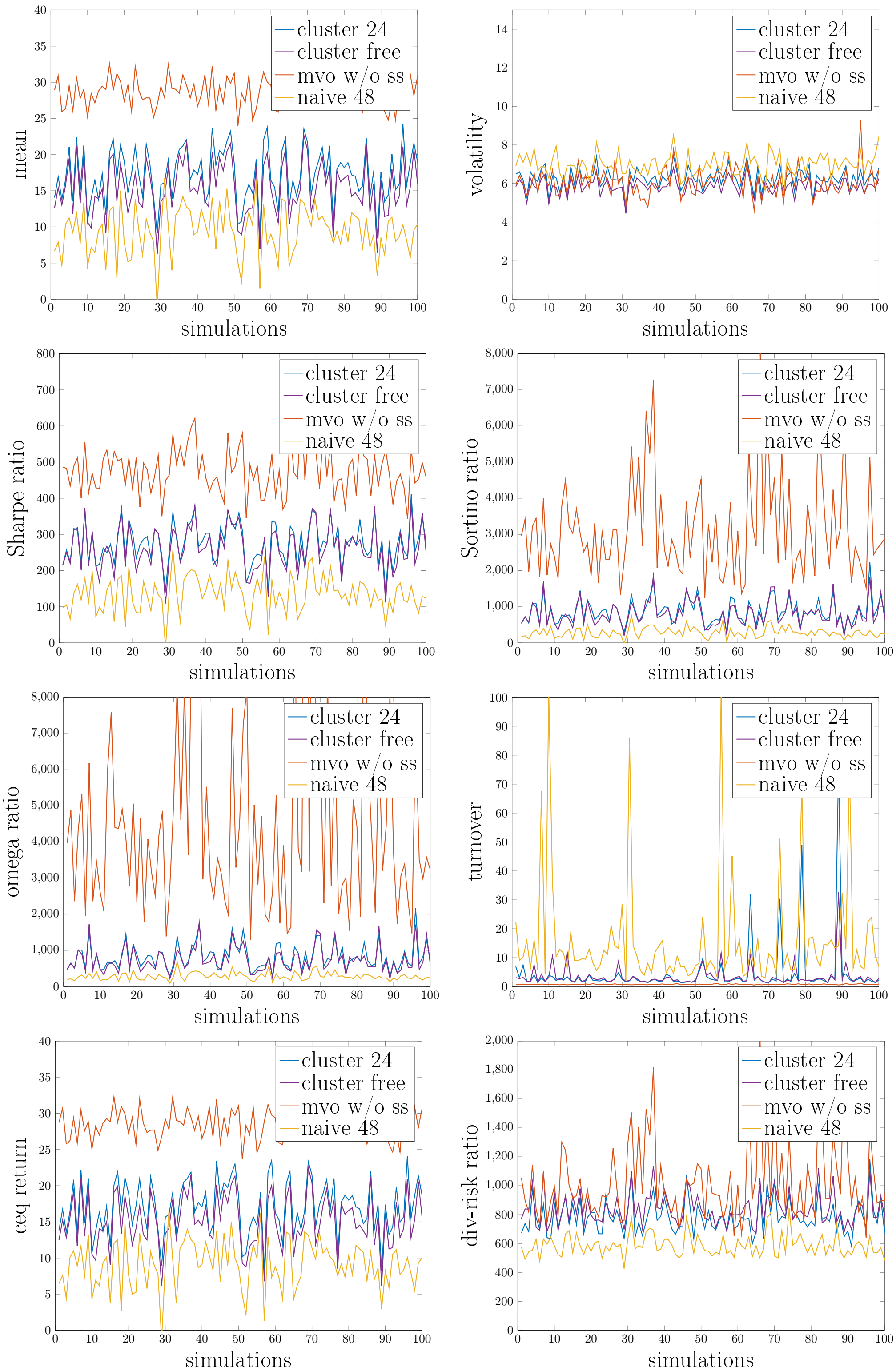


Figure D.18: Model 5.28: Annual performance (in percent, except turnover). Simulation model 5, 48 assets, 28 correlation groups, 180 months estimation window, 360 months total testing period. First 100 simulation sets.

