Can Optimized Portfolios Beat 1/N?

This dissertation is presented in part fulfillment
of the requirement for the completion of an
MSc in Economics
in the Department of Economics, University of Konstanz,
and an
MSc in Economics and Econometrics
in the School of Economics, University of Nottingham.
The work is the sole responsibility of the candidate.

By: Valerius Disch

Period of Completion: April 21, 2018 - August 21, 2018

Word Count: 14,777 / 15,000

1st Assessor: Professor Marcel Fischer, University of Konstanz
2nd Assessor: Professor Patrick Marsh, University of Nottingham

Konstanz, August 21, 2018

Abstract

This analysis conducts an out-of-sample horse race between the naive diversification approach and 15 optimized portfolio strategies across four US and two European equity data sets. Not a single optimized strategy achieves to consistently outperform the naive diversification benchmark in a statistically significant manner based on a total of six performance evaluation criteria. However, unconstrained strategies related to the sample-based mean-variance and minimum-variance strategy exhibit superior performances in two US data sets. Timing strategies attain consistently good results for US industry data sets. Performance behavior appears to be sensitive to respective data sets, yet independent of performance evaluation criteria.

Contents

List of Figures

List of Tables

List of Abbreviations

AMEX	=	American Stock Exchange
AR(p)	=	Autoregressive Model of Order p
BS	=	Bayes-Stein
BSsc	=	Bayes-Stein Short Sale Constraint
CAPM	=	Capital Asset Pricing Model
CE	=	Certainty Equivalent Return
CR	=	Calmar Ratio
DD	=	Drawdown
et al.	=	et alia, “and others”
EU	=	Europe
EW	=	Equally-Weighted
EW/MinV	=	Equally-Weighted and Minimum-Variance
EW/MV	=	Equally-Weighted and Mean-Variance
EW/MV/MinV	=	Equally-Weighted and Mean-Variance and Minimum-Variance
HAC	=	Heteroskedasticity and Autocorrelation
i.e.	=	id est, “that is to say”
IS	=	In-Sample
JK	=	Jobson and Korkie
LPM	=	Lower Partial Moments
LW	=	Ledoit and Wolf
MAR	=	Minimum Acceptable Return
MaxDD	=	Maximum Drawdown
MinV	=	Minimum-Variance
MinVsc	=	Minimum-Variance Short Sale Constraint
MKT	=	Excess Return on the Market
ML	=	Maximum Likelihood
MP	=	MacKinlay-Pástor
MSCI	=	Morgan Stanley Capital International
MV	=	Mean-Variance
MVlsc	=	Mean-Variance Long-Short Sale Constraint
MV/MinV	=	Mean-Variance and Minimum-Variance
MVsc	=	Mean-Variance Short Sale Constraint
NASDAQ	=	National Association of Securities Dealers Automated Quotations
NYSE	=	New York Stock Exchange
OR	=	Omega Ratio
RoVaR	=	Return on Value-at-Risk
RRT	=	Reward-to-Risk Timing
SoR	=	Sortino Ratio
SR	=	Sharpe Ratio
US	=	United States
USA	=	United States of America
VaR	=	Value-at-Risk
VT	=	Volatility Timing
VW	=	Value-Weighted

1. Introduction

In 1952, Harry Markowitz published his work Portfolio Selection and arguably set the foundation of what is nowadays known as Modern Portfolio Theory. Regarding a set of risky assets, Markowitz (1952) rationalizes an investment behavior that favors expected return and dismisses risk measured in terms of variance of return. Since returns typically tend to increase with risk, this constitutes a fundamental trade-off which, however, can be addressed by a rational and risk-averse investor via fixing a level of variance of return while maximizing expected return. Although theoretically appealing, there is no guarantee that following this proposed investment behavior indeed results in desired outcomes when tested in a real data environment. A natural choice of a desired outcome is the outperformance of an optimized portfolio strategy relative to the naive diversification or simply $1/N$ approach. The latter describes an investment schedule which invests equal shares of wealth in all available risky assets of which the number is assumed to be N.

In this spirit, this analysis conducts an out-of-sample horse race between the naive diversification approach and 15 optimized portfolio strategies based on the mean-variance framework. By doing so, several contributions to the existing related portfolio optimization literature are made. First, similar studies such as DeMiguel et al. (2009b) are updated by respecting well-established mean-variance model extensions as well as rather newly proposed approaches resulting in a comprehensive selection of optimized portfolio strategies. Second, the data sets are not exclusively restricted to the United States (US) but also account for European equity data and hence allow for a comparison between US and European specific results. Third, the scope of performance evaluation criteria is extended and contains an additional number of four reward-to-risk ratios alongside the standard Sharpe ratio and certainty equivalent return. Besides, the test statistic introduced by Ledoit and Wolf (2008) to decide on the pairwise difference between the Sharpe ratios of two different strategies is adjusted so as to be also valid in case of certainty equivalent returns. This facilitates the identification of statistically significant results and explicitly accounts for nonnormally distributed return data. The results of this analysis are beneficial not only for private investors but also for professional practitioners who strive for a more favorable return-to-risk trade-off and have an interest in answering the question of whether optimized portfolios can beat $1/N$.

The main finding of this analysis can be summarized in the fact that not a single optimized portfolio strategy achieves to consistently outperform the naive diversification benchmark in a statistically significant manner. However, performances across different data sets are very heterogeneous. In case of two US equity data sets formed on bivariate sorts of market equity and the book-to-market equity ratio, unconstrained strategies related to the sample-based mean-variance and minimum-variance strategy exhibit superior performances. In contrast, timing strategies only attain consistently good results in the context of US equity data classified according to US industries. The European data sets are generally short of any outperforming optimized strategy. Eventually, the performance behavior appears to be very sensitive to respective data sets. The relative performance rankings of the strategies under consideration, however, are independent of the performance evaluation criteria, i.e. the Sharpe ratio is in general an adequate choice although relying on the assumption of normally distributed return data.

This analysis is organized as follows. Section 2 reviews relevant literature and presents important empirical results. Section 3 describes the data. Section 4 introduces the optimized portfolio strategies and Section 5 lists the performance evaluation criteria. Results are presented and discussed in Section 6. Section 7 contains robustness checks and Section 8 concludes.

2. Literature Review

The past few years of portfolio optimization research can be characterized by a series of publications conducting horse races between different competing optimized portfolio strategies. In this context, the work by DeMiguel et al. (2009b) is of particular interest and also serves as the basis on which this analysis is build.

DeMiguel et al. (2009b) compare the performance of 12 different optimized portfolio strategies relative to the naive diversification benchmark with respect to the out-of-sample Sharpe ratio, certainty equivalent return, and turnover. The underlying data are composed of monthly excess stock returns for mainly US equities and cover the period from July 1963 to November 2004. By applying a rolling-window estimation approach, DeMiguel et al. (2009b) come to the somewhat surprising conclusion that none of the strategies under consideration achieves to consistently outperform the naive diversification approach. This result is valid throughout all performance criteria. DeMiguel et al. (2009b) attribute the bad performance of the optimized strategies to the characteristic occurrence of estimation errors when estimating expected returns and variance-covariance matrices in the process of building these optimal strategies.

Tu and Zhou (2011) subsequently introduce four novel optimized portfolio strategies that are optimized combinations of already existing approaches. In particular, the naive diversification approach is combined with the sample-based mean-variance strategy as well as with the strategies suggested by Jorion (1986), MacKinlay and Pástor (2000), and Kan and Zhou (2007), respectively. The general setting and estimation methodology is very similar to DeMiguel et al. (2009b) which ensures a certain degree of comparability. Tu and Zhou (2011) show evidence of an outperformance of the optimally combined portfolio strategies relative to their uncombined counterparts in terms of Sharpe ratios and certainty equivalent returns. Some of the newly proposed optimized portfolio strategies even achieve to outperform naive diversification. This is especially true in case of large sample sizes. These findings are promising and rationalize the use of a subset of their proposed models also in this analysis.

Pflug et al. (2012) focus on the two levels of uncertainty prevailing in the context of portfolio optimization. First, the actual realizations of the asset returns are uncertain and second, and even more critical, the data generating process itself is not known exactly. The higher the degree of this model uncertainty the more the theoretically optimal portfolio strategy resembles the naive diversification approach. On the basis of different risk measures, Pflug et al. (2012) subsequently demonstrate that naive diversification is even optimal in case of a risk-averse investor who cares about both levels of uncertainty in combination with a very limited or even nonexistent knowledge of the data generating process. As a result, the empirical success of the naive diversification approach cannot only be attributed to general estimation inaccuracies when estimating moments of an assumed data generating distribution but also to wrong inferences about the data generating process in the first place. Following this argument, focus should be either put on the improvement of the estimation precision of already existing models or the introduction of new optimized portfolio strategies including a justified modeling of the data generating process.

An interesting remark is made by Kirby and Ostdiek (2012) who argue that the design of the comparisons between the different portfolio strategies typically favors naive diversification. The general assumption of a representative mean-variance investor leads to optimized portfolio strategies that strive for high expected returns. This, however, can result in extreme portfolio weight positions that take advantage of even minor asset return differences. In case of an imprecise estimation of these return rates, the resulting portfolio can exhibit poor out-of-sample performance. Hence, Kirby and Ostdiek (2012) propose the idea of imposing a target return on the optimized portfolio strategies equal to the out-of-sample return obtained by the naive diversification approach. This indeed improves the performance of the optimized portfolio strategies relative to naive diversification and dampens the portfolio weights to less extreme positions. In addition, Kirby and Ostdiek (2012) introduce a volatility and reward-to-risk timing strategy which achieve superior results relative to the naive diversification benchmark for US equity data covering the period from July 1963 to December 2008. The performance of these two newly introduced strategies is encouraging as it suggests that optimized portfolio strategies can actually outperform naive diversification irrespective of the reliance on sufficiently long sample sizes.

An extension of the work by DeMiguel et al. (2009b) to a broader set of international asset classes including stocks, bonds, and commodities is given by Jacobs et al. (2014). For the time period from February 1973 to December 2008, a total of eleven mean-variance based optimized portfolio strategies are compared to the naive diversification approach. As a first result, Jacobs et al. (2014) confirm previous findings that hardly any optimized portfolio strategy consistently outperforms naive diversification with respect to equity data. Second, these results transmit to the other asset classes under consideration for which naive diversification also seems to be the most successful approach in achieving high out-of-sample Sharpe ratios.

In the same spirit, Bessler et al. (2017) examine the period from January 1993 to December 2011 with respect to an international portfolio consisting of stocks, bonds, and commodity indices and compare the naive diversification approach to the optimized sample-based mean-variance, minimum-variance, and Bayes-Stein strategy. In accordance with DeMiguel et al. (2009b), the Sharpe ratios belonging to the optimized strategies do not significantly outperform naive diversification. Additionally, the Omega ratio as well as the maximum drawdown are included as further reward-to-risk measures and according to these criteria the obtained results are suggestive of a superior performance regarding the optimized portfolio strategies.

This analysis aims at continuing this series of portfolio optimization research by conducting a horse between optimized portfolio strategies that incorporates the main insights of previous empirical work with respect to both the selection of optimized strategies as well as performance evaluation criteria.

3. Data

This analysis is based on a total of six data sets containing monthly excess returns over a risk-free rate defined as the one-month US Treasury bill rate. There are four samples based on US data and two samples referring to European data. The returns of the US data sets are a combination of New York Stock Exchange (NYSE), American Stock Exchange (AMEX), and National Association of Securities Dealers Automated Quotations (NASDAQ) stocks and are value-weighted portfolios either formed based on bivariate sorts on market equity and the book-to-market equity ratio or ordered by industry affiliation. The US data cover the time period from January 1970 to December 2017. The choice of the time period is determined by data availability. The US industry data are only complete as of the year 1970. The return series for the other US samples originally include more observations but are adjusted for the same time period as the US industry data to ensure comparability. Independent of the actual data availability, all data sets are balanced such that the time period always starts in January and ends in December. This is done to guarantee an equal distribution of months to mitigate calendar effects such as the well-known January effect. As a result, the US portfolios each contain a total of 576 months of excess return data representing 48 years of excess returns, respectively. The returns of the European data sets are based on data provided by Morgan Stanley Capital International (MSCI) and Bloomberg and are value-weighted portfolios formed based on bivariate sorts on market equity and the book-to-market equity ratio. The time period ranges from January 1991 to December 2017 and thus includes 324 observations altogether corresponding to 27 years of monthly excess return data. All data belong to the single-asset class of stocks and are drawn from Kenneth R. French's data library[1]. The

[1] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Table 3.1: Overview Data Sets. This table lists a total of six data sets that are included in this analysis all of which contain monthly excess returns over a risk-free rate defined as the one-month US Treasury bill rate. The first four samples refer to US data while the remaining two samples are based on European data. The abbreviations are introduced to refer to the specific data sets in the text. The variable N indicates the total number of available risky assets included in each data set. The time period indicates the duration of the data collection and the last column of the table lists the underlying number of observations included in each data set. The data are obtained from Kenneth R. French's website `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`. A more detailed description of the data and the process of creating the different portfolios can be found in Appendix A.

#	Data Set	Abbreviation	N	Time Period	Obs.
	United States of America (USA)				
1	Six US portfolios formed on market equity and the book-to-market equity ratio	PF6 USA	6	01/1970 - 12/2017	576
2	Twenty-five US portfolios formed on market equity and the book-to-market equity ratio	PF25 USA	25	01/1970 - 12/2017	576
3	Ten US industry portfolios	IND10 USA	10	01/1970 - 12/2017	576
4	Forty-eight US industry portfolios	IND48 USA	48	01/1970 - 12/2017	576
	Europe (EU)				
5	Six European portfolios formed on market equity and the book-to-market equity ratio	PF6 EU	6	01/1991 - 12/2017	324
6	Twenty-five European portfolios formed on market equity and the book-to-market equity ratio	PF25 EU	25	01/1991 - 12/2017	324

data sets employed in this analysis are very similar to the samples typically used in the related empirical portfolio optimization literature and therefore a high degree of comparability with respect to the obtained results is guaranteed. Table 3.1 presents an overview of the data considered in this analysis.

The remainder of this section is dedicated to the description of PF6 USA and PF6 EU as well as their respective value-weighted excess market return (MKT) series to obtain a deeper understanding of the inherent characteristics of the data sets underlying this analysis. In case of the PF6 portfolios, the total stock data are divided into a small and a big group with respect to a specific level of market equity of which each group is again subdivided into either a low, medium, or high group with respect to a specific level of book-to-market equity ratio. A more detailed description of the data and the process of creating the different portfolios can be found in Appendix A. Table 3.2 exhibits descriptive statistics for the US and European MKT series and PF6 portfolios, respectively. Several interesting facts appear when examining Table 3.2. Focusing on the MKT series of the USA and Europe, a first observation refers to the mean of the US MKT series which is substantially higher relative to that of Europe. In addition, the European market is characterized by a higher standard deviation and more extreme values as indicated by the corresponding minimum and maximum values. These facts summarize in a European MKT Sharpe ratio of approximately 0.12 which is lower than the respective value of 0.17 for the USA. Based on this reward-to-risk ratio, the descriptive statistics are suggestive of an outperformance of the European market by the US market. A second observation is

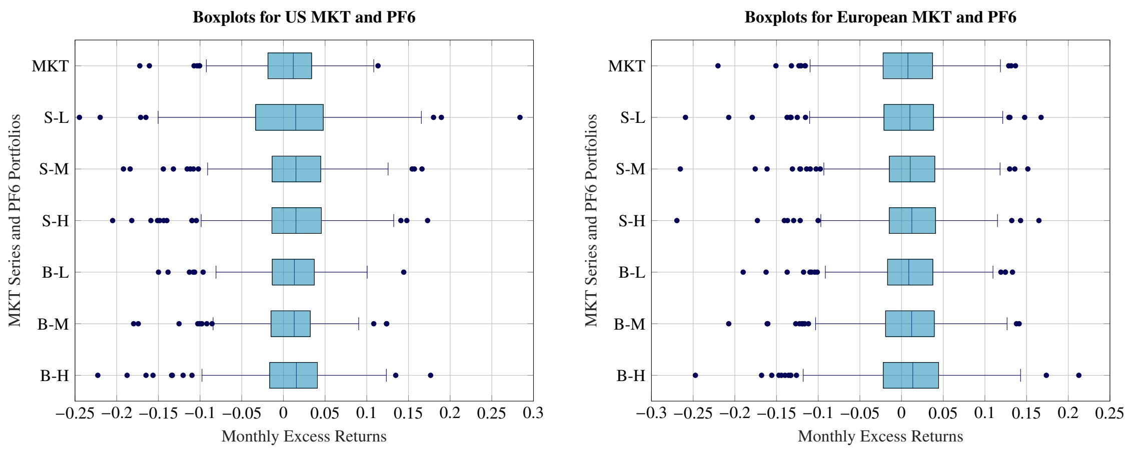
Table 3.2: Descriptive Statistics for the US and European MKT Series and PF6 Portfolios. This table lists the MKT series and the six portfolios for the US and Europe, i.e. PF6 USA and PF6 EU, respectively. The six portfolios are constructed by means of bivariate sorts of the total stock data on market equity with two groups, small and big, and on book-to-market equity ratio with three groups, low, medium, and high. The US portfolios are truncated so as to match the time period of the European portfolio. Hence, the data span from January 1991 to December 2017 containing 324 observations of monthly excess returns representing 27 years. This table displays the mean, standard deviation, sample-bias corrected skewness, sample-bias corrected kurtosis, minimum value, maximum value, and Sharpe ratio of each return series.