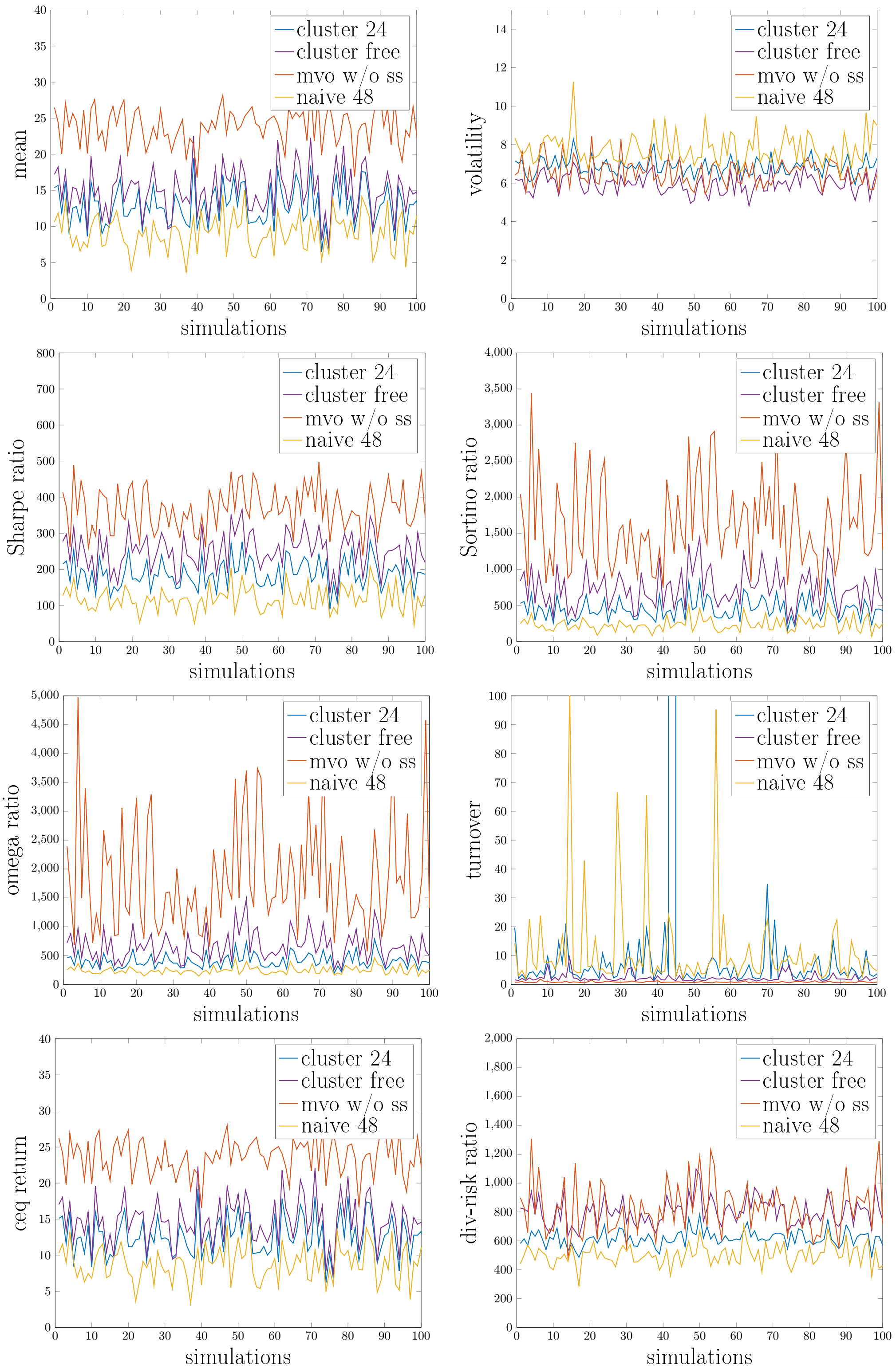


Figure D.19: Model 7.20: Annual performance (in percent, except turnover). Simulation model 7, 48 assets, 20 correlation groups, 180 months estimation window, 360 months total testing period. First 100 simulation sets.