USA		**Small**			**Big**		
Statistics	MKT	Low	Medium	High	Low	Medium	High
Mean	0.0072	0.0090	0.0125	0.0135	0.0097	0.0094	0.0102
Standard Deviation	0.0417	0.0657	0.0499	0.0531	0.0417	0.0410	0.0499
Skewness	−0.7082	−0.2178	−0.4815	−0.6509	−0.4712	−0.7527	−0.8022
Kurtosis	4.5338	4.5721	4.6741	4.8114	4.0160	5.6602	5.7157
Minimum	−0.1723	−0.2447	−0.1919	−0.2049	−0.1499	−0.1797	−0.2227
Maximum	0.1135	0.2838	0.1663	0.1729	0.1444	0.1239	0.1766
Sharpe Ratio	0.1727	0.1370	0.2505	0.2542	0.2326	0.2293	0.2044
Europe		**Small**			**Big**		
Statistics	MKT	Low	Medium	High	Low	Medium	High
Mean	0.0057	0.0056	0.0082	0.0104	0.0065	0.0086	0.0086
Standard Deviation	0.0483	0.0530	0.0487	0.0501	0.0457	0.0495	0.0572
Skewness	−0.5951	−0.8329	−0.9231	−0.7339	−0.5482	−0.6478	−0.4904
Kurtosis	4.8052	5.9839	6.8226	6.4994	4.6034	4.7857	4.8253
Minimum	−0.2202	−0.2592	−0.2654	−0.2695	−0.1900	−0.2073	−0.2473
Maximum	0.1367	0.1674	0.1515	0.1648	0.1331	0.1411	0.2127
Sharpe Ratio	0.1180	0.1057	0.1684	0.2076	0.1422	0.1737	0.1503

that throughout all portfolios the minimum and maximum values are relatively extreme in relation to the corresponding means and standard deviations. This behavior is indicative of outliers. Also, all portfolios show a negative skewness as well as a level of kurtosis that exceeds the reference level of a standard normal distribution of three. These observations cast a first doubt on the general assumption prevailing in the empirical literature of having normally distributed asset returns.

A graphical representation of some distributional characteristics of the US and European MKT series and PF6 portfolios can be obtained by boxplots as shown in Fig. 3.1. The lengths of the boxes for the European sample appear to be more homogenous relative to the US sample suggesting that the European portfolios share more similar distributional characteristics. Again, the number of outliers in all boxplots speak in favor of return data which are not normally distributed.

Table 3.3 yields correlations of the US MKT series and PF6 portfolios. All correlations are highly positive and even the lowest correlation coefficient between the US B-M and S-L portfolios still amounts to 0.62. In addition, the correlations between the US and European portfolios are highly positive ranging from 0.62 for the S-H portfolio to 0.79 for the MKT series. Facing an investment universe that only consists of highly correlated assets limits the potential benefits of diversification and might result in lower levels of



(a) Boxplots for the US MKT series and PF6 portfolios. The US portfolios are truncated so as to match the time period of the European portfolios. Hence, the data spans from January 1991 to December 2017 containing 324 observations of monthly excess returns representing 27 years.

(b) Boxplots for the European MKT series and PF6 portfolios. The data spans from January 1991 to December 2017 containing 324 observations of monthly excess returns representing 27 years.

Figure 3.1: Boxplots for the US and European MKT Series and PF6 Portfolios. This figure shows boxplots for the US and European MKT series and PF6 portfolios, respectively. The vertical line in each box denotes the median of the distribution. The left and right bounds of the box indicate the 25th and 75th percentiles, respectively. The whiskers cover the range of the observations that are still within the 1.5 interquartile range of the lower and upper quartile. Outliers are denoted by a dot.

reward-to-risk ratios due to higher levels of risk.

Fig. 3.2 displays the monthly cumulative MKT series for the USA and Europe. The US sample starts in January 1970, whereas the European sample only starts in January 1991 due to limited data availability. Both series are upward sloping and show a clear positive trend. Moreover, both return patterns resemble each other, however, the US cumulative returns are higher relative to the European returns for most of the underlying time period. Several major drops in the observed return patterns can be attributed to corresponding economic crises such as the dot-com bubble in the beginning of the 21st century as well

Table 3.3: Correlation Table for the US and European MKT Series and PF6 Portfolios. This table lists correlations for the US and European MKT series and PF6 portfolios, respectively. The US portfolios are truncated so as to match the time period of the European portfolios. Hence, the data span from January 1991 to December 2017 containing 324 observations of monthly excess returns representing 27 years. This table contains several different correlations. In fact, columns 2 (MKT) to 7 (B-H) span two correlation matrices at the same time. The lower left triangular matrix containing plain numbers represents the correlation matrix of the US sample. The upper right triangular matrix containing italic numbers represents the correlation matrix of the European sample. The last column USA/EU yields the correlations of the portfolios between the USA and Europe.

Factors	MKT	S-L	S-M	S-H	B-L	B-M	B-H	USA / EU
MKT	**1.00**	*0.89*	*0.92*	*0.89*	*0.95*	*0.99*	*0.96*	0.79
S-L	0.83	**1.00**	*0.95*	*0.88*	*0.85*	*0.85*	*0.81*	0.66
S-M	0.85	0.91	**1.00**	*0.96*	*0.84*	*0.89*	*0.87*	0.66
S-H	0.81	0.84	0.96	**1.00**	*0.78*	*0.87*	*0.89*	0.62
B-L	0.97	0.78	0.75	0.69	**1.00**	*0.93*	*0.85*	0.77
B-M	0.89	0.62	0.77	0.77	0.82	**1.00**	*0.94*	0.78
B-H	0.85	0.63	0.79	0.82	0.76	0.89	**1.00**	0.72

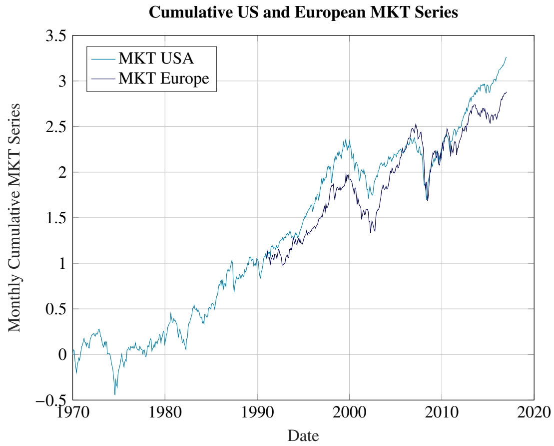
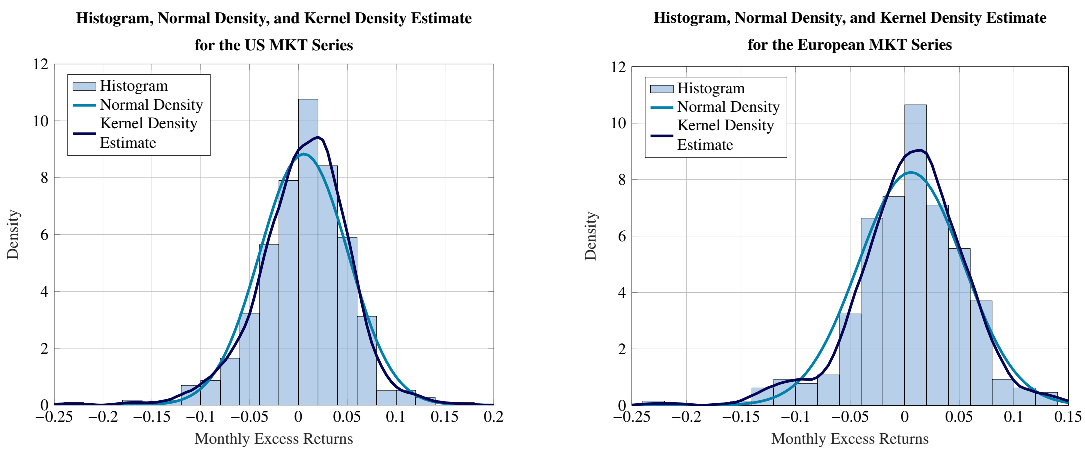


Figure 3.2: Monthly Cumulative MKT Series for the US and Europe. This figure shows the monthly cumulative MKT series for the US and Europe. The US sample covers the period from January 1970 to December 2017 containing 576 observations of monthly excess returns representing 48 years. The European sample covers the period from January 1991 to December 2017 containing 324 observations of monthly excess returns representing 27 years.

as the global financial crisis in 2007 and 2008.

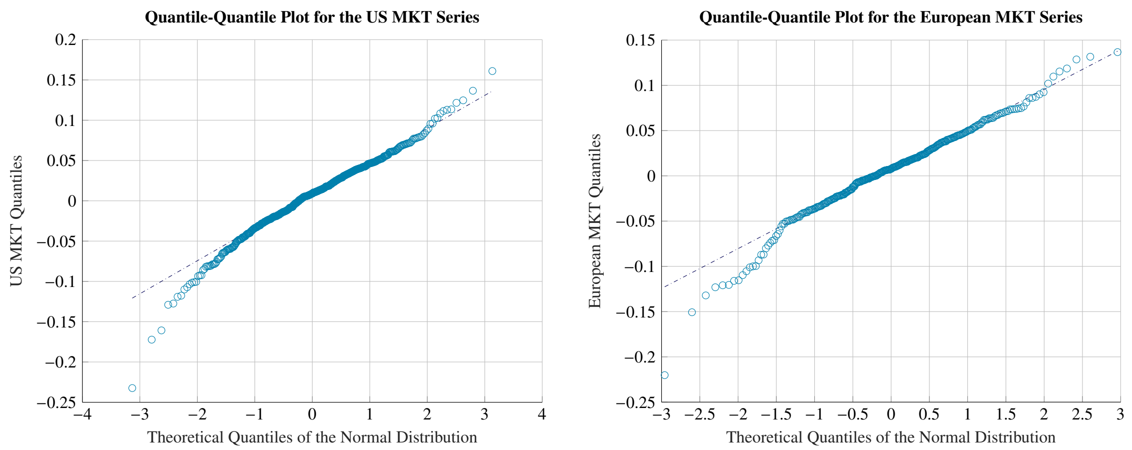
Since the actual distribution of the underlying return data critically influences the choice of appropriate statistical procedures, several different tests are performed. Tests for normality with respect to a specific sample distribution can rely either on a graphical approach or more formally utilizing actual statistical tests. Fig. 3.3 depicts the histogram, a fitted normal distribution, and a kernel density estimate for the US and the European MKT series. A visual inspection of the fitted normal distribution relative to the kernel density estimate does not yield a huge discrepancy neither for the US nor for the European sample. Only the European return distribution exhibits slightly fat tails, especially on the left-hand side of the kernel density estimate. These fat tails are characteristic of re-



(a) Histogram, normal density, and kernel density estimate for the US MKT series based on a total of 576 monthly excess return observations from January 1970 to December 2017. The kernel function is the Epanechnikov with a bandwidth of 0.0124.

(b) Histogram, normal density, and kernel density estimate for the European MKT series based on a total of 324 monthly excess return observations from January 1991 to December 2017. The kernel function is the Epanechnikov with a bandwidth of 0.0147.

Figure 3.3: Histogram, Normal Density, and Kernel Density Estimate for the US and European MKT Series This figure shows the histogram, normal density, and kernel density estimate for the US and European MKT series. The kernel function is the Epanechnikov with unbounded support.



(a) Quantile-quantile plot for the US MKT series covering the period from January 1970 to December 2017 containing 576 observations of monthly excess returns representing 48 years.

(b) Quantile-quantile plot for the European MKT series covering the period from January 1991 to December 2017 containing 324 observations of monthly excess returns representing 27 years.

Figure 3.4: Quantile-Quantile Plot for the US and European MKT Series. This figure shows the quantile-quantile plots for the US and European MKT series. The straight line represents perfectly normally distributed data and serves as a benchmark for comparison with the actual MKT series.

turn series and the reason why return data is often better described by a leptokurtic rather than a normal distribution.

Quantile-quantile plots such as given in Fig. 3.4 allow for a comparison of two distributions. To be precise, the actual return data of the US and European MKT samples are compared to a corresponding normally distributed sample represented by the straight line, respectively. Clearly, both ends of both return series do not fall along the straight line, hence indicating that the MKT series are not normally distributed.

The assumption of normally distributed return data is also examined by employing two statistical tests. Both the Jarque-Bera test as well as the Lilliefors test are two-sided goodness-of-fit tests of whether the sample data stem from a normal distribution with unknown and hence estimated mean and variance. The Jarque-Bera test compares the pre-

Table 3.4: Normality and Stationarity Tests for the US and European MKT Series. This table lists the Jarque-Bera and the Lilliefors test for normality as well as the augmented Dickey-Fuller test for stationarity of a time series. The normality tests challenge the null hypothesis of having a normally distributed sample to the alternative hypothesis of the sample data being nonnormal. The test for stationarity challenges the null hypothesis of having a time series characterized by $\text{MKT}_t = \text{MKT}_{t-1} + \epsilon_t$, i.e. the time series contains a unit root and is nonstationary, to the alternative hypothesis of the time series being stationary and following an autoregressive (AR) process of order one, AR(1), i.e. $\text{MKT}_t = \phi \text{MKT}_{t-1} + \epsilon_t$ with $\phi < 1$ and white noise ϵ_t without any drift or deterministic time trend. The tests are applied to the US and European monthly MKT series covering the period January 1970 to December 2017 and January 1991 to December 2017, respectively. This table displays the test statistics of each test and the corresponding p-values.

	USA		**Europe**	
	Statistic	p-Value	Statistic	p-Value
Tests for Normality				
Jarque-Bera	123.2157	0.0010	60.7163	0.0010
Lilliefors	0.0553	0.0010	0.0663	0.0016
Test for Stationarity				
Augmented Dickey-Fuller	−21.9360	0.0010	−15.9355	0.0010

dicted skewness and kurtosis of the normal distribution to the actual data sample, whereas the Lilliefors test compares the empirical distribution function of the data sample with the cumulative distribution function of the predicted normal distribution. As shown in Table 3.4 and already recommended by the visual inspection of the MKT series, both tests reject the null hypothesis that the sample distribution is normal at the 5% significance level. This is true for the US as well as for the European MKT series. Table 3.4 also presents the test statistic and the corresponding p-value of the augmented Dickey-Fuller test for stationarity. The test rejects the null hypothesis at the 5% significance level suggesting a stationary MKT series for both the US and Europe. This result also implies that the data generating parameters such as the mean and variance do not change over time.

To sum up, the above findings are suggestive of an outperformance of the European market by the US market based on both a higher mean return and smaller standard deviation. In addition, the specific process of creating the portfolios considered in this analysis results in high positive correlations of assets within the different data sets. All return series exhibit the common characteristics of having a negative skewness and high kurtosis. This speaks in favor of frequent small return gains combined with the possibility of few rather extreme return losses. In fact, statistical tests reject the null hypothesis of having normally distributed return series, an information that needs to be accounted for to ensure reliable statistical inferences.

4. Optimized Portfolio Strategies and 1/N

This section presents a brief overview of different approaches on how to create optimized portfolios when facing estimation risk as is done in the related empirical literature. Subsequently, the optimized portfolio strategies employed in this analysis are introduced. However, to begin with, some notation and the general setting are defined.

4.1. Mean-Variance Framework

Return and risk measured in terms of variance of return are the key characteristics when considering assets and portfolios of assets. Due to their pivotal role with respect to portfolio optimization, it is crucial to understand the basic mechanisms at work when assets with different properties are combined to create a portfolio of assets. Let $r = (r_1, \ldots, r_N)^T$ denote the vector of uncertain excess returns for an investment universe consisting of N risky assets over the risk-free rate r_f, i.e. $r = R - r_f \mathbf{1}_N$ with the complementary vector of nonexcess returns $R \in \mathbb{R}^{N \times 1}$ and $\mathbf{1}_N$ is a vector of ones with dimension N. The expected excess returns of the assets $\mu(r_1, \ldots, r_N)$ can be stated as

$$\mu = (E(r_1), \ldots, E(r_N))^T. \tag{4.1}$$

The symmetric $N \times N$ variance-covariance matrix of the returns $\Sigma(r_1, \dots, r_N)$ is given by

$$\begin{aligned}\Sigma &= E((r-\mu)(r-\mu)^T)\\ &= \begin{pmatrix} V(r_1) & \mathrm{Cov}(r_1,r_2) & \dots & \mathrm{Cov}(r_1,r_N)\\ \mathrm{Cov}(r_2,r_1) & V(r_2) & \dots & \mathrm{Cov}(r_2,r_N)\\ \vdots & \vdots & \ddots & \vdots\\ \mathrm{Cov}(r_N,r_1) & \mathrm{Cov}(r_N,r_2) & \dots & V(r_N)\end{pmatrix}\\ &= \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1N}\\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2N}\\ \vdots & \vdots & \ddots & \vdots\\ \sigma_{N1} & \sigma_{N2} & \dots & \sigma_N^2\end{pmatrix}\end{aligned} \tag{4.2}$$

with the variance of asset i denoted by σ_i^2 and $\sigma_{ij} = \sigma_i\sigma_j\rho_{ij}$ is the covariance between asset i and j with correlation coefficient ρ_{ij}. A portfolio can be completely characterized by specifying the proportion that each available risky asset takes in the portfolio. This information can be stored in $x = (x_1, \dots, x_N)^T$ with the individual weighting of asset i denoted by x_i such that $x_f = 1 - \mathbf{1}_N^T x$ is the investment in the risk-free asset and the representative investor is fully invested, i.e. the weights sum up to one. Consequently, the return of a portfolio in excess of the risk-free rate $r_p(x_1, \dots, x_N)$ takes the form

$$r_p = \sum_{i=1}^{N} x_i r_i = x^T r. \tag{4.3}$$

The expected excess return of a portfolio $\mu_p(x_1, \dots, x_N)$ consisting of N different assets can be stated as

$$\mu_p = E(r_p) = \sum_{i=1}^{N} x_i E(r_i) = x^T \mu. \tag{4.4}$$

The corresponding variance of a portfolio $\sigma_p^2(x_1, \dots, x_N)$ can be specified as

$$\sigma_p^2 = \sum_{i=1}^{N} x_i^2\sigma_i^2 + \sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N} x_i x_j \sigma_i \sigma_j \rho_{ij} = x^T \Sigma x, \tag{4.5}$$

where Σ is expected to be positive definite, i.e. $x^T\Sigma x > 0$, $x \neq 0$, to guarantee that the assets under consideration are not redundant, i.e. are linearly independent with respect to their return pattern. Although the existence of such an asset would not alter the result of the underlying optimization problem the invertibility of the variance-covariance matrix needs to be secured due to computational issues.

For the representative investor to behave in accordance with the expected utility hypothesis at least one of the two assumptions have to be fulfilled. Either the investor is subject to have quadratic utility, although other utilities can be evaluated utilizing a second-order approximation, or the distributions of asset returns are jointly normally distributed, i.e. can be completely characterized by only the first two moments, while the individual returns are all independent and identically distributed, respectively. Pennacchi (2008) provides a general proof. Quadratic utility implies an increasing relative and absolute risk aversion with increasing investor's wealth, a fact that seems to contradict realistic behavior as argued in Cohn et al. (1975) and Morin and Suarez (1983). Normally distributed returns refer to the class of elliptical distributions such as the normal Gaussian distribution. However, empirically returns rather follow a leptokurtic distribution that cannot entirely be described only by its mean and variance in accordance to the findings of Section 3. Also, investors aiming at maximizing expected utility not necessarily only care about the first two moments of the underlying distribution but might have also preferences for higher moments. Despite these potential limitations and assumptions, the beauty of a model is not based on the exact replication of the complex world in all detail but rather on abstraction and simplification while still maintaining the actual purpose of a model that is providing meaningful insights.

In this spirit and given a time period of T observations, the representative investor chooses a composition x_t at every period t, $t = 1, \ldots, T$, of the available N risky assets to create a portfolio so as to maximize expected utility, $E(u)$, i.e.

$$\max_{x_t} E(u(x_t)) = x_t^T \mu_t - \frac{\gamma}{2} x_t^T \Sigma_t x_t. \tag{4.6}$$

The expected utility is an increasing function in the portfolio's expected excess return and decreasing in the portfolio's variance with the investor's specific risk-aversion parameter γ representing the strength of this fundamental trade-off between expected return and risk. It is not necessary to explicitly constrain the weights to sum to one as this condition is already implicitly incorporated in the optimization problem due to the fact that excess returns rather than nonexcess returns are considered, see Appendix B. This optimization problem has a closed-form solution that can be obtained by taking the first derivative of Eq. (4.6) with respect to x_t and takes the form

$$x_t = \frac{1}{\gamma} \Sigma_t^{-1} \mu_t. \tag{4.7}$$

In general, this solution implies an allocation to the unique risky portfolio as well as to the risk-free asset with a larger share invested in the risk-free asset the higher the investor's risk-aversion parameter. Finally, the adjusted weights

$$\omega_t = \frac{x_t}{\left|\mathbf{1}_N^T x_t\right|} = \frac{\frac{1}{\gamma}\Sigma_t^{-1}\mu_t}{\left|\mathbf{1}_N^T \frac{1}{\gamma}\Sigma_t^{-1}\mu_t\right|} = \frac{\Sigma_t^{-1}\mu_t}{\left|\mathbf{1}_N^T \Sigma_t^{-1}\mu_t\right|} \tag{4.8}$$

yield the relative amount allocated to the risky portfolio at date t. The rescaling by the absolute value guarantees a proportional investment between the risky portfolio and the risk-free asset across the different strategies such that performance differentials do not depend on an unequal investment share with respect to the risk-free asset.

4.2. Estimation Risk

The optimization problem in Eq. (4.6) crucially depends on the estimation of expected returns, variances, and covariances. The number of parameters one would need to estimate to compute the optimal mean-variance portfolio when considering 100 risky assets already amounts to 5,150[2]. Apart from the fact that the estimation of a large number of quantities can be rather time consuming and a computational challenging task, there is also a certain level of risk inherent in every estimation. In fact, there is no guarantee that the obtained estimates coincide with the actual true underlying values. Already Merton (1980) shows evidence that the estimation of the variance-covariance matrix is much more precise relative to the estimation of the corresponding expected returns. Chopra and Ziemba (1993) further specify that the most critical estimations with respect to estimation accuracy are the expected returns, followed by the variance estimates and lastly the estimates of the covariances. The difficulty in estimating expected returns has its origin in the typically observed relatively small average return values in combination with relatively high return volatilities such that only a very long time series of observations can guarantee an accurate estimation. Broadie (1993) calculates that 26 years of monthly return data are needed to correctly distinguish with a probability of 90% between two assets with different expected monthly returns of 1 and 1.5%, respectively. This result assumes normally distributed returns, a common standard deviation of 7%, and a correlation coefficient of 0.5. This example nicely illustrates how difficult it is to obtain precise return estimates or rather how return estimates critically depend on the number of observations available for estimation. In contrast, less than five years of monthly data are required to distinguish with a probability of 90% between two assets with different standard deviations of 6 and 7%, respectively. This computation is based on normally distributed returns with a common mean of 1% and a correlation coefficient of 0.5, see Broadie (1993). Still, having a large data set also introduces new challenges to the estimation process since the probability of the time series being nonstationary, for example, increases with the length of the data set as argued in Jobson and Korkie (1980). This constitutes a trade-off between estimation precision and estimation validity that is omnipresent in the context of out-of-sample portfolio optimization. Over the past decades, the related empirical literature has developed different econometric approaches on how to mitigate this estimation risk.

[2]For a set of $N \in \mathbb{N}$ risky assets, N returns r_i, N variances σ_i^2, and $N(N-1)/2$ corresponding covariances σ_{ij}, $i = 1, \ldots, N$, $i \neq j$, need to be estimated resulting in a total of $2N + N(N-1)/2$ parameters. Additional parameters needed are the risk-free rate r_f and the risk-aversion parameter γ.

4.3. Approaches to Fight Estimation Risk

The Bayesian approach starts from the premise that the parameters of the data generating process are not known to the representative investor. As a result, the investor is faced with the task of incorporating an educated guess about the data generating process into the expected utility maximization problem. These a priori inferences, or priors, represent the investor's beliefs about the parameters of the data generating process before being presented with evidence. The class of priors is commonly divided into uninformative or diffuse priors and informative priors. The former such as proposed in Barry (1974), Klein and Bawa (1976), and Brown (1979) do not systematically incorporate any specific information about the parameters of the return distribution. Since the Bayesian approach explicitly accounts for potential estimation errors, the already risky assets become even more risky. The risk of the risk-free asset, however, remains zero by assumption. In comparison to the mean-variance plug-in approach, the Bayesian solution hence typically invests relatively less in the risky assets and more in the risk-free asset. Nevertheless, the results obtained from these general diffuse priors are very similar to the mean-variance approach and even coincide in case of an infinitely long time series of observations, see Brandt (2009). In contrast, informative priors do incorporate specific information. The arguably most well-known approach is based on Black and Litterman (1992).

First introduced by Stein (1956) and then extended by James and Stein (1961), the shrinkage technique also aims at reducing potential estimation errors. The basic idea of shrinkage, nicely illustrated in Efron and Morris (1977), is that the quality of the estimate of the expected returns should increase if the sample mean is shrinked toward a common value or grand mean, i.e. the mean of the means across all variables. In particular, the shrunked mean is supposed to be less likely affected by extreme observations relative to the sample mean. This approach can be understood as a special case of the Bayesian technique in which the prior is the shrinkage target. The resulting reduction in the estimate's variance is assumed to be more beneficial than the harm that occurs by shrinking and thus biasing the estimate. Shrinkage can thus be seen as a trade-off between either incurring a bias or having low levels of variance. In general, the shrinkage factor is an increasing function in the number of assets and a decreasing function in the number of time periods considered. Jobson and Korkie (1980), Jorion (1985), and Jorion (1986) focus on shrinking expected returns, whereas Frost and Savarino (1986), Ledoit and Wolf (2003), and Ledoit and Wolf (2004) extend the traditional approach and apply the shrinking technique on variances and covariances. Kourtis et al. (2012) avoid the detour and directly shrink the inverse of the variance-covariance matrix. Moreover, shrinkage is not only limited to the moments of the asset return distribution but can also be applied to optimal portfolio weights. Applications can be found in Golosnoy and Okhrin (2007) and Frahm and Memmel (2010).

A further specification of a prior is given by factor models. Factor models try to facilitate the estimation of expected returns by identifying a limited number of variables able to explain the observed variation in asset return cross-sections or time series. The Capital Asset Pricing Model (CAPM) as suggested by Sharpe (1964), Lintner (1965), and Mossin (1966) represents a special case of a single-factor model and links the expected asset return to the expected excess market return. The number of parameters one would need to estimate to compute the optimal mean-variance portfolio when considering 100 risky assets decreases from 5,150 based on the mean-variance plug-in approach to 302[3]. This reduction in the number of parameters to be estimated illustrates the advantages of factor models with respect to the estimation procedure. Applications of this approach can be found in Pástor (2000) and Pástor and Stambaugh (2000). However, the workhorse of the empirical factor model literature is the Fama-French three-factor model as proposed in Fama and French (1993) including its extensions to a four-factor model as in Carhart (1997) and a five-factor model as in Fama and French (2015).

Additional strands consider different specific portfolio restrictions and hence are supposed to model real-world trading in a more realistic fashion. By doing so, short sale constraints play a key role. In this context, Jorion (1992) proposes an interesting thought experiment. He considers a world with two assets in which asset A and B are assumed to have an estimated average return of 10.1 and 9.9%, respectively, although both assets have a true return of 10%. A short sale restricted investor would choose to invest mainly in asset A based on its higher return. Without short-sale constraints, however, the investor will heavily buy asset A and short asset B to take advantage of the return differential of 0.2%. Since the resulting profit of the portfolio is increasing with the long position in asset A and the short position in asset B, the optimization process can result in very extreme investment positions. In response to that finding, constraining the weights to be nonnegative not only mitigates the chance of adopting extreme portfolio weights due to estimation error but also leads to more realistic outcomes. DeMiguel et al. (2009a) extend this approach by focusing on the minimum-variance portfolio while constraining different norms of the portfolio weight vector to be smaller than a given threshold. Furthermore, it is also possible to restrict short sales only to specific single assets or to allow short sales in general but only up to a limited degree. Other real-world examples of restrictions that investors might face are specific constitutional regulations or policy constraints that only allow the investment in socially responsible assets such as given by companies that do not violate human rights, support corruption, or negatively affect the environment.

4.4. Portfolio Strategies

The selection of models included in this analysis aims at fulfilling two objectives. First, past empirical literature is respected by including already well-established optimized portfolio strategies such as the sample-based mean-variance and the Bayes-Stein strategy.

[3] There are only $3N + 2$ parameters that need to be estimated namely N returns, N betas, N variances, the expected market return, and the variance of the market return. Additional parameters needed are the risk-free rate r_f and the risk-aversion parameter γ.

Table 4.1: Overview Portfolio Strategies. This table lists a total of 16 portfolio strategies that are included in this analysis. The first model is the naive diversification approach that also serves as the benchmark model on which the performances of the remaining 15 optimized portfolio strategies are measured against. The abbreviations are introduced to refer to the specific model in the text. The last column of the table lists the number of estimated parameters needed to implement a specific model under the assumption of an investment universe consisting of N different risky assets. This number excludes the estimation of the risk-free rate r_f and the risk aversion parameter γ.

#	Portfolio Strategy	Abbreviation	Number of Estimations
	Naive Diversification		
1	Equally-weighted	EW	0
	Sample-Based Mean Variance		
2	Sample based mean-variance	MV	$(N^2 + 3N)/2$
	Moment Restrictions		
3	Minimum-variance	MinV	$(N^2 + N)/2$
4	Value-weighted	VW	0
5	MacKinlay-Pástor	MP	$(N^2 + 3N)/2$
	Bayesian Approach		
6	Bayes-Stein	BS	$(N^2 + 3N)/2$
	Portfolio Weight Constraints		
7	Mean-variance short sale constraint	MVsc	$(N^2 + 3N)/2$
8	Minimum-variance short sale constraint	MinVsc	$(N^2 + N)/2$
9	Bayes-Stein short sale constraint	BSsc	$(N^2 + 3N)/2$
10	Mean-variance long-short sale constraint	MVlsc	$(N^2 + 3N)/2$
	Optimal Combinations of Strategies		
11	Mean-variance and minimum-variance	MV/MinV	$(N^2 + 3N)/2$
12	Equally-weighted and minimum-variance	EW/MinV	$(N^2 + N)/2$
13	Equally-weighted and mean-variance	EW/MV	$(N^2 + 3N)/2$
14	Equally-weighted and mean-variance and minimum-variance	EW/MV/MinV	$(N^2 + 3N)/2$
	Timing Strategies		
15	Volatility timing	VT	N
16	Reward-to-risk timing	RRT	$2N$

Second, prior research is updated by also considering rather newly proposed strategies such as the volatility and reward-to-risk timing. Furthermore, different theoretically appealing models such as the minimum-variance strategy are included to serve as equal competitors in the horse race between optimized portfolio strategies. Above all, the naive diversification approach is empirically very successful in terms of out-of-sample portfolio performance although completely lacking any inherent optimization and thus defines the benchmark to beat. Table 4.1 presents an overview of the optimized portfolio strategies considered in this analysis. All models are assigned a class that describes the underlying applied technique of creating the specific portfolios. However, this classification is not mutually exclusive such that models might also be attributed to different classes. Still, the intention is to provide a first impression on the different characteristics inherent in the models under consideration. In addition, Table 4.1 also displays the number of estimated parameters needed to implement a specific model under the assumption of an investment universe consisting of N different risky assets. These numbers are very homogenous throughout the strategies and mainly adopt the value of $\frac{N^2+3N}{2}$ which corresponds to the number of estimates needed to establish the mean-variance strategy or $\frac{N^2+N}{2}$ which refers to the number of estimates needed for the implementation of the minimum-variance strategy. This information, however, is only of quantitative nature and completely lacks qualitative aspects. To be precise, the difference in the number of parameters needed to be estimated between the mean-variance and the minimum-variance approach is rather limited and only amounts to N. Nevertheless, this difference represents the N arguably most difficult expected return estimations such that the gain in estimation accuracy by excluding these N return estimates might be larger than ex ante expected by only looking at the quantitative differences. In addition, some models repeatedly utilize the estimated parameters and hence potential estimation errors can have more adverse effects in comparison to only a limited or even single usage of estimation results. The following presentation of the optimized portfolio strategies predominantly focuses on the respective underlying idea of optimization and states the parameters needed to implement a specific strategy. Explicit optimization procedures are note stated but can be reviewed by consulting the corresponding references mentioned in the text.

4.4.1. Equally-Weighted

The equally-weighted (EW) approach or simply naive diversification or $1/N$ describes a schedule that invests an equal share of $1/N$ in each of the N available risky assets, i.e. $x_i = 1/N$, $i = 1, \ldots, N$. The relative importance of this investment approach lies in its simplicity and widespread use. By studying the American 401(k) defined contribution saving plan which is characterized by the fact that participants themselves decide on their retirement saving schedule and thus pension income, Benartzi and Thaler (2001) find that a considerable fraction of participants equally distribute their contributions across the available investment opportunities. This result shows evidence that investment decisions even as important as deciding on the financial security for retirement do not necessarily rely on more sophisticated portfolio strategies. Huberman and Jiang (2006) subsequently add some more information and show that participants hardly ever invest in more than four funds irrespective of the actual number of available investment opportunities. Benartzi and Thaler (2007), in turn, discover that the part of the sign-up form used to indicate in which funds to invest only consists of four lines. Hence, to choose more funds an additional form needs to be requested representing an effort that might hinder further investments. This observation can explain the limited number of chosen funds, however, it does not rationalize the widespread use of equally dividing the investment. In particular, the question arises whether people even have a sufficient degree of financial literacy and willpower to consistently execute more sophisticated strategies. And yet, nothing has been said about the actual performance of the naive diversification approach relative to optimized portfolio strategies. To start with, some thoughts about diversification.

Diversification in general can result in a reduction of a portfolio's overall volatility hence representing a desired effect. The variance of a portfolio $\sigma_p^2(x_1, \ldots, x_N)$ consisting of N risky assets can be stated as

$$\sigma_p^2 = \sum_{i=1}^{N} x_i^2 \sigma_i^2 + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} x_i x_j \sigma_i \sigma_j \rho_{ij} \tag{4.9}$$

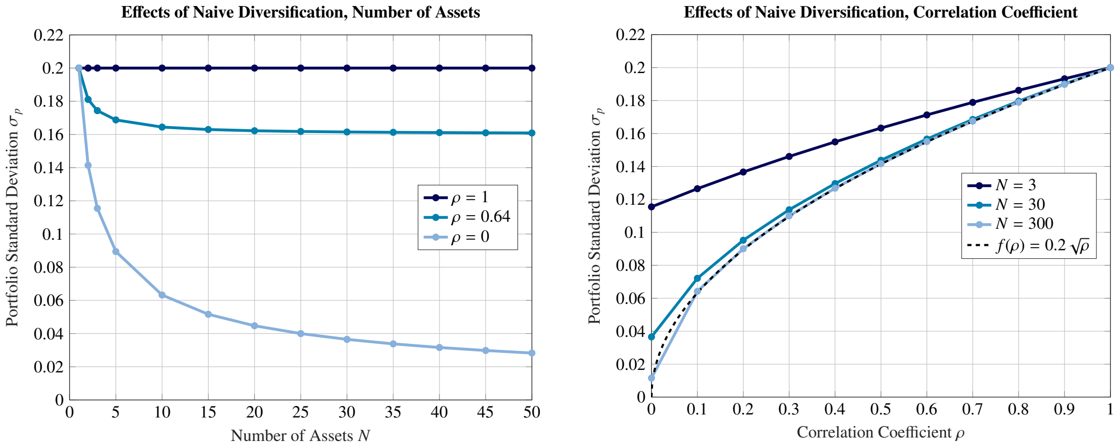
and thus is not only the sum of the portfolio's individual assets but also depends on the covariance. In fact, the correlation coefficient ρ is the key component that rationalizes the possibility of diversification in the first place. Without loss of generality, assume that $x_i > 0$ and $x_j > 0$, $i, j = 1, \ldots, N$, and $0 \leq \rho_{ij} \leq 1$ for all i and j, $i \neq j$. In general, diversification benefits increase the lower the correlation between the assets so that the minimum variance is obtained for $\rho_{ij} = 0$. However, also the number of assets used to create a specific portfolio heavily influences the level of overall risk. In case of naive diversification, i.e. $x_i = 1/N$, and under the assumptions of homogenous individual variances, i.e. $\sigma_i^2 = \sigma^2$ for all i, and homogenous correlation coefficients, i.e. $\rho_{ij} = \rho$ for all i and j, $i \neq j$, Eq. (4.9) reduces to

$$\begin{aligned} \sigma_p^2 &= N \frac{1}{N^2} \sigma^2 + (N^2 - N) \frac{1}{N^2} \sigma^2 \rho \\ &= \frac{1}{N} \sigma^2 + \sigma^2 \rho - \frac{1}{N} \sigma^2 \rho. \end{aligned} \tag{4.10}$$

The first derivative of Eq. (4.10) with respect to the number of assets N negatively depends on the expression $1/N^2$, i.e. including more assets to the portfolio keeps reducing the risk, but this effect decreases with an increasing number of assets. Utilizing Eq. (4.10), it is also true that

$$\lim_{N \to \infty} \sigma_p^2 = \sigma^2 \rho = \mathrm{Cov}(r_i, r_j). \tag{4.11}$$

Fig. 4.1a depicts this relationship by plotting the standard deviation σ_p of a naive diversified portfolio against the number of assets included in the portfolio. Based on Eq. (4.10), it can be seen that in case of perfect positive correlation, i.e. $\rho = 1$, the variance of the portfolio is equal to the assumed homogenous variance of $\sigma = 0.2$ and no diversification effect can be seen. In case of no correlation at all, i.e. $\rho = 0$, the variance vanishes in the limit of $N \to \infty$ and no risk remains. The intermediate case considers a correlation coefficient of $\rho = 0.64$. According to Eq. (4.11) and for $N \to \infty$, the portfolio's standard deviation σ_p is at most $1 - \sqrt{\rho}$ percentage smaller relative to σ and hence Eq. (4.11) establishes an upper bound for the diversification effect. The maximum reduction of the portfolio's standard deviation in this scenario therefore amounts to 20% of $\sigma = 0.2$. Indeed, Fig. 4.1a illustrates the stagnation of risk reduction at a standard deviation level of 0.16 as the number of assets increases. This level defines the cutoff between the unsystematic or idiosyncratic risk that can be diversified away by increasing the number of assets



(a) Effects of naive diversification with focus on the number of assets. The graph contains three examples for different correlation parameter values of $\rho = 1$, $\rho = 0.64$, and $\rho = 0$, respectively.