Figure D.20: Model 7.24: Annual performance (in percent, except turnover). Simulation model 7, 48 assets, 24 correlation groups, 180 months estimation window, 360 months total testing period. First 100 simulation sets.

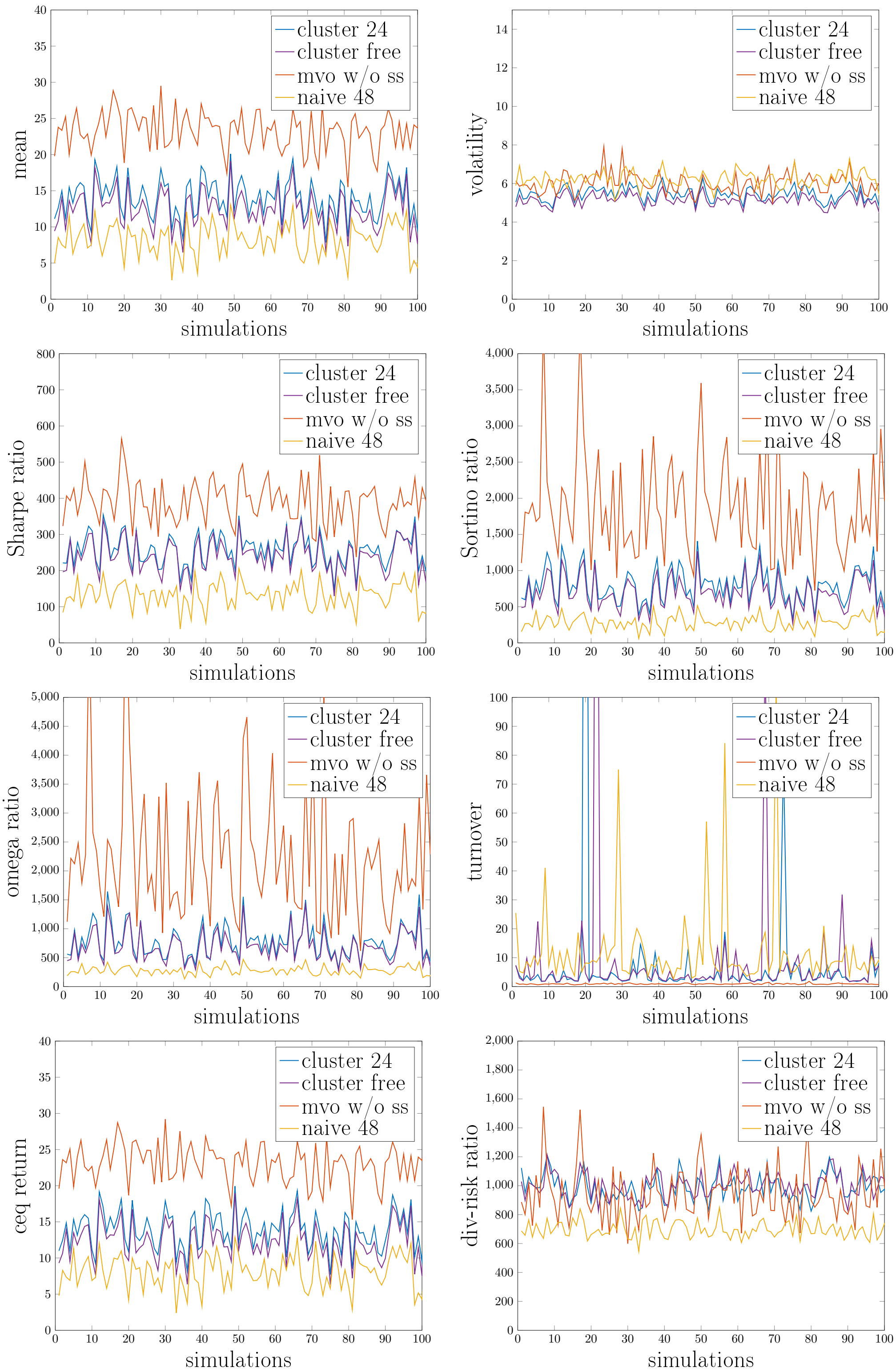


Figure D.21: Model 7.28: Annual performance (in percent, except turnover). Simulation model 7, 48 assets, 28 correlation groups, 180 months estimation window, 360 months total testing period. First 100 simulation sets.