(b) Effects of naive diversification with focus on the correlation coefficient. The graph contains three examples for different numbers of assets with $N = 3$, $N = 30$, and $N = 300$, respectively. In addition, the function $f(\rho) = 0.2\sqrt{\rho}$ serves as a reference for the case of $N \to \infty$.

Figure 4.1: Effects of Naive Diversification on the Portfolio's Standard Deviation. This figure shows effects of naive diversification with respect to the portfolio's standard deviation. It is assumed that all N assets have identical variances, i.e. $\sigma_i^2 = \sigma^2$ for all i, and identical correlation coefficients, i.e. $\rho_{ij} = \rho$ for all i and j, $i \neq j$, with $i, j = 1, \ldots, N$. The standard deviation is set to $\sigma = 0.2$.

included in the portfolio and the systematic or market risk of 0.16 that cannot be diversified away. Fig. 4.1a also suggests that a sound diversification can already be achieved by investing in approximately 20 assets. However, this picture is a bit oversimplified in the sense that empirically more assets are required to obtain well-diversified portfolios. Respecting transaction costs, Statman (1987) states that at least 30 stocks are needed to establish a sufficient diversified portfolio. This number is obtained by comparing the costs of diversification, i.e. transactions costs, to the benefits of diversification, i.e. reduction in risk. Adding more assets to the portfolio is beneficial as long as the benefits of doing so outweigh the costs. A more general proof that diversification pays out can be found in Samuelson (1967).

Fig. 4.1b depicts the relation between the portfolio's standard deviation and the correlation coefficient ρ. As the correlation coefficient approaches zero, the portfolio's standard deviation is decreasing. This observation is independent of the number of assets. Based on Eq. (4.11), the standard deviation of the portfolio σ_p converges to $\sigma\sqrt{\rho}$ for $N \to \infty$. Since the square root is a concave function, the diversification effect and thus the decrease in risk increases relatively more for low values of ρ.

To sum up, the naive diversification approach is not only effectively used in real-world investment decisions but also theoretically reasonable at least in terms of diversification aspects. The naive diversification approach is a portfolio strategy that, at any date t, invests $1/N$ in each of the N available risky assets, i.e. $x_t^{\text{EW}} = 1/N \in \mathbb{R}^{N \times 1}$. As a result, this strategy is completely independent of the underlying data and does not require any estimation procedure thus lacking any sort of estimation risk. The relative vector of portfolio weights is given by

$$\omega_t^{\text{EW}} = \frac{x_t^{\text{EW}}}{\left|\mathbf{1}_N^T x_t^{\text{EW}}\right|} = \frac{\frac{1}{N}\mathbf{1}_N}{\left|\mathbf{1}_N^T \frac{1}{N}\mathbf{1}_N\right|} = \frac{1}{N}\mathbf{1}_N. \tag{4.12}$$

4.4.2. Sample-Based Mean-Variance

The weights of the sample-based mean-variance (MV) strategy are in accordance with Eq. (4.8), i.e. the relative vector of portfolio weights is

$$\omega_t^{\text{MV}} = \frac{\Sigma_t^{-1}\mu_t}{\left|\mathbf{1}_N^T \Sigma_t^{-1}\mu_t\right|}. \tag{4.13}$$

In fact, based on an estimation period of M observations, the excess mean return and the variance-covariance matrix are substituted with their corresponding sample counterparts such that the excess sample mean return $\hat{\mu}_t$ is calculated as

$$\hat{\mu}_t = \frac{1}{M}\sum_{\tau=t-M+1}^{t} \mu_\tau \tag{4.14}$$

and the sample variance-covariance matrix $\hat{\Sigma}_t$ is obtained via

$$\hat{\Sigma}_t = \frac{1}{M-N-2}\sum_{\tau=t-M+1}^{t} (\mu_\tau - \hat{\mu}_t)(\mu_\tau - \hat{\mu}_t)^T . \tag{4.15}$$

The particular form of the sample variance-covariance matrix is chosen based on the fact that its inverse $\hat{\Sigma}_t^{-1}$ is an unbiased estimator of Σ_t^{-1} in case of $M > N + 4$, see Kan and Zhou (2007). This strategy does not incorporate any specific estimation technique and therefore is exposed to potential estimation errors.

4.4.3. Minimum-Variance

The minimum-variance (MV) strategy refers to the optimization problem given by

$$\min_{\omega_t^{\text{MinV}}} \frac{1}{2}(\omega_t^{\text{MinV}})^T \Sigma_t \omega_t^{\text{MinV}} \quad \text{subject to} \quad \mathbf{1}_N^T \omega_t^{\text{MinV}} = 1. \tag{4.16}$$

The solution to this optimization problem is

$$\omega_t^{\text{MinV}} = \frac{\Sigma_t^{-1}\mathbf{1}_N}{\mathbf{1}_N^T \Sigma_t^{-1}\mathbf{1}_N}, \tag{4.17}$$

see Appendix C. This strategy is implemented by substituting the variance-covariance matrix with its sample-based counterpart $\hat{\Sigma}_t$. In contrast to the sample-based mean-variance strategy, the weights of the minimum-variance strategy are only dependent on the sample variance-covariance matrix and avoid estimating the arguably most difficult and critical expected returns. Therefore, this strategy is relatively less exposed to estimation errors.

In fact, the minimum-variance portfolio strategy does not generally belong to the class of optimized approaches with respect to the mean-variance trade-off. This would be only true in case of identical expected returns across all assets. Nevertheless, the minimum-variance strategy is worth to be considered due to its theoretically limited estimation error and because some of the following optimized portfolio strategies do incorporate this strategy.

4.4.4. Value-Weighted

The two-fund separation theorem according to Tobin (1958) states that there exists a unique efficient risky market portfolio which every mean-variance investor should hold. The separation of the total investment between this risky market portfolio and the risk-free asset is then only dependent on the investor's individual risk-aversion parameter γ. The CAPM by Sharpe (1964), Lintner (1965), and Mossin (1966) can be understood as a model whose equilibrium outcome coincides with this very efficient risky market portfolio. In fact, it is exactly the value-weighted market portfolio in which each asset is represented according to its relative value of market equity. The expected excess return of this value-weighted (VW) portfolio strategy is directly obtained from the data sets and given by the US and European MKT series, respectively.

4.4.5. MacKinlay-Pástor

MacKinlay and Pástor (2000) start from the assumption that asset returns can be perfectly described by factor models such as described in Section 4.3. However, in most or even all reasonably scenarios the exact model is not known, i.e. risk factors are missing. Nevertheless, the mispricing of asset returns based on the lack of a risk factor translates to the variance-covariance matrix of the residuals and this knowledge can be exploited to improve the estimation of the expected returns. By doing so, the variance-covariance matrix of the returns transforms more and more into a simple identity matrix of the form $\sigma_t^2 I_N$ the higher the level of uncertainty about the underlying asset pricing factor model. According to the MacKinlay and Pástor (2000) (MP) strategy, the variance-covariance matrix of returns is expressed as

$$\Sigma_t^{\text{MP}} = \alpha_t \mu_t \mu_t^T + \sigma_t^2 I_N \tag{4.18}$$

where α_t is a term that represents the model misspecification. The optimal weights can then be reformulated as

$$\begin{aligned} x_t^{\text{MP}} &= \frac{1}{\gamma} (\Sigma_t^{\text{MP}})^{-1} \mu_t^{\text{MP}} \\ &= \frac{1}{\gamma} \frac{1}{\sigma_t^2} \left(I_N - \frac{\alpha_t \mu_t^{\text{MP}} (\mu_t^{\text{MP}})^T}{\sigma_t^2 + \alpha_t (\mu_t^{\text{MP}})^T \mu_t^{\text{MP}}} \right) \mu_t^{\text{MP}} \\ &= \frac{1}{\gamma} \frac{1}{\sigma_t^2 + \alpha_t (\mu_t^{\text{MP}})^T \mu_t^{\text{MP}}} \mu_t^{\text{MP}}. \end{aligned} \tag{4.19}$$

The estimates for α_t, μ_t^{MP}, and σ_t^2 can be obtained via the process of a likelihood maximization, however, this demanding optimization can be avoided by applying a semi-analytical solution as presented in Tu and Zhou (2011) yielding

$$\hat{x}_t^{\mathrm{MP}} \approx \frac{1}{\gamma} \frac{\tilde{\mu}_t}{\hat{\lambda}_{1,t} - \tilde{\mu}_t^T \tilde{\mu}_t}. \tag{4.20}$$

In this context, $\hat{\lambda}_{1,t}$ denotes the largest eigenvalue and $\hat{v}_{1,t}$ the corresponding eigenvector of the matrix $\hat{\Sigma}_t^{\mathrm{ML}} + \hat{\mu}_t^{\mathrm{ML}} (\hat{\mu}_t^{\mathrm{ML}})^T$, where the maximum likelihood (ML) estimates of the sample mean and the variance-covariance matrix are given by $\hat{\mu}_t^{\mathrm{ML}}$ and $\hat{\Sigma}_t^{\mathrm{ML}}$, respectively. The adjusted return is expressed as $\tilde{\mu}_t = \hat{v}_{1,t} \hat{\mu}_t^{\mathrm{ML}} \hat{v}_{1,t}$ such that the relative vector of portfolio weights can be stated as

$$\hat{\omega}_t^{\mathrm{MP}} = \frac{\hat{x}_t^{\mathrm{MP}}}{\left|\mathbf{1}_N^T \hat{x}_t^{\mathrm{MP}}\right|} = \frac{\frac{1}{\gamma} \frac{\tilde{\mu}_t}{\hat{\lambda}_{1,t} - \tilde{\mu}_t^T \tilde{\mu}_t}}{\left|\mathbf{1}_N^T \frac{1}{\gamma} \frac{\tilde{\mu}_t}{\hat{\lambda}_{1,t} - \tilde{\mu}_t^T \tilde{\mu}_t}\right|} = \frac{\tilde{\mu}_t}{\left|\mathbf{1}_N^T \tilde{\mu}_t\right|}. \tag{4.21}$$

4.4.6. Bayes-Stein

The Bayes-Stein (BS) portfolio strategy is a combination of shrinkage and Bayesian estimation. Following Jorion (1986), the expected sample mean return estimator is obtained by shrinking the expected sample mean return toward the mean of the expected sample return estimate of the minimum-variance portfolio $\hat{\mu}_{p,t}^{\mathrm{MinV}} = \hat{\mu}_t^T \hat{\omega}_t^{\mathrm{MinV}}$, i.e.

$$\hat{\mu}_t^{\mathrm{BS}} = (1 - \hat{\phi}_t)\hat{\mu}_t + \hat{\phi}_t \hat{\mu}_{p,t}^{\mathrm{MinV}}, \tag{4.22}$$

where the shrinkage factor $0 < \hat{\phi}_t < 1$ takes the form

$$\hat{\phi}_t = \frac{N+2}{(N+2) + M(\hat{\mu}_t - \hat{\mu}_{p,t}^{\mathrm{MinV}} \mathbf{1}_N)^T \hat{\Sigma}_t^{-1} (\hat{\mu}_t - \hat{\mu}_{p,t}^{\mathrm{MinV}} \mathbf{1}_N)}. \tag{4.23}$$

Supported by empirical evidence, Jorion (1985) argues that shrinking the expected sample mean toward the minimum-variance portfolio improves the estimation due to the fact that the minimum-variance portfolio is the optimal choice in case of identical expected returns of all assets. The Bayes-Stein approach additionally estimates the variance-covariance matrix by applying a traditional Bayesian estimation technique to account for potential estimation error. This procedure leads to an estimate of the form

$$\hat{\Sigma}_t^{\mathrm{BS}} = \hat{\Sigma}_t \left(1 + \frac{1}{M + \hat{\kappa}_t}\right) + \frac{\hat{\kappa}_t}{M(M + 1 + \hat{\kappa}_t)} \frac{\mathbf{1}_N \mathbf{1}_N^T}{\mathbf{1}_N^T \hat{\Sigma}_t^{-1} \mathbf{1}_N} \tag{4.24}$$

with $\hat{\kappa}_t = M\hat{\phi}_t / (1 - \hat{\phi}_t)$ being the precision parameter of the prior

$$p(\mu_t | \Sigma_t, \mu_{p,t}^{\mathrm{MinV}}, \kappa_t) \propto \exp\left(-\frac{1}{2}(\mu_t - \mu_{p,t}^{\mathrm{MinV}} \mathbf{1}_N)^T (\kappa_t \Sigma_t^{-1})(\mu_t - \mu_{p,t}^{\mathrm{MinV}} \mathbf{1}_N)\right). \tag{4.25}$$

As a result, the solution of the underlying optimization problem stated in Eq. (4.7) transforms to

$$\hat{x}_t^{\text{BS}} = \frac{1}{\gamma}(\hat{\Sigma}_t^{\text{BS}})^{-1}\hat{\mu}_t^{\text{BS}}. \tag{4.26}$$

The relative vector of portfolio weights is given by

$$\hat{\omega}_t^{\text{BS}} = \frac{\hat{x}_t^{\text{BS}}}{\left|\mathbf{1}_N^T \hat{x}_t^{\text{BS}}\right|} = \frac{\frac{1}{\gamma}(\hat{\Sigma}_t^{\text{BS}})^{-1}\hat{\mu}_t^{\text{BS}}}{\left|\mathbf{1}_N^T \frac{1}{\gamma}(\hat{\Sigma}_t^{\text{BS}})^{-1}\hat{\mu}_t^{\text{BS}}\right|} = \frac{(\hat{\Sigma}_t^{\text{BS}})^{-1}\hat{\mu}_t^{\text{BS}}}{\left|\mathbf{1}_N^T(\hat{\Sigma}_t^{\text{BS}})^{-1}\hat{\mu}_t^{\text{BS}}\right|}. \tag{4.27}$$

4.4.7. Portfolio Weight Constraints

The portfolio weight constraints considered in this analysis result in four specifications. On the one hand, short sale constraints are applied to the sample-based mean variance (MVsc), minimum-variance (MinVsc), and Bayes-Stein (BSsc) strategy, respectively, i.e. all weights are forced to adopt positive values. As a result, the mean-variance optimization problem alters to a quadratic optimization problem including linear inequality constraints and can be stated as

$$\max_{x_t} E(u(x_t)) = x_t^T \mu_t - \frac{\gamma}{2} x_t^T \Sigma_t x_t \quad \text{subject to} \quad x_t \geq 0. \tag{4.28}$$

Unfortunately, there exists no closed-form solution to this optimization problem such that a standard numerical quadratic optimization program is applied to obtain the respective final weight vectors.

DeMiguel et al. (2009b) show that adding a nonnegativity weight constraint to the mean-variance strategy is effectively a form of shrinkage of the expected returns toward the average of all the expected returns. In the same spirit, Jagannathan and Ma (2003) argue that the nonnegativity weight restriction of the minimum-variance strategy in fact shrinks the larger elements of the variance-covariance matrix toward zero and hence results in similar outcomes as the approaches focusing solely on the shrinkage of the variance-covariance matrix such as in Ledoit and Wolf (2004). As always, shrinking can potentially affect the result in two ways. Specification error introduced by the shrinking process can render the results biased, whereas the level of estimation error decreases relative to nonshrinking. This trade-off is assumed to be in favor of the shrinking technique such that the gain in estimation accuracy outweighs the loss incurred from obtaining biased results.

In addition, a long-short sale constraint is considered that not only constraints short sales but also long sales. This strategy is often applied in the context of mutual and hedge funds and arguably represents are more realistic scenario relative to no constraints at all or short sale only constraints, see also Michaud (1993) and Sorensen et al. (2007). In fact, these portfolio weight constraint strategies might come very close to the truth in terms of mimicking actual investment behavior. The specific long-short sale constraint strategy considered in this analysis takes the form

$$\max_{x_t} E(u(x_t)) = x_t^T \mu_t - \frac{\gamma}{2} x_t^T \Sigma_t x_t \quad \text{subject to} \quad -0.3 \leq x_t \leq 1.3. \tag{4.29}$$

This represents the so-called 130/30 strategy, i.e. the portfolio can be exposed up to 130% in long equity positions and up to 30% in short equity positions such that the total exposure to equity still amounts to a 100% long position. This strategy is only considered with respect to the sample-based mean-variance strategy and is referred to as MVlsc.

4.4.8. Mean-Variance and Minimum-Variance

Kan and Zhou (2007) argue that in case of estimation uncertainty with respect to the data generating parameters, it is not optimal to hold the efficient market portfolio of risky assets in combination with the risk-free asset as proposed by Markowitz (1952) and Tobin (1958). The idea is that when estimating the efficient portfolio with error, including additional portfolios could be beneficial in terms of reducing estimation risk through diversification as long as the estimation errors of the efficient market portfolio and the newly introduced portfolio are not perfectly correlated. In particular, including a third portfolio to the risky and risk-free assets, the minimum-variance portfolio, can help to reduce estimation risk. The minimum-variance portfolio is chosen because it does not rely on the estimation of the expected mean returns and hence can be estimated with a relatively high precision. Kan and Zhou (2007) refer to this strategy as three-fund separation and formulate the weights as a weighted sum of the sample-based mean-variance portfolio that already incorporates the risk-free asset and the minimum-variance portfolio (MV/MinV), i.e.

$$\hat{x}_t^{\text{MV/MinV}} = \frac{1}{\gamma}\left(\hat{c}_t \hat{\Sigma}_t^{-1} \hat{\mu}_t + \hat{d}_t \hat{\Sigma}_t^{-1} \mathbf{1}_N\right). \tag{4.30}$$

The individual weights $\hat{c}_t$ and $\hat{d}_t$ are chosen so as to maximize the expected out-of-sample performance with respect to the utility of a mean-variance investor. It is

$$\hat{c}_t = \frac{(M-N-1)(M-N-4)}{M(M-2)} \frac{\hat{\psi}^2_{\text{adj.},t}}{\hat{\psi}^2_{\text{adj.},t} + \frac{N}{M}} \tag{4.31}$$

and

$$\hat{d}_t = \frac{(M-N-1)(M-N-4)}{M(M-2)} \frac{\frac{N}{M}}{\hat{\psi}^2_{\text{adj.},t} + \frac{N}{M}} \hat{\mu}_{p,t}^{\text{MinV}} \tag{4.32}$$

with $\hat{\mu}_{p,t}^{\text{MinV}} = \hat{\mu}_t^T \hat{\omega}_t^{\text{MinV}}$ and

$$\hat{\psi}^2_{\text{adj.},t} = \frac{(M-N-1)\hat{\psi}_t^2 - (N-1)}{M} + \frac{2(\hat{\psi}_t^2)^{\frac{N-1}{2}}(1+\hat{\psi}_t^2)^{-\frac{M-2}{2}}}{M B_{\hat{\psi}_t^2/(1+\hat{\psi}_t^2)}\left(\frac{N-1}{2}, \frac{M-N+1}{2}\right)} \tag{4.33}$$

being an adjusted estimator of

$$\hat{\psi}_t^2 = (\hat{\mu}_t - \hat{\mu}_{p,t}^{\text{MinV}}\mathbf{1}_N)^T\hat{\Sigma}_t^{-1}(\hat{\mu}_t - \hat{\mu}_{p,t}^{\text{MinV}}\mathbf{1}_N). \tag{4.34}$$

The term $B_{\hat{\psi}_t^2/(1+\hat{\psi}_t^2)}\left(\frac{N-1}{2}, \frac{M-N+1}{2}\right)$ refers to the incomplete beta function

$$B_x(a,b) = \int_0^x y^{a-1}(1-y)^{b-1}dy. \tag{4.35}$$

The first fraction of the adjusted estimator $\hat{\psi}_{\text{adj.},t}^2$ represents the unbiased estimator of $\hat{\psi}_t^2$ and the second fraction an adjustment in case the unbiased estimator is too small. The resulting portfolio can thus be characterized by

$$\hat{x}_t^{\text{MV/MinV}} = \frac{1}{\gamma}\frac{(M-N-1)(M-N-4)}{M(M-2)}\left(\left(\frac{\hat{\psi}_{\text{adj.},t}^2}{\hat{\psi}_{\text{adj.},t}^2 + \frac{N}{M}}\right)\hat{\Sigma}_t^{-1}\hat{\mu}_t + \left(\frac{\frac{N}{M}}{\hat{\psi}_{\text{adj.},t}^2 + \frac{N}{M}}\right)\hat{\mu}_{p,t}^{\text{MinV}}\hat{\Sigma}_t^{-1}\mathbf{1}_N\right). \tag{4.36}$$

The relative vector of portfolio weights takes the form

$$\begin{aligned}\hat{\omega}_t^{\text{MV/MinV}} &= \frac{\hat{x}_t^{\text{MV/MinV}}}{\left|\mathbf{1}_N^T\hat{x}_t^{\text{MV/MinV}}\right|} \\ &= \frac{\frac{1}{\gamma}\frac{(M-N-1)(M-N-4)}{M(M-2)}\left(\left(\frac{\hat{\psi}_{\text{adj.},t}^2}{\hat{\psi}_{\text{adj.},t}^2+\frac{N}{M}}\right)\hat{\Sigma}_t^{-1}\hat{\mu}_t + \left(\frac{\frac{N}{M}}{\hat{\psi}_{\text{adj.},t}^2+\frac{N}{M}}\right)\hat{\mu}_{p,t}^{\text{MinV}}\hat{\Sigma}_t^{-1}\mathbf{1}_N\right)}{\left|\mathbf{1}_N^T\frac{1}{\gamma}\frac{(M-N-1)(M-N-4)}{M(M-2)}\left(\left(\frac{\hat{\psi}_{\text{adj.},t}^2}{\hat{\psi}_{\text{adj.},t}^2+\frac{N}{M}}\right)\hat{\Sigma}_t^{-1}\hat{\mu}_t + \left(\frac{\frac{N}{M}}{\hat{\psi}_{\text{adj.},t}^2+\frac{N}{M}}\right)\hat{\mu}_{p,t}^{\text{MinV}}\hat{\Sigma}_t^{-1}\mathbf{1}_N\right)\right|} \\ &= \frac{\left(\frac{\hat{\psi}_{\text{adj.},t}^2}{\hat{\psi}_{\text{adj.},t}^2+\frac{N}{M}}\right)\hat{\Sigma}_t^{-1}\hat{\mu}_t + \left(\frac{\frac{N}{M}}{\hat{\psi}_{\text{adj.},t}^2+\frac{N}{M}}\right)\hat{\mu}_{p,t}^{\text{MinV}}\hat{\Sigma}_t^{-1}\mathbf{1}_N}{\left|\mathbf{1}_N^T\left(\left(\frac{\hat{\psi}_{\text{adj.},t}^2}{\hat{\psi}_{\text{adj.},t}^2+\frac{N}{M}}\right)\hat{\Sigma}_t^{-1}\hat{\mu}_t + \left(\frac{\frac{N}{M}}{\hat{\psi}_{\text{adj.},t}^2+\frac{N}{M}}\right)\hat{\mu}_{p,t}^{\text{MinV}}\hat{\Sigma}_t^{-1}\mathbf{1}_N\right)\right|}\end{aligned} \tag{4.37}$$

Intuitively, the weight put on the minimum-variance portfolio is an increasing function in the number of assets and a decreasing function in the number of observations as both scenarios tend to increase the estimation error when calculating the efficient portfolio. A higher risk of estimation error leads to a greater reliance on the minimum-variance portfolio that can be calculated with a relatively higher accuracy. However, it is important to note that the determination of the weights $\hat{c}_t$ and $\hat{d}_t$ regarding the optimal combination of the optimal portfolios is also exposed to estimation risk. In fact, as long as the process of optimally combining already optimal strategies introduces an additional level of estimation risk, it is not ex ante clear that these potential errors are completely offset by the theoretically predicted diversification gains.

4.4.9. Equally-Weighted and Minimum-Variance

This strategy as proposed in DeMiguel et al. (2009b) follows the idea of Kan and Zhou (2007). However, instead of optimally combining the mean-variance portfolio strategy with the minimum-variance equivalent, DeMiguel et al. (2009b) focus on the optimal combination of the naive diversification approach and the minimum-variance strategy (EW/MinV). Since the former does not rely on any estimation procedure at all and the minimum-variance strategy only deals with uncertainty regarding the estimation of the variance-covariance matrix, the strategy as suggested by DeMiguel et al. (2009b) is supposed to be exposed to only very limited estimation risk. The relative vector of portfolio weights is given by

$$\hat{\omega}_t^{\text{EW/MinV}} = \hat{c}_t \frac{1}{N} \mathbf{1}_N + \hat{d}_t \hat{\Sigma}_t^{-1} \mathbf{1}_N \tag{4.38}$$

subject to $\mathbf{1}_N^T \hat{\omega}_t^{\text{EW/MinV}} = 1$. The individual weights $\hat{c}_t$ and $\hat{d}_t$ are chosen so as to maximize the expected out-of-sample performance with respect to the utility of a mean-variance investor. It is

$$\hat{c}_t = 1 - \hat{d}_t \mathbf{1}_N^T \hat{\Sigma}_t^{-1} \mathbf{1}_N \tag{4.39}$$

and

$$\hat{d}_t = \frac{(M-N-2)(\mathbf{1}_N^T \hat{\Sigma}_t \mathbf{1}_N)(\mathbf{1}_N^T \hat{\Sigma}_t^{-1} \mathbf{1}_N) - N^2 M}{N^2(M-N-2)k(\mathbf{1}_N^T \hat{\Sigma}_t^{-1} \mathbf{1}_N) - 2MN^2(\mathbf{1}_N^T \hat{\Sigma}_t^{-1} \mathbf{1}_N) + (M-N-2)(\mathbf{1}_N^T \hat{\Sigma}_t^{-1} \mathbf{1}_N)^2(\mathbf{1}_N^T \hat{\Sigma}_t \mathbf{1}_N)} \tag{4.40}$$

with $k = \frac{M^2(M-2)}{(M-N-1)(M-N-2)(M-N-4)}$.

4.4.10. Equally-Weighted and Mean-Variance

In the spirit of Kan and Zhou (2007) and DeMiguel et al. (2009b), Tu and Zhou (2011) propose a new set of combinations of optimal portfolio strategies. First, the naive diversification approach is combined with the mean-variance strategy (EW/MV). The optimal combination is determined by the coefficient δ that is chosen so as to maximize the expected out-of-sample performance with respect to the utility of a mean-variance investor, i.e.

$$\hat{x}_t^{\text{EW/MV}} = (1 - \hat{\delta}_t)\hat{\omega}_t^{\text{EW}} + \hat{\delta}\hat{\omega}_t^{\text{MV}}. \tag{4.41}$$

The combination of the mean-variance strategy with the equally-weighted approach can be understood as a shrinkage toward naive diversification resulting in a trade-off between bias and estimation precision. In this context, the shrinkage toward the naive diversification approach will bias the plain mean-variance solution but render the estimation more precise. Since the optimal combination parameter δ is also subject to estimation risk, it is ex ante not clear, that the resulting out-of-sample excess return performance will be indeed superior to either the naive diversification approach or the mean-variance strategy. Tu and Zhou (2011), however, attribute a rather limited estimation risk to the determination of δ such that a superior performance of their strategy is expected. It is

$$\hat{\delta}_t = \frac{\hat{\pi}_{1,t}}{\hat{\pi}_{1,t} + \hat{\pi}_{2,t}} \tag{4.42}$$

with

$$\hat{\pi}_{1,t} = (\hat{\omega}_t^{\text{EW}})^T \hat{\Sigma}_t \hat{\omega}_t^{\text{EW}} - \frac{2}{\gamma}(\hat{\omega}_t^{\text{EW}})^T \hat{\mu}_t + \frac{1}{\gamma^2}\hat{\theta}^2_{\text{adj.},t} \tag{4.43}$$

and

$$\hat{\pi}_{2,t} = \frac{1}{\gamma^2}(c_1 - 1)\hat{\theta}^2_{\text{adj.},t} + \frac{c_1}{\gamma^2}\frac{N}{M}, \tag{4.44}$$

with $c_1 = \frac{(M-2)(M-N-2)}{(M-N-1)(M-N-4)}$, where $\hat{\theta}^2_{\text{adj.},t}$ is an adjusted estimator of $\hat{\theta}^2_t = \hat{\mu}_t^T \hat{\Sigma}_t^{-1} \hat{\mu}_t$ satisfying

$$\hat{\theta}^2_{\text{adj.},t} = \frac{(M-N-2)\hat{\theta}^2_t - N}{M} + \frac{2(\hat{\theta}^2_t)^{\frac{N}{2}}(1+\hat{\theta}^2_t)^{-\frac{M-2}{2}}}{M B_{\hat{\theta}^2_t/(1+\hat{\theta}^2_t)}\left(\frac{N}{2}, \frac{M-N}{2}\right)}. \tag{4.45}$$

The incomplete beta function $B_{\hat{\theta}^2_t/(1+\hat{\theta}^2_t)}\left(\frac{N}{2}, \frac{M-N}{2}\right)$ is defined as in Eq. (4.35). The condition $T > N+4$ needs to be fulfilled to ensure the existence of the second moment of the inverse of the variance-covariance matrix. The relative vector of portfolio weights is given by

$$\hat{\omega}_t^{\text{EW/MV}} = \frac{\hat{x}_t^{\text{EW/MV}}}{\left|\mathbf{1}_N^T \hat{x}_t^{\text{EW/MV}}\right|} = \frac{(1-\hat{\delta}_t)\hat{\omega}_t^{\text{EW}} + \hat{\delta}\hat{\omega}_t^{\text{MV}}}{\left|\mathbf{1}_N^T\left((1-\hat{\delta}_t)\hat{\omega}_t^{\text{EW}} + \hat{\delta}\hat{\omega}_t^{\text{MV}}\right)\right|}. \tag{4.46}$$

4.4.11. Equally-Weighted and Mean-Variance and Minimum-Variance

A second combination of optimal portfolio strategies proposed by Tu and Zhou (2011) is the extension of the already optimally combined MV/MinV strategy by Kan and Zhou (2007) to the inclusion of the naive diversification approach (EW/MV/MinV). To be precise, it is

$$\hat{x}_t^{\text{EW/MV/MinV}} = (1-\hat{\delta}_t)\hat{\omega}_t^{\text{EW}} + \hat{\delta}_t\hat{\omega}_t^{\text{MV/MinV}}, \tag{4.47}$$

where $\hat{\delta}_t$ is again chosen so as to maximize the expected out-of-sample performance with respect to the utility of a mean-variance investor and takes the form

$$\hat{\delta}_t = \frac{\hat{\pi}_{1,t} - \hat{\pi}_{13,t}}{\hat{\pi}_{1,t} - 2\hat{\pi}_{13,t} + \hat{\pi}_{3,t}}. \tag{4.48}$$

It is $\hat{\pi}_{1,t}$ as in Eq. (4.43) and

$$\begin{aligned}\hat{\pi}_{13,t} = {} & \frac{1}{\gamma^2}\hat{\theta}^2_{\text{adj.},t} - \frac{1}{\gamma}(\hat{\omega}_t^{\text{EW}})^T\hat{\mu}_t \\ & + \frac{1}{\gamma c_1}\Big(\left(\hat{\psi}^2_{\text{adj.},t}(\hat{\omega}_t^{\text{EW}})^T\hat{\mu}_t + (1-\hat{\psi}^2_{\text{adj.},t})\hat{\mu}_{p,t}^{\text{MinV}}(\hat{\omega}_t^{\text{EW}})^T\mathbf{1}_N\right) \\ & - \frac{1}{\gamma}\left(\hat{\psi}^2_{\text{adj.},t}\hat{\mu}_t^T\hat{\Sigma}_t^{-1}\hat{\mu}_t + (1-\hat{\psi}^2_{\text{adj.},t})\hat{\mu}_{p,t}^{\text{MinV}}\hat{\mu}_t^T\hat{\Sigma}_t^{-1}\mathbf{1}_N\right)\Big)\end{aligned} \tag{4.49}$$

and

$$\hat{\pi}_{3,t} = \frac{1}{\gamma^2}\hat{\theta}^2_{\text{adj.},t} - \frac{1}{\gamma^2 c_1}\left(\hat{\theta}^2_{\text{adj.},t} - \frac{N}{M}\hat{\psi}^2_{\text{adj.},t}\right) \tag{4.50}$$

with $\hat{\psi}_{\text{adj.},t}$ as in Eq. (4.33), $\hat{\theta}_{\text{adj.},t}$ as in Eq. (4.45), and $c_1 = \frac{(M-2)(M-N-2)}{(M-N-1)(M-N-4)}$. The relative vector of portfolio weights is given by

$$\hat{\omega}_t^{\text{EW/MV/MinV}} = \frac{\hat{x}_t^{\text{EW/MV/MinV}}}{\left|\mathbf{1}_N^T \hat{x}_t^{\text{EW/MV/MinV}}\right|} = \frac{(1-\hat{\delta}_t)\hat{\omega}_t^{\text{EW}} + \hat{\delta}\hat{\omega}_t^{\text{MV/MinV}}}{\left|\mathbf{1}_N^T\left((1-\hat{\delta}_t)\hat{\omega}_t^{\text{EW}} + \hat{\delta}\hat{\omega}_t^{\text{MV/MinV}}\right)\right|}. \tag{4.51}$$

4.4.12. Volatility Timing

The volatility timing (VT) strategy by Kirby and Ostdiek (2012) can be stated as

$$\hat{\omega}_{it}^{\text{VT}} = \frac{(1/\hat{\sigma}_{it}^2)^\eta}{\sum\limits_{i=1}^{N}(1/\hat{\sigma}_{it}^2)^\eta} \quad \text{with} \quad i = 1, \ldots, N, \tag{4.52}$$

where $\hat{\sigma}_{it}$ is the sample volatility of the ith excess risky return and $\eta \geq 0$. This approach coincides with the solution of the mean-variance strategy when ignoring expected returns and under the assumption of a diagonal variance-covariance matrix. Assuming a diagonal variance-covariance matrix, however, is equivalent to assuming all the pairwise correlation coefficients of risky asset to be zero and thus representing an extreme form of shrinking the variance-covariance matrix to a diagonal. The potential bias that comes with this form of extreme shrinkage is opposed to the estimation precision improvement as there are $N(N-1)/2$ parameters less to estimate. The assumption, as usual, is that gains in estimation precision outweigh the incurred bias. The strategy of Kirby and Ostdiek (2012) comes with some more advantages. In fact, the weights proposed in Eq. (4.52) neither result in negative weights nor do they rely on the inversion of the variance-covariance matrix, two characteristics that potentially lead to a higher estimation accuracy. The parameter η can be understood as a tuning parameter that decides on the sensitivity of the portfolio weights with respect to changes in volatility. In the limit of $\eta \to 0$, the portfolio weights coincide with the weights of the naive diversification approach. The benchmark case representing the mean-variance approach under the assumption of a diagonal variance-covariance matrix is $\eta = 1$.