List of Figures

List of Tables

References

Alp, Ö. S. and R. Korn (2011). Continuous-time mean-variance portfolio optimization in a jump-diffusion market. *Decisions in Economics and Finance* **34** (1), 21–40.

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis.* Vol. 2. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York.

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical Finance* **9** (3), 203–228.

Bjerring, T. T., O. Ross, and A. Weissensteiner (2017). Feature selection for portfolio optimization. *Annals of Operations Research* **256** (1), 21–40.

Branger, N., K. Lucivjanska, and A. Weissensteiner (2019). Optimal granularity for portfolio choice. *Journal of Empirical Finance* **50**, 125–146.

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

Carmichael, B., G. B. Koumou, and K. Moran (2015). Unifying portfolio diversification measures using Rao's quadratic entropy. *Cahier de Recherche/Working Paper* **15**, 08.

Cont, R., R. Deguest, and G. Scandolo (2010). Robustness and sensitivity analysis of risk measurement procedures. *Quantitative Finance* **10** (6), 593–606.

Cover, T. M. (1991). Universal portfolios. *Mathematical Finance* **1** (1), 1–29.

Deelstra, G., J. Dhaene, and M. Vanmaele (2011). An overview of comonotonicity and its applications in finance and insurance. In: *Advanced Mathematical Methods for Finance.* Springer, Berlin, Heidelberg, 155–179.

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

Dhaene, J., M. Denuit, M. Goovaerts, R. Kaas, and D. Vyncke (2002). The concept of comonotonicity in actuarial science and finance: Theory. *Insurance: Mathematics and Economics* **31** (1), 3–33.

Dykstra, R. L. (1983). An algorithm for restricted least squares regression. *Journal of the American Statistical Association* **78** (384), 837–842.

Föllmer, H. and A. Schied (2002). Convex measures of risk and trading constraints. *Finance and Stochastics* **6** (4), 429–447.

Glunt, W., T. L. Hayden, S. Hong, and J. Wells (1990). An alternating projection algorithm for computing the nearest Euclidean distance matrix. *SIAM Journal on Matrix Analysis and Applications* **11** (4), 589–600.

Goldstein, M. A., J. McCarthy, and A. G. Orlov (2019). The core, periphery, and beyond: Stock market comovements among EU and non-EU countries. *Financial Review* **54** (1), 5–56.

Haff, L. (1979). An identity for the Wishart distribution with applications. *Journal of Multivariate Analysis* **9** (4), 531–544.

Hastie, T., R. Tibshirani, and J. Friedman (2017). *The Elements of Statistical Learning.* Springer Series in Statistics. New York, NY, USA: Springer New York Inc.

Higham, N. J. (2002). Computing the nearest correlation matrix - a problem from finance. *IMA Journal of Numerical Analysis* **22** (3), 329–343.

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

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

Luenberger, D. G. (1997). *Optimization by Vector Space Methods*. John Wiley & Sons, Inc, New York.

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

Marti, G., S. Andler, F. Nielsen, and P. Donnat (2016). Clustering financial time series: How long is enough? *arXiv Preprint arXiv:1603.04017*.

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

Pflug, G. C., A. Pichler, and D. Wozabal (2012). The 1/N investment strategy is optimal under high model ambiguity. *Journal of Banking and Finance* **36** (2), 410–417.

Rachev, S., S. Ortobelli, S. Stoyanov, F. J. Fabozzi, and A. Biglova (2008). Desirable properties of an ideal risk measure in portfolio theory. *International Journal of Theoretical and Applied Finance* **11** (01), 19–54.

Rao, C. and T. Nayak (1985). Cross entropy, dissimilarity measures, and characterizations of quadratic entropy. *IEEE Transactions on Information Theory* **31** (5), 589–593.