4.4.13. Reward-to-Risk Timing

This strategy is also proposed by Kirby and Ostdiek (2012) and related to their VT strategy. However, the reward-to-risk timing (RRT) strategy also accounts for expected returns, i.e.

$$\hat{\omega}_{it}^{\text{RRT}} = \frac{\left(\max(\hat{\mu}_{it}, 0)/\hat{\sigma}_{it}^2\right)^\eta}{\sum\limits_{i=1}^{N}\left(\max(\hat{\mu}_{it}, 0)/\hat{\sigma}_{it}^2\right)^\eta} \quad \text{with} \quad i = 1, \ldots, N. \tag{4.53}$$

In this specific context, the expected returns are restricted to be positive, i.e. the representative investor does not invest in assets that yield a negative expected return. This constraint is chosen so as to ensure a high degree of comparability between the VT and

RRT strategy. Both strategies result in positive weight vectors and hence reduce the risk of adopting extreme positions while mitigating the risk of high estimation errors at the same time.

5. Methodology

This section is dedicated to the methodology used to perform the horse race between the different portfolio strategies. In fact, the data sets employed in this analysis are divided into an in-sample period used for estimating the data generating parameters and an out-of-sample period for which monthly excess returns are determined. Subsequently, these out-of-sample return forecasts are evaluated based on different performance evaluation criteria. To start with, the following describes the actual estimation procedure based on a rolling-window approach. Next, the employed performance evaluation criteria are introduced and discussed.

5.1. Estimation Procedure

The estimation procedure of the out-of-sample forecasting is based on a rolling window approach and can be described as a four-step procedure. First, the length of the estimation window M needs to be defined. Here, the estimation window consists of $M = 120$ consecutive monthly excess return observations corresponding to 10 years of return data. Second, starting at time $t = M$, all parameters required for the calculation of the optimal portfolio weights $\hat{\omega}_t^i$ at time t for strategy i, $i = \text{EW}, \text{MV}, \ldots, \text{RRT}$, are estimated based on the M previous months, i.e. $[t - M + 1, \ldots, t]$. Typically, these parameters are the sample mean and the variance-covariance matrix. Third, the optimal portfolio weights $\hat{\omega}_t^i$ are obtained and used to calculate the corresponding estimated out-of-sample portfolio returns for time $t + 1$. Fourth, the estimation window is shifted ahead one period by including the next return observation corresponding to date $t + 1$ and dropping the earliest return so as to keep the length of the estimation window constant, i.e. $[t - M + 2, \ldots, t + 1]$. Finally, the process continues with step two and rolls through the data until the end of the return series is reached. Given a time series of size T, this approach results in an estimated out-of-sample return series of length $T - M$.

The rolling-window approach has several degrees of freedom that can potentially affect the final results. It is ex ante not evident, for example, how to optimally choose the length of the estimation window. A shorter estimation window might lead to inaccurate estimations of the data generating parameters, whereas a longer estimation window might be insensitive to potential structural breaks, i.e. time variations in the sample moments. The main motivation for applying the rolling window approach of size $M = 120$ is given by the fact that this appears to be the working horse of the related empirical literature and hence preserves comparability. Shorter and longer estimation windows are included in the robustness checks.

5.2. Performance Evaluation Criteria

Several different performance evaluation criteria are employed in this analysis to assess the actual performance of the out-of-sample monthly excess return series obtained from the rolling window approach with respect to the optimized portfolio strategies introduced in Section 4.4. The majority of performance criteria fall into the classification of reward-to-risk ratios of which the Sharpe ratio is arguably the most famous and widely used. However, out-of-sample return forecasts typically result in time series that are not normally distributed. For this reason, this analysis also considers reward-to-risk ratios that do not critically rely on this normality assumption of asset returns. In fact, whereas the reward measures throughout all ratios exclusively focus on the out-of-sample mean returns, the risk measures incorporate several competing approaches and hence create different risk-adjusted performance measures. Moreover, the focus is set on reward-to-risk ratios that do not excessively rely on the estimation of additional quantities which might further question the reliability of the obtained results such as for example the Treynor ratio. Following Frahm et al. (2012), this analysis does also not consider turnover which empirically is a common performance measure providing information about the transaction costs of realizing the respective portfolios in every period. The reason for not including this measurement lies in the fact that the setting of this analysis is a rather static one focusing on a repeated single period portfolio choice problem and thus does not actually incorporate the very nature of portfolio rebalancing. Instead, the certainty equivalent return is included which provides information about the risk-free return a representative investor demands to be indifferent between this risk-free return and the corresponding returns resulting from the respective different portfolio strategies.

5.2.1. Sharpe Ratio

Introduced in Sharpe (1966) and generalized in Sharpe (1994), the out-of-sample Sharpe ratio (SR) of portfolio strategy i, $i = \text{EW}, \text{MV}, \ldots, \text{RRT}$, is defined as

$$\widehat{\text{SR}}^i = \frac{E(\hat{R}_p^i) - r_f}{\hat{\sigma}_p^i} = \frac{\hat{\mu}_p^i}{\hat{\sigma}_p^i}. \tag{5.1}$$

The numerator yields the difference between the expected out-of-sample nonexcess mean portfolio returns for strategy i given by $E(\hat{R}_p^i)$ and the risk-free rate r_f, i.e. $\hat{\mu}_p^i = E(\hat{R}_p^i) - r_f$, or rather

$$\hat{\mu}_p^i = \frac{1}{T-M} \sum_{t=M+1}^{T} \hat{\mu}_t^i, \tag{5.2}$$

where $\hat{\mu}_p^i$ denotes the expected out-of-sample excess mean portfolio return. The denominator contains the corresponding sample portfolio standard deviation

$$\hat{\sigma}_p^i = \sqrt{\frac{1}{T-M-1}\sum_{t=M+1}^{T}(\hat{\mu}_t^i - \hat{\mu}_p^i)^2}. \tag{5.3}$$

A higher positive Sharpe ratio is preferred over a lower positive Sharpe ratio as the former yields a more favorable return-to-risk trade-off. Negative Sharpe ratios indicate an outperformance of the portfolio strategy by the risk-free asset and thus represent an undesired event. As a further, however, not attainable benchmark, the in-sample (IS) Sharpe ratio based on the in-sample mean and variance-covariance matrix, i.e. $\hat{\mu}_{\text{IS}}$ and $\hat{\Sigma}_{\text{IS}}$, is calculated as

$$\widehat{\text{SR}}_{\text{IS}}^{\text{MV}} = \frac{\hat{\mu}_{p,\text{IS}}^{\text{MV}}}{\hat{\sigma}_{p,\text{IS}}^{\text{MV}}} = \frac{\hat{\mu}_{\text{IS}}^T \hat{\omega}_{\text{IS}}^{\text{MV}}}{\sqrt{(\hat{\omega}_{\text{IS}}^{\text{MV}})^T \hat{\Sigma}_{\text{IS}} \hat{\omega}_{\text{IS}}^{\text{MV}}}}. \tag{5.4}$$

The in-sample weights are denoted by $\hat{\omega}_{\text{IS}}^{\text{MV}}$ based on a time period of $[M+1,\dots,T]$, i.e. the originally out-of-sample period is now treated as if it were known in advance and hence optimization on this period results in the highest possible Sharpe ratio. In the same manner, the remaining performance evaluation criteria are evaluated with respect to their in-sample performance.

The Sharpe ratio is also not free from criticism. First, it assumes normally distributed returns. Hence, time series of returns that violate this assumption might lead to biased estimates of Sharpe ratios. Second, risk is represented by the standard deviation, i.e. volatility. Volatility itself, however, is not necessarily an unwanted characteristic but rather a positive feature when restricted on its upside potential. The price function of a long call option, for example, positively depends on the volatility of the underlying market and illustrates the beneficial properties that volatility can have. Still, its conceptual simplicity and widespread use render the Sharpe ratio a valuable tool for evaluating risk-adjusted portfolio performance.

To finally answer the question of whether there are optimized portfolio strategies that can consistently outperform the naive diversification approach, it is necessary to establish a statistically informed decision criterion. As a first step toward this direction and in accordance with DeMiguel et al. (2009b), the test statistic proposed by Jobson and Korkie (1981) (JK) adjusted for the correction suggested by Memmel (2003) is applied to decide on the pairwise difference between the out-of-sample Sharpe ratio of the naive diversification approach, SR^{EW}, and strategy i, SR^i, with $i = \text{MV}, \text{MinV}, \dots, \text{RRT}$. It is

$$\widehat{\text{SR}}^{\text{EW}} - \widehat{\text{SR}}^i \sim \mathcal{N}(0, \hat{\Theta}) \tag{5.5}$$

with

$$\hat{\Theta} = \frac{1}{T-M}\left(2\left(1-(\rho^{\text{EW},i})^2\right) + \frac{1}{2}\left((\widehat{\text{SR}}^{\text{EW}})^2 + (\widehat{\text{SR}}^i)^2 - 2\widehat{\text{SR}}^{\text{EW}}\widehat{\text{SR}}^i(\rho^{\text{EW},i})^2\right)\right). \tag{5.6}$$

The out-of-sample correlation coefficient of the two portfolios under consideration is denoted by $\rho^{\text{EW},i}$. The null hypotheses of the one-sided tests of the form $H_0^i : \widehat{\text{SR}}^{\text{EW}} - \widehat{\text{SR}}^i \geq 0$ are then decided on the test statistics

$$\hat{z}^i_{\text{JK}} = \frac{\widehat{\text{SR}}^{\text{EW}} - \widehat{\text{SR}}^i}{\sqrt{\hat{\Theta}}}. \tag{5.7}$$

These test statistics refer to a sample size of $T - M$. Given the assumption of independent and identically normally distributed out-of-sample portfolio returns, the test statistics $\hat{z}^i_{\text{JK}}$ are asymptotically distributed as a standard normal. However, out-of-sample portfolio returns are at best approximately normally distributed but in general rather display fat tails and resemble a leptokurtic distribution. This, in turn, might lead to wrong inferences with respect to p-values. To account for this drawback, this analysis also considers a studentized time series bootstrap approach based on the Delta method and as introduced by Ledoit and Wolf (2008)[4] (LW) that is independent of the normality assumption. Let

$$f(v) = \text{SR}^{\text{EW}} - \text{SR}^i \tag{5.8}$$

denote the difference between the Sharpe ratios of the naive diversification approach and the optimized strategy i with $v = (\mu^{\text{EW}}, \mu^i, \gamma^{\text{EW}}, \gamma^i)$ and the raw variances $\gamma^{\text{EW}} = E((r^{\text{EW}})^2)$ and $\gamma^i = E((r^i)^2)$. It follows that

$$f(\mu^{\text{EW}}, \mu^i, \gamma^{\text{EW}}, \gamma^i) = \frac{\mu^{\text{EW}}}{\sqrt{\gamma^{\text{EW}} - (\mu^{\text{EW}})^2}} - \frac{\mu^i}{\sqrt{\gamma^i - (\mu^i)^2}}. \tag{5.9}$$

According to Ledoit and Wolf (2008) and the Delta method, this results in

$$\sqrt{T - M}(\widehat{\text{SR}}^{\text{EW}} - \widehat{\text{SR}}^i - (\text{SR}^{\text{EW}} - \text{SR}^i)) \overset{d}{\sim} \mathcal{N}(0, \nabla^T f(v) \Psi \nabla f(v)) \tag{5.10}$$

with

$$\nabla^T f(v) = \left(\frac{\gamma^{\text{EW}}}{(\gamma^{\text{EW}} - (\mu^{\text{EW}})^2)^{3/2}}, -\frac{\gamma^i}{(\gamma^i - (\mu^i)^2)^{3/2}}, -\frac{1}{2} \frac{\mu^{\text{EW}}}{(\gamma^{\text{EW}} - (\mu^{\text{EW}})^2)^{3/2}}, \frac{1}{2} \frac{\mu^i}{(\gamma^i - (\mu^i)^2)^{3/2}} \right). \tag{5.11}$$

The null hypotheses of the two-sided tests of the form $H_0^i : \widehat{\text{SR}}^{\text{EW}} - \widehat{\text{SR}}^i = 0$ are then decided based on the test statistics

$$\hat{z}^i_{\text{LW}} = \frac{\widehat{\text{SR}}^{\text{EW}} - \widehat{\text{SR}}^i - (\text{SR}^{\text{EW}} - \text{SR}^i)}{\sqrt{\nabla^T f(\hat{v}) \hat{\Psi} \nabla f(\hat{v}) / (T - M)}}. \tag{5.12}$$

The estimate of the matrix Ψ is obtained via a heteroskedasticity and autocorrelation (HAC) robust kernel estimation and the p-values are calculated as in Ledoit and Wolf (2008) with a block length of $b = 5$ and $K = 1{,}000$ bootstrap iterations.

[4]Code is freely available at `https://www.econ.uzh.ch/en/people/faculty/wolf/publications.html`.

5.2.2. Sortino Ratio

Since individuals are typically loss-averse but not gain-averse, a downside risk measure could be more appropriate in capturing the investor's fundamental trade-off between return and risk. Originated from Bawa (1975) and Fishburn (1977), lower partial moments (LPM) only focus on the negative deviations of returns, i.e. the downside risk, with respect to a minimum acceptable return (MAR) r^{MAR}. The lower partial moment of order n for strategy i is

$$\widehat{\text{LPM}}_n^i(r^{\text{MAR}}) = \frac{1}{T-M} \sum_{t=M+1}^{T} \max(r^{\text{MAR}} - \hat{\mu}_t^i, 0)^n. \tag{5.13}$$

The order n determines the weighting of the return's deviation from the minimum acceptable return and translates into a representation of a risk-averse investor for values larger than one. In this analysis, the minimum acceptable return is set to zero. However, due to the fact that $\hat{\mu}_t^i$ refers to excess returns, it is possible to equivalently state the lower partial moment in terms of nonexcess returns and a minimum acceptable return equal to the risk-free rate. The Sortino ratio (SoR) is obtained by substituting the standard deviation in the denominator of the Sharpe ratio with the lower partial moment of order two. Based on Bawa (1978), Harlow (1991), and Sortino and Van Der Meer (1991), the Sortino ratio emerges as

$$\widehat{\text{SoR}}^i = \frac{\hat{\mu}_p^i}{\sqrt{\frac{1}{T-M} \sum_{t=M+1}^{T} \max(r^{\text{MAR}} - \hat{\mu}_t^i, 0)^2}}. \tag{5.14}$$

5.2.3. Omega Ratio

The Omega ratio (OR) is a measure that does not require any assumption on the underlying distribution of the specific time series and hence can be referred to as a nonparametric ratio. In addition, the Omega ratio considers all moments of the return distribution as opposed to the Sharpe ratio that only incorporates the first two moments. Established in Keating and Shadwick (2002), the Omega ratio as formulated in Kaplan and Knowles (2004) is

$$\widehat{\text{OR}}^i = \frac{\hat{\mu}_p^i}{\frac{1}{T-M} \sum_{t=M+1}^{T} \max(r^{\text{MAR}} - \hat{\mu}_t^i, 0)} + 1. \tag{5.15}$$

A higher Omega ratio is preferred over a lower Omega ratio and, for example, a return series yielding a higher excess kurtosis has ceteris paribus a lower Omega ratio relative to a return series with a lower excess kurtosis.

5.2.4. Calmar Ratio

The Calmar ratio (CR) is a somewhat more extreme risk-adjusted performance evaluation measure as it only considers the most adverse development of the out-of-sample return series. Nevertheless, investors might be more concerned about a single severe loss relative to many small losses. The Calmar ratio developed by Young (1991) is the expected excess out-of-sample mean portfolio return over the maximum drawdown (MaxDD) for strategy i, i.e.

$$\widehat{\mathrm{CR}}^i = \frac{\hat{\mu}_p^i}{-\widehat{\mathrm{MaxDD}}_p^i}. \tag{5.16}$$

The drawdown (DD) for strategy i at time t is the percentage loss of the portfolio's value since the previous peak until time t, i.e.

$$\widehat{\mathrm{DD}}_{p,t}^i = \max\left(0, \max_{\tau\in[M+1,t]}\left(\frac{\hat{p}_\tau^i - \hat{p}_t^i}{\hat{p}_\tau^i}\right)\right). \tag{5.17}$$

The running maximum value of the portfolio in the time interval $[M+1,t]$ is denoted by $\hat{p}_\tau^i$ and $\hat{p}_t^i$ is the actual value of the portfolio at time t. The maximum drawdown for strategy i is the largest loss that can occur during the entire out-of-sample period $[M+1,T]$ and hence is given by

$$\widehat{\mathrm{MaxDD}}_p^i = \max_{t\in[M+1,T]}(\widehat{\mathrm{DD}}_{p,t}^i) = \max_{t\in[M+1,T]}\left(0, \max_{\tau\in[M+1,t]}\left(\frac{\hat{p}_\tau^i - \hat{p}_t^i}{\hat{p}_\tau^i}\right)\right). \tag{5.18}$$

Similar to the other risk-adjusted performance measures, higher values of the Calmar ratio are preferred over lower values.

5.2.5. Return on Value-at-Risk

Value-at-risk (VaR) provides information about the maximum possible loss of an investment during the out-of-sample period that is not exceeded with a probability of $1-\alpha$. Formally, the value-at-risk can be defined as

$$\mathrm{Prob}(\Delta p \leq -\mathrm{VaR}_p) = 1-\alpha, \tag{5.19}$$

where p denotes the value of the portfolio. Under the assumption of normally distributed returns, the value-at-risk can be obtained by

$$\widehat{\mathrm{VaR}}_p^i = -(\hat{\mu}_p^i + z_\alpha \hat{\sigma}_p^i) \tag{5.20}$$

with z_α representing the α-quantile of the standard normal distribution. This analysis applies a value of $\alpha = 5\%$. The excess return on value-at-risk (RoVaR) as proposed by Dowd (2000) is defined as

$$\widehat{\mathrm{RoVaR}}^i = \frac{\hat{\mu}_p^i}{\widehat{\mathrm{VaR}}_p^i}. \tag{5.21}$$

The value-at-risk strategy is occasionally criticized for not being a coherent risk measure as defined in Artzner et al. (1999). In particular, the axiom of subadditivity is violated, i.e. the value-at-risk of two portfolios after they have been merged can exceed the sum of the individual portfolio's values-at-risk. In addition, the value-at-risk criterion does not provide any information about the potential loss that might occur beyond the cut-off threshold specified by α. The return on value-at-risk is included in this analysis due to its popular use in the financial industry with respect to risk management and regarding financial reporting and regulatory guidelines concerning a bank's required capital set by the Bank for International Settlements in their Basel Accords.

5.2.6. Certainty Equivalent Return

The certainty equivalent return (CE) is the risk-free return a representative investor demands to be indifferent between this risk-free return and the corresponding returns resulting from the different portfolio strategies, respectively. In this analysis, the certainty equivalent is defined as in DeMiguel et al. (2009b), i.e.

$$\widehat{\mathrm{CE}}^i = \hat{\mu}_p^i - \frac{\gamma}{2}(\hat{\sigma}_p^i)^2. \tag{5.22}$$

This actually represents the expected utility of a mean-variance investor with respect to Eq. (4.6). However, this expected utility can serve as an approximation for the certainty equivalent as shown in Christensen and Feltham (2003). In this context, the investor's risk-aversion parameter γ is set to one.

Similar to the Sharpe ratio and also in accordance with DeMiguel et al. (2009b), the Delta method is applied to decide on the pairwise difference between the out-of-sample certainty equivalent returns of the naive diversification approach, $\mathrm{CE}^{\mathrm{EW}}$, and strategy i, CE^i, with $i = \mathrm{MV}, \mathrm{MinV}, \ldots, \mathrm{RRT}$. Let

$$f(v) = \mathrm{CE}^{\mathrm{EW}} - \mathrm{CE}^i \tag{5.23}$$

with $v = (\mu^{\mathrm{EW}}, \mu^i, \gamma^{\mathrm{EW}}, \gamma^i)$ and the raw variances $\gamma^{\mathrm{EW}} = E((r^{\mathrm{EW}})^2)$ and $\gamma^i = E((r^i)^2)$. Under the assumption of independent and identically normally distributed returns, it is

$$\sqrt{T-M}\left(\widehat{\mathrm{CE}}^{\mathrm{EW}} - \widehat{\mathrm{CE}}^i - (\mathrm{CE}^{\mathrm{EW}} - \mathrm{CE}^i)\right) \overset{d}{\sim} \mathcal{N}(0, \nabla^T f(v)\Omega\nabla f(v)) \tag{5.24}$$

with

$$\nabla^T f(v) = \left(-(1+\gamma\mu^{\mathrm{EW}}), 1+\gamma\mu^i, \frac{\gamma}{2}, -\frac{\gamma}{2}\right) \tag{5.25}$$

and

$$\Omega = \begin{pmatrix} (\sigma^{\mathrm{EW}})^2 & \sigma^{\mathrm{EW},i} & 0 & 0 \\ \sigma^{i,\mathrm{EW}} & (\sigma^i)^2 & 0 & 0 \\ 0 & 0 & 2(\sigma^{\mathrm{EW}})^4 & 2(\sigma^{\mathrm{EW},i})^2 \\ 0 & 0 & 2(\sigma^{i,\mathrm{EW}})^2 & (\sigma^i)^4 \end{pmatrix}. \tag{5.26}$$

The null hypotheses of the one-sided tests of the form $H_0^i : \widehat{\mathrm{CE}}^{\mathrm{EW}} - \widehat{\mathrm{CE}}^i \geq 0$ are then decided on the test statistics

$$\hat{z}_{\Delta}^{i} = \frac{\widehat{\text{CE}}^{\text{EW}} - \widehat{\text{CE}}^{i} - (\text{CE}^{\text{EW}} - \text{CE}^{i})}{\sqrt{\nabla^{T} f(\hat{v})\hat{\Omega}\nabla f(\hat{v})/(T-M)}}. \tag{5.27}$$

These test statistics, however, also critically depend on the assumption of having normally distributed out-of-sample excess returns. To account for this drawback, the Sharpe ratio bootstrap methodology introduced by Ledoit and Wolf (2008) is adjusted for a certainty equivalent return setting. Hence, the null hypotheses of the two-sided tests of the form $H_0^i : \widehat{\text{CE}}^{\text{EW}} - \widehat{\text{CE}}^{i} = 0$ are decided based on the test statistics given in Eq. (5.27) but the matrix $\hat{\Omega}$ is substituted with the HAC robust matrix $\hat{\Psi}$. The p-values are calculated as in Ledoit and Wolf (2008) with a block length of $b = 5$ and $K = 1,000$ bootstrap iterations.

Finally, it is important to note that the underlying mean-variance framework of this analysis favors the Sharpe ratio as well as the certainty equivalent return relative to the other risk-adjusted performance measures. This is due to the fact that the optimized portfolio strategies are all designed so as to favor expected return and dismiss risk measured in terms of variance of return. Since the other performance criteria measure risk not in terms of variance, the portfolio strategies considered in this analysis are strictly speaking not optimized with respect to these additional risk measures. In fact, if the representative investor would only care about risk, for example, measured as the value-at-risk then the underlying maximization problem should be altered such that the variance is substituted with the value-at-risk. An application of such a mean-VaR approach can be found in Fabozzi et al. (2010). Nevertheless, the evaluation of optimized portfolio strategies based on the mean-variance framework but assessed by performance evaluation criteria that not only account for risk measured in terms of variance of return can provide valuable insights into the strategies' performance behavior.

6. Empirical Results and Discussion

This analysis can be understood as a horse race between different out-of-sample optimized portfolio strategies taking place on venues represented by different data sets where the role of the jockeys is assigned to the different performance evaluation criteria. This section now presents the results of these competitions starting with the findings regarding the Sharpe ratio as displayed in Table 6.1.

A first observation refers to the consistently highest Sharpe ratios of the MV IS strategy across all data sets indicated by the ranking -1-. This insight, however, comes at no surprise and is by construction since the MV IS optimization relies on future information which in reality is not actually available by the time of portfolio creation. Still, it defines an upper bound for the maximum attainable Sharpe ratio in case of perfect foresight and hence can be used to assess the severity of performance decrease based on this information asymmetry when compared to the MV strategy. Whereas the decrease in case of the

Table 6.1: Sharpe Ratios. This table lists the out-of-sample Sharpe ratios (SRs) of the naive diversification (EW) approach, the in-sample SRs for the mean-variance (MV IS) approach, and the out-of-sample SRs of the 15 optimized portfolio strategies for all six data sets under consideration. The first number in parentheses represents the p-value of the one-sided JK test on the pairwise difference between the out-of-sample SR of EW and strategy i, $i = \text{MV}, \text{MinV}, \ldots, \text{RRT}$. The second number in parentheses represents the p-value of the corresponding two-sided LW bootstrap test. SRs that outperform the EW benchmark at a significance level of $\alpha = 0.1$ with respect to LW are highlighted in bold. The numbers indicated by -·- denote the relative ranking of the specific strategy within each data set based on the value of the SR. All results rely on a rolling-window approach with window length $M = 120$, risk-aversion parameter $\gamma = 1$, and tuning parameter $\eta = 1$ (only VT and RRT).

Strategy	PF6 USA $N = 6$	PF25 USA $N = 25$	IND10 USA $N = 10$	IND48 USA $N = 48$	PF6 EU $N = 6$	PF25 EU $N = 25$
EW	0.2396 -12-	0.2386 -12-	0.2663 -9-	0.2379 -6-	0.1379 -14-	0.1461 -11-
MV IS	0.4460 -1-	0.6351 -1-	0.3399 -1-	0.4350 -1-	0.3896 -1-	0.6324 -1-
MV	**0.3912** (0.00) -6- (0.01)	**0.3942** (0.00) -7- (0.03)	0.2208 (0.17) -13- (0.44)	0.1037 (0.02) -14- (0.09)	0.1861 (0.23) -8- (0.58)	0.2227 (0.20) -4- (0.45)
MinV	**0.3964** (0.00) -4- (0.00)	**0.4302** (0.00) -5- (0.00)	0.3194 (0.08) -3- (0.21)	0.2078 (0.27) -8- (0.56)	**0.2219** (0.02) -2- (0.09)	0.1383 (0.44) -14- (0.85)
VW	0.1525 (0.00) -16- (0.00)	0.1525 (0.00) -17- (0.00)	0.1525 (0.00) -16- (0.00)	0.1525 (0.00) -11- (0.00)	0.0964 (0.00) -16- (0.00)	0.0964 (0.00) -17- (0.00)
BS	**0.4122** (0.00) -2- (0.00)	**0.4682** (0.00) -3- (0.00)	0.2827 (0.34) -8- (0.75)	0.1564 (0.08) -10- (0.27)	0.2040 (0.13) -3- (0.35)	0.2261 (0.18) -3- (0.41)
MVsc	0.2640 (0.08) -11- (0.21)	0.2658 (0.08) -10- (0.23)	0.1653 (0.00) -15- (0.03)	0.1342 (0.00) -12- (0.02)	0.1688 (0.06) -11- (0.28)	0.1627 (0.18) -9- (0.42)
MinVsc	0.2665 (0.07) -10- (0.13)	**0.2756** (0.03) -9- (0.05)	0.3091 (0.07) -4- (0.20)	0.2823 (0.08) -4- (0.17)	**0.1802** (0.01) -9- (0.08)	0.1803 (0.05) -8- (0.20)
BSsc	0.2683 (0.05) -9- (0.13)	0.2605 (0.11) -11- (0.33)	0.2010 (0.02) -14- (0.07)	0.1780 (0.04) -9- (0.13)	0.1700 (0.04) -10- (0.27)	0.1595 (0.23) -10- (0.50)
MVlsc	0.2793 (0.09) -8- (0.20)	0.2278 (0.42) -13- (0.84)	0.0953 (0.00) -17- (0.00)	0.0556 (0.00) -15- (0.02)	0.1874 (0.06) -7- (0.29)	0.2033 (0.16) -7- (0.43)
MP	0.2302 (0.00) -13- (0.03)	0.2270 (0.00) -14- (0.03)	0.2351 (0.00) -12- (0.03)	0.2233 (0.00) -7- (0.00)	0.1343 (0.00) -15- (0.03)	0.1420 (0.00) -12- (0.02)
MV/MinV	**0.4122** (0.00) -3- (0.00)	**0.4712** (0.00) -2- (0.00)	0.2630 (0.47) -10- (0.94)	0.0476 (0.00) -16- (0.01)	0.1946 (0.15) -6- (0.43)	0.2201 (0.18) -5- (0.40)
EW/MinV	**0.3826** (0.00) -7- (0.00)	**0.3938** (0.00) -8- (0.00)	**0.3215** (0.03) -2- (0.08)	0.2648 (0.13) -5- (0.27)	0.1988 (0.03) -4- (0.14)	0.1404 (0.42) -13- (0.83)
EW/MV	**0.3962** (0.00) -5- (0.00)	**0.4086** (0.00) -6- (0.00)	0.2623 (0.41) -11- (0.83)	0.1278 (0.00) -13- (0.00)	0.1951 (0.10) -5- (0.30)	0.2337 (0.10) -2- (0.24)
EW/MV/MinV	0.0162 (0.00) -17- (0.04)	**0.4501** (0.00) -4- (0.00)	0.2959 (0.17) -6- (0.41)	0.0026 (0.00) -17- (0.02)	−0.0816 (0.00) -17- (0.04)	0.2050 (0.26) -6- (0.56)
VT	0.2282 (0.01) -14- (0.10)	0.2102 (0.00) -15- (0.00)	**0.2972** (0.00) -5- (0.03)	**0.2886** (0.00) -2- (0.01)	0.1422 (0.27) -12- (0.56)	0.1287 (0.03) -15- (0.06)
RRT	0.2178 (0.00) -15- (0.02)	0.1934 (0.00) -16- (0.00)	**0.2862** (0.02) -7- (0.10)	**0.2879** (0.00) -3- (0.01)	0.1412 (0.32) -13- (0.68)	0.1211 (0.01) -16- (0.02)

PF6 USA portfolio is rather limited, performance drops are substantial for the remaining data sets. The Sharpe ratio decreases, for example, from 0.64 to 0.39 for PF25 USA. Already the comparison between MV IS and MV illustrates the losses in performance due to estimation errors and the resulting challenge of obtaining statistically significant out-of-sample outperformances relative to EW.

A second observation considers the limited number of optimized portfolio strategies that achieve to indeed significantly outperform EW. Out of the 90 listed optimized out-of-sample portfolio Sharpe ratios only 21 show a significant outperformance based on the two-sided p-value of the studentized bootstrap Ledoit and Wolf (2008) test for a significance level of $\alpha = 0.1$. Moreover, these few positive findings are neither attributable to a single optimized strategy nor to a specific data set. In fact, the results already now suggest that no optimized portfolio strategy consistently outperforms EW and confirms the findings of the related empirical literature which judges EW a tough benchmark to beat. It appears as if the inherent estimation risks completely outweigh theoretically proposed performance gains. In case of EW/MV/MinV and PF6 EU, a negative Sharpe ratio is obtained, i.e. the performance of this particular strategy is even inferior relative to an investment in the risk-free rate.

The findings of these first two observations are disappointing, yet not completely unexpected. The general performance behavior of the optimized strategies is very heterogeneous across data sets and deserves a closer examination. Although inferior to MV IS, the MV strategy substantially and significantly outperforms EW in case of PF6 and PF25 USA. Focusing on PF6 USA, MV outperforms EW by a factor of 1.6. Comparably high and significant performance levels are obtained for MinV, BS, MV/MinV, EW/MinV, and EW/MV. For PF25 USA, MV/MinV even outperforms EW by a factor of more than 2. In fact, all unconstrained optimized portfolio strategies apart from VW and MP appear to significantly outperform EW for PF6 and PF25 USA. No such outperformance can be found neither for the two US industry data sets nor for the two European portfolios. The VW strategies achieve identical levels of Sharpe ratios for the US and European data sets, respectively, since the value of the portfolios belonging to the same region is independent of the specific disposition of the portfolios. However, no superior performance is achieved across the data sets. VT and RRT are not explicitly designed as short sale constraint strategies but effectively are since they only attain positive weights. The two timing strategies yield significant results in case of both US industry portfolios, i.e. IND10 and IND48 USA. They are also the only strategies that consistently achieve a significantly higher Sharpe ratio for these two data sets. The good performance of the timing strategies with respect to the US industry portfolios can be related to specific characteristics inherent in these data sets. IND48, for example, exhibits the highest dispersion in mean and variance across the included assets of all data sets under consideration. Exactly these characteristics positively affect the performance of the timing strategies.

To shed some more light on the actual optimization process, Table 6.2 yields aver-

Table 6.2: Average Minimum and Maximum Sharpe Ratio Portfolio Weights. This table lists the average out-of-sample minimum and maximum portfolio weights for every strategy and all six data sets under consideration. This information is missing for the value-weighted (VW) strategy as only data about the final portfolio returns are available but no specific corresponding portfolio weights. The numbers are obtained by averaging the weights of every single asset over the out-of-sample estimation period. Presented, however, are only the minimum and maximum values of these averages. All results rely on a rolling-window approach with window length $M = 120$, risk-aversion parameter $\gamma = 1$, and tuning parameter $\eta = 1$ (only VT and RRT).

	PF6 USA $N = 6$		PF25 USA $N = 25$		IND10 USA $N = 10$		IND48 USA $N = 48$		PF6 EU $N = 6$		PF25 EU $N = 25$	
Strategy	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.
EW	0.17	0.17	0.04	0.04	0.10	0.10	0.02	0.02	0.17	0.17	0.04	0.04
MV	−2.02	2.32	−1.53	1.04	−0.44	0.68	−0.34	0.39	−1.93	2.46	−8.37	9.27
MinV	−1.02	1.17	−0.44	0.66	−0.40	0.44	−0.16	0.38	−0.53	1.44	−0.56	1.50
VW	-	-	-	-	-	-	-	-	-	-	-	-
BS	−1.62	1.86	−1.11	0.88	−0.43	0.45	−0.23	0.30	−1.15	2.00	−5.68	6.81
MVsc	0.00	0.62	0.00	0.25	0.00	0.36	0.00	0.21	0.00	0.75	0.00	0.34
MinVsc	0.00	0.38	0.00	0.24	0.00	0.40	0.00	0.34	0.00	0.58	0.00	0.45
BSsc	0.00	0.59	0.00	0.20	0.00	0.41	0.00	0.21	0.00	0.70	0.00	0.30
MVlsc	−0.30	0.97	−0.30	0.83	−0.28	0.59	−0.30	0.79	−0.30	0.96	−0.30	1.01
MP	0.13	0.22	0.03	0.06	0.05	0.14	0.01	0.03	0.15	0.19	0.03	0.05
MV/MinV	−1.49	1.70	−0.75	0.66	−0.46	0.55	−0.33	0.53	−0.89	2.06	−3.77	5.04
EW/MinV	−0.82	1.01	−0.30	0.48	−0.27	0.35	−0.06	0.19	−0.37	1.22	−0.30	0.86
EW/MV	−1.31	1.64	−0.61	0.39	−0.03	0.29	−0.15	0.23	−0.75	1.85	−2.99	3.44
EW/MV/MinV	−2.49	2.74	−0.63	0.59	−0.26	0.40	−0.26	0.18	−0.99	1.43	−6.06	7.49
VT	0.09	0.27	0.01	0.16	0.03	0.45	0.00	0.19	0.06	0.35	0.01	0.22
RRT	0.08	0.33	0.01	0.20	0.03	0.39	0.00	0.21	0.07	0.37	0.01	0.24

age minimum and maximum portfolio weights corresponding to the results regarding the Sharpe ratio. In case of EW, the lower and upper extreme portfolio weights coincide by construction and represent the equal investment across all available assets. The minimum and maximum weights of MP closely resemble the corresponding EW weights which also translates into similar Sharpe ratios. Estimation risk inherent in the process of determining the MP weights, however, renders the performance of MP to be inferior relative to EW.