Rokach, L. and O. Maimon (2005). Clustering methods. In: *Data Mining and Knowledge Discovery Handbook*. Springer, Boston, MA, 321–352.

Rollinger, T. N. and S. T. Hoffman (2013). Sortino: A 'Sharper' ratio. *Chicago, IL: Red Rock Capital*.

Sharpe, W. F. (1994). The Sharpe ratio. *Journal of Portfolio Management* **21** (1), 49–58.

Timmermann, A. (2000). Moments of Markov switching models. *Journal of Econometrics* **96** (1), 75–111.

Tu, J. and G. Zhou (2011). Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies. *Journal of Financial Economics* **99** (1), 204–215.

Scientific and Professional Career

Education

06/2000	Abitur, Leininger Gymnasium Grünstadt
10/2000 - 09/2003	Grundstudium, Ruprecht-Karls-Universität Heidelberg, degree program Mathematics with minor subject Computer Science
04/2003	Vordiplom
10/2003 - 09/2006	parental leave
10/2006 - 03/2012	Hauptstudium, Technische Universität Kaiserslautern, degree program Business Mathematics, specialization in Financial Mathematics
03/2012	Diploma Thesis, 'Convergence of Binomial Models for Option Pricing'
since 10/2012	PhD student of Prof. Dr. Jörn Saß

Teaching Experience: Tutorials and Seminars

10/2001 - 03/2002	Analysis 1
10/2012 - 03/2013	Non-Life Insurance
04/2013 - 09/2013	Insurance Seminar
10/2013 - 03/2014	Mathematik für Wirtschaftswissenschaftler
04/2014 - 09/2014	Mathematik für Bauingenieure
10/2014 - 03/2015	Praktische Mathematik: Stochastik
04/2015 - 09/2015	Biostatistik für Chemiker
10/2015 - 03/2016	Mathematik für Wirtschaftswissenschaftler
04/2016 - 09/2016	Mathematik für Bauingenieure
10/2016 - 03/2017	Mathematik für Wirtschaftswissenschaftler
04/2017 - 09/2017	Foundation of Financial Mathematics
10/2017 - 03/2018	Project Seminar "Advanced Modelling in Actuarial and Financial Mathematics"

Others

04/2018 - 03/2020	academic staff member for the distance-learning master degree program Financial Engineering

Wissenschaftlicher und Beruflicher Werdegang

Ausbildung

06/2000	Abitur, Leininger Gymnasium Grünstadt
10/2000 - 09/2003	Grundstudium, Ruprecht-Karls-Universität Heidelberg, Diplomstudiengang Mathematik mit Nebenfach Informatik
04/2003	Vordiplom
10/2003 - 09/2006	Elternzeit
10/2006 - 03/2012	Hauptstudium, Technische Universität Kaiserslautern, Diplomstudiengang Wirtschaftsmathematik mit Vertiefungsrichtung Finanzmathematik
03/2012	Diplomarbeit, 'Convergence of Binomial Models for Option Pricing'
seit 10/2012	Doktorandin von Prof. Dr. Jörn Saß

Lehre: Betreuung von Übungen und Seminaren

10/2001 - 03/2002	Analysis 1
10/2012 - 03/2013	Non-Life Insurance
04/2013 - 09/2013	Insurance Seminar
10/2013 - 03/2014	Mathematik für Wirtschaftswissenschaftler
04/2014 - 09/2014	Mathematik für Bauingenieure
10/2014 - 03/2015	Praktische Mathematik: Stochastik
04/2015 - 09/2015	Biostatistik für Chemiker
10/2015 - 03/2016	Mathematik für Wirtschaftswissenschaftler
04/2016 - 09/2016	Mathematik für Bauingenieure
10/2016 - 03/2017	Mathematik für Wirtschaftswissenschaftler
04/2017 - 09/2017	Foundation of Financial Mathematics
10/2017 - 03/2018	Project Seminar "Advanced Modelling in Actuarial and Financial Mathematics"

Weiteres

04/2018 - 03/2020	Wissenschaftliche Mitarbeiterin im Master-Fernstudiengang Financial Engineering