A concern often raised in the context of portfolio optimization relates to extreme portfolio weight positions that are obtained thus exposing the optimization problem to more adverse effects of estimation errors. Michaud (1989) even argues that mean-variance optimization can be referred to as the act of error maximization since assets with a wrongly estimated high expected return and a wrongly estimated low variance tend to be assigned large portfolio weights and vice versa. Consequently, the resulting portfolios exhibit an upward bias with respect to the estimate of the portfolio mean and a downward bias with respect to the portfolio variance. However, portfolio weight positions in this analysis are limited. The most extreme value is attained for the PF25 EU data set with respect to MV and yields an average maximum long position in an asset of approximately 930%. Apart from the PF25 EU data set for which more extreme portfolio weight positions are realized, weights for the unconstrained optimized portfolio strategies hardly ever exceed the absolute value of 250%. The latter number might represent an extreme value in real-world trading but is rather limited regarding the theoretical portfolio optimization literature. An interesting comparison can be conducted by examining MV, MVsc, and MVlsc since these strategies represent three different levels of constraining portfolio weights. MV is not subject to any constraints and hence results in the most extreme portfolio weight positions also relative to all other strategies. MVsc prohibits short selling completely such that all portfolio weights are forced to adopt at least a zero holding position. Moreover, the additional restriction of the portfolio weights to sum to one effectively leads to a long sale constraint with threshold one. Indeed, MVsc as well as the other short sale constraint strategies only obtain weights between zero and one. In addition, lower average minimum weights frequently adopt a zero holding position, i.e. at least one asset is not included in the respective portfolio. MVlsc is a somewhat intermediate scenario that allows short sales but only up to a threshold of −30%. This lower bound is often attained allowing the respective average maximum return to exceed the boundary of 100% which, however, only happens in case of PF25 EU. Although MVlsc utilizes the possibility of realizing a specific level of short selling, the resulting performance with respect to Sharpe ratios is comparable. As far as the results of this analysis are concerned, short and long sale constraints limit the out-of-sample Sharpe ratio performance.

Results for the Sortino and Omega ratios are listed in Table 6.3. The major drawback of these performance evaluation criteria as well as for the Calmar ratio and the return on value-at-risk is the lack of an appropriate statistical test to differentiate between the EW benchmark performance and the respective optimized portfolio strategies. This issue is clearly under-researched. Nevertheless, valuable information can be extracted when comparing the results of the Sortino ratio to the Sharpe ratio. Apart from EW/MV/MinV in combination with PF6 EU, all Sortino ratios exceed the corresponding Sharpe ratios. This behavior is expected since the Sortino ratio only considers downside risk defined as negative deviations from the minimum acceptable return of zero. However, the Sortino ratios are less than twice as high as the corresponding Sharpe ratios as would be the case for normally distributed returns. This is again suggestive of the negative skewness inherent in the return data such that by considering only adverse deviations from the minimum acceptable return still more than half of the volatility for the whole distribution is captured. Optimized portfolio strategies that achieve good results for the Sharpe ratio also continue to exhibit the highest values in case of the Sortino ratio. Results for the Omega ratio are in line with previous findings. Since the Omega ratio is a nonparametric reward-to-risk ratio but achieves comparable outcomes, the loss of considering only the first two moments of the return distribution as for the Sharpe ratio does not seem to heavily affect performance behavior.

Table 6.4 presents the results for the Calmar ratios and the returns on value-at-risk. In

Table 6.3: Sortino and Omega Ratios. This table lists the out-of-sample Sortino ratios (SoRs) and Omega ratios (OR) of the naive diversification (EW) approach, the in-sample SoRs and ORs for the mean-variance (MV IS) approach, and the out-of-sample SoRs and ORs of the 15 optimized portfolio strategies for all six data sets under consideration. The numbers indicated by -·- denote the relative ranking of the specific strategy within each data set based on the value of the SoR and OR. All results rely on a rolling-window approach with window length $M = 120$, risk-aversion parameter $\gamma = 1$, and tuning parameter $\eta = 1$ (only VT and RRT).

Sortino Ratios						
Strategy	PF6 USA $N = 6$	PF25 USA $N = 25$	IND10 USA $N = 10$	IND48 USA $N = 48$	PF6 EU $N = 6$	PF25 EU $N = 25$
EW	0.3622 -12-	0.3614 -12-	0.4217 -10-	0.3679 -6-	0.1977 -14-	0.2094 -11-
MV IS	0.8556 -1-	1.5820 -1-	0.5943 -1-	0.9093 -1-	0.7433 -1-	1.6951 -1-
MV	0.7372 -4-	0.7365 -7-	0.3459 -13-	0.1506 -14-	0.2793 -7-	0.4176 -2-
MinV	0.7224 -5-	0.7871 -5-	0.5575 -2-	0.3344 -8-	0.3230 -2-	0.2029 -13-
VW	0.2213 -16-	0.2213 -17-	0.2213 -16-	0.2213 -11-	0.1363 -16-	0.1363 -17-
BS	0.7827 -2-	0.9405 -3-	0.4695 -7-	0.2418 -10-	0.3001 -3-	0.4164 -3-
MVsc	0.3987 -11-	0.4032 -10-	0.2464 -15-	0.1992 -13-	0.2456 -11-	0.2360 -9-
MinVsc	0.4183 -9-	0.4366 -9-	0.5270 -4-	0.4633 -2-	0.2601 -9-	0.2606 -8-
BSsc	0.4058 -10-	0.3952 -11-	0.3092 -14-	0.2693 -9-	0.2473 -10-	0.2306 -10-
MVlsc	0.4413 -8-	0.3608 -13-	0.1401 -17-	0.0813 -15-	0.2746 -8-	0.3162 -7-
MP	0.3471 -13-	0.3433 -14-	0.3651 -12-	0.3442 -7-	0.1924 -15-	0.2033 -12-
MV/MinV	0.7827 -3-	0.9533 -2-	0.4360 -9-	0.0733 -16-	0.2850 -5-	0.3879 -5-
EW/MinV	0.6692 -7-	0.6713 -8-	0.5469 -3-	0.4138 -5-	0.2847 -6-	0.1975 -14-
EW/MV	0.7052 -6-	0.7684 -6-	0.4194 -11-	0.2147 -12-	0.2862 -4-	0.4083 -4-
EW/MV/MinV	0.0171 -17-	0.8476 -4-	0.4993 -5-	0.0028 -17-	−0.0855 -17-	0.3805 -6-
VT	0.3448 -14-	0.3182 -15-	0.4809 -6-	0.4585 -4-	0.2023 -12-	0.1816 -15-
RRT	0.3288 -15-	0.2938 -16-	0.4601 -8-	0.4605 -3-	0.2009 -13-	0.1705 -16-

Omega Ratios						
Strategy	PF6 USA $N = 6$	PF25 USA $N = 25$	IND10 USA $N = 10$	IND48 USA $N = 48$	PF6 EU $N = 6$	PF25 EU $N = 25$
EW	1.8590 -12-	1.8548 -12-	2.0080 -9-	1.8789 -6-	1.4411 -14-	1.4752 -11-
MV IS	3.1504 -1-	5.5764 -1-	2.4311 -1-	3.3102 -1-	2.9362 -1-	5.6607 -1-
MV	2.7561 -5-	2.8840 -7-	1.7662 -13-	1.3334 -14-	1.6787 -7-	1.9233 -2-
MinV	2.7442 -6-	3.0296 -5-	2.2574 -3-	1.7156 -8-	1.7916 -2-	1.4648 -13-
VW	1.4856 -16-	1.4856 -17-	1.4856 -16-	1.4856 -12-	1.2906 -16-	1.2906 -17-
BS	2.8822 -3-	3.4604 -3-	2.0461 -8-	1.5318 -10-	1.7473 -3-	1.9011 -3-
MVsc	2.0111 -10-	2.0047 -10-	1.5698 -15-	1.4435 -13-	1.5588 -11-	1.5356 -9-
MinVsc	2.0010 -11-	2.0543 -9-	2.2244 -4-	2.0921 -4-	1.6131 -9-	1.6161 -8-
BSsc	2.0364 -9-	1.9698 -11-	1.7340 -14-	1.6136 -9-	1.5703 -10-	1.5256 -10-
MVlsc	2.1317 -8-	1.8301 -13-	1.2965 -17-	1.1646 -15-	1.6498 -8-	1.7287 -7-
MP	1.8124 -13-	1.7989 -14-	1.8536 -12-	1.8106 -7-	1.4272 -15-	1.4586 -14-
MV/MinV	2.8827 -2-	3.5023 -2-	1.9621 -11-	1.1474 -16-	1.7136 -5-	1.8382 -6-
EW/MinV	2.6660 -7-	2.8057 -8-	2.2876 -2-	2.0001 -5-	1.6902 -6-	1.4662 -12-
EW/MV	2.7933 -4-	3.0148 -6-	1.9689 -10-	1.5061 -11-	1.7175 -4-	1.8939 -4-
EW/MV/MinV	1.1538 -17-	3.3063 -4-	2.1387 -6-	1.0143 -17-	0.5361 -17-	1.8543 -5-
VT	1.8023 -14-	1.7277 -15-	2.1984 -5-	2.1654 -2-	1.4620 -12-	1.4093 -15-
RRT	1.7527 -15-	1.6588 -16-	2.1339 -7-	2.1591 -3-	1.4571 -13-	1.3802 -16-

accordance to the previous reward-to-risk ratios, unconstrained strategies related to the sample-based mean-variance and minimum-variance strategy as well as BS exhibit high values. VT and RRT continue to achieve high relative rankings for the two US industry data sets.

Results for the certainty equivalent returns are listed in Table 6.5. Since the certainty equivalent return is a common performance evaluation criterion, statistical tests to decide on the pairwise difference between the EW benchmark performance and the respective optimized strategies are available. Typically, the outperformance is decided on the Delta method as in DeMiguel et al. (2009b). However, this analysis additionally adjusted the Sharpe ratio studentized bootstrap methodology as introduced in Ledoit and Wolf (2008) to the context of certainty equivalent returns. This more general test allows for statistical inferences that do not crucially depend on the normality assumption of return distributions and hence allows for more reliable conclusions. Indeed, all out-of-sample portfolio return time series are statistically different from a normal distribution at a significance level of $\alpha = 0.01$ based on the Jarque-Bera test (results are not reported) and hence justify the use of a test statistic which accounts for this fact. As a result, a total of 17 certainty equivalent returns are superior relative to the EW benchmark of which 16 are related to either the PF6 or PF25 USA data sets. Superior optimized portfolio strategies are similar to the Sharpe ratio setting and are mainly MV, MinV, BS, MV/MinV, EW/MinV, and EW/MV. The main result, however, is unchanged and summarizes in the lack of an optimized portfolio strategy that achieves to consistently outperform the naive diversification benchmark. An explanation of the good performances of EW can be found when reconsidering the underlying data sets. All assets included in the data sets are obtained by combining several individual assets into portfolios such that the final data set can be better described as a portfolio of portfolios. Hence, investing naively in a selection of portfolios typically leads to better results relative to naively investing in a selection of individual assets that exhibit a high level of idiosyncratic volatility. The differences between the US and European PF6 and PF25 portfolios with respect to optimized portfolio strategies that achieve significant outperformances might be attributed to the different underlying time period available for estimation. The probability of the optimized strategies to outperform EW increases with an increasing out-of-sample estimation period as estimation precision increases and hence the European data sets face a greater challenge due to their relatively limited number of observations.

As a final step and to somehow circumvent the shortcoming of not having reliable statistical test statistics to differentiate between the reward-to-risk ratios of Sortino, Omega, Calmar, and the return on value-at-risk, a rank correlation investigation is applied in the style of Eling (2008). Table 6.6 presents the rank correlations with respect to the performance evaluation criteria for the two US and European PF6 and PF25 data sets. In case of PF6 USA, the lowest rank correlation still amounts to a value of 0.95 between SR and CR as well as CR and RoVaR. Equally high and significant correlations are obtained for PF25

Table 6.4: Calmar Ratios and Returns on Value-at-Risk. This table lists the out-of-sample Calmar ratios (CRs) and returns on value-at-risk (RoVaR) of the naive diversification (EW) approach, the in-sample CRs and RoVaRs for the mean-variance (MV IS) approach, and the out-of-sample CRs and RoVaRs of the 15 optimized portfolio strategies for all six data sets under consideration. The numbers indicated by -·- denote the relative ranking of the specific strategy within each data set based on the value of the CR and RoVaR. All results rely on a rolling-window approach with window length $M = 120$, risk-aversion parameter $\gamma = 1$, and tuning parameter $\eta = 1$ (only VT and RRT).

Calmar Ratios						
Strategy	PF6 USA $N = 6$	PF25 USA $N = 25$	IND10 USA $N = 10$	IND48 USA $N = 48$	PF6 EU $N = 6$	PF25 EU $N = 25$
EW	0.0209 -14-	0.0215 -13-	0.0226 -11-	0.0207 -6-	0.0124 -14-	0.0131 -11-
MV IS	0.0883 -1-	0.1711 -1-	0.0482 -1-	0.0684 -1-	0.0579 -1-	0.1682 -1-
MV	0.0805 -2-	0.0404 -7-	0.0252 -10-	0.0120 -14-	0.0227 -2-	0.0639 -2-
MinV	0.0473 -6-	0.0487 -6-	0.0382 -2-	0.0197 -7-	0.0208 -4-	0.0118 -13-
VW	0.0124 -16-	0.0124 -17-	0.0124 -16-	0.0124 -13-	0.0086 -16-	0.0086 -17-
BS	0.0754 -3-	0.0825 -2-	0.0345 -4-	0.0132 -11-	0.0220 -3-	0.0566 -3-
MVsc	0.0223 -10-	0.0226 -10-	0.0131 -15-	0.0131 -12-	0.0155 -10-	0.0147 -9-
MinVsc	0.0218 -11-	0.0225 -11-	0.0312 -6-	0.0255 -4-	0.0156 -9-	0.0153 -8-
BSsc	0.0225 -9-	0.0224 -12-	0.0167 -14-	0.0176 -9-	0.0154 -11-	0.0143 -10-
MVlsc	0.0271 -8-	0.0294 -9-	0.0097 -17-	0.0119 -15-	0.0181 -7-	0.0238 -7-
MP	0.0208 -15-	0.0211 -14-	0.0185 -13-	0.0195 -8-	0.0121 -15-	0.0128 -12-
MV/MinV	0.0668 -4-	0.0794 -3-	0.0314 -5-	0.0058 -16-	0.0205 -5-	0.0415 -5-
EW/MinV	0.0408 -7-	0.0381 -8-	0.0376 -3-	0.0220 -5-	0.0178 -8-	0.0117 -14-
EW/MV	0.0512 -5-	0.0659 -5-	0.0218 -12-	0.0142 -10-	0.0198 -6-	0.0356 -6-
EW/MV/MinV	0.0007 -17-	0.0697 -4-	0.0281 -7-	0.0001 -17-	−0.0414 -17-	0.0489 -4-
VT	0.0212 -12-	0.0203 -15-	0.0273 -8-	0.0285 -3-	0.0127 -12-	0.0112 -15-
RRT	0.0210 -13-	0.0193 -16-	0.0261 -9-	0.0294 -2-	0.0126 -13-	0.0106 -16-
Returns on Value-at-Risk						
Strategy	PF6 USA $N = 6$	PF25 USA $N = 25$	IND10 USA $N = 10$	IND48 USA $N = 48$	PF6 EU $N = 6$	PF25 EU $N = 25$
EW	0.1705 -12-	0.1696 -12-	0.1932 -9-	0.1690 -6-	0.0915 -14-	0.0975 -11-
MV IS	0.3760 -1-	0.6602 -1-	0.2645 -1-	0.3888 -1-	0.3176 -1-	0.7000 -1-
MV	0.3120 -6-	0.3152 -7-	0.1550 -13-	0.0673 -14-	0.1276 -8-	0.1566 -4-
MinV	0.3176 -4-	0.3542 -5-	0.2410 -3-	0.1446 -8-	0.1560 -2-	0.0918 -14-
VW	0.1022 -16-	0.1022 -17-	0.1022 -16-	0.1022 -11-	0.0623 -16-	0.0623 -17-
BS	0.3344 -2-	0.3979 -3-	0.2075 -8-	0.1051 -10-	0.1416 -3-	0.1594 -3-
MVsc	0.1912 -11-	0.1927 -10-	0.1117 -15-	0.0889 -12-	0.1143 -11-	0.1097 -9-
MinVsc	0.1933 -10-	0.2012 -9-	0.2314 -4-	0.2072 -4-	0.1231 -9-	0.1231 -8-
BSsc	0.1949 -9-	0.1882 -11-	0.1392 -14-	0.1213 -9-	0.1153 -10-	0.1074 -10-
MVlsc	0.2046 -8-	0.1608 -13-	0.0615 -17-	0.0350 -15-	0.1285 -7-	0.1410 -7-
MP	0.1627 -13-	0.1601 -14-	0.1667 -12-	0.1571 -7-	0.0889 -15-	0.0945 -12-
MV/MinV	0.3344 -3-	0.4015 -2-	0.1903 -10-	0.0298 -16-	0.1342 -6-	0.1545 -5-
EW/MinV	0.3031 -7-	0.3148 -8-	0.2429 -2-	0.1919 -5-	0.1375 -4-	0.0933 -13-
EW/MV	0.3173 -5-	0.3305 -6-	0.1897 -11-	0.0843 -13-	0.1346 -5-	0.1656 -2-
EW/MV/MinV	0.0099 -17-	0.3767 -4-	0.2194 -6-	0.0016 -17-	−0.0473 -17-	0.1424 -6-
VT	0.1611 -14-	0.1465 -15-	0.2205 -5-	0.2128 -2-	0.0947 -12-	0.0849 -15-
RRT	0.1526 -15-	0.1333 -16-	0.2106 -7-	0.2122 -3-	0.0939 -13-	0.0795 -16-

Table 6.5: Certainty Equivalent Returns. This table lists the out-of-sample certainty equivalent returns (CEs) of the naive diversification (EW) approach, the in-sample CEs for the mean-variance (MV IS) approach, and the out-of-sample CEs of the 15 optimized portfolio strategies for all six data sets under consideration. The first number in parentheses represents the p-value of the one-sided JK test on the pairwise difference between the out-of-sample CE of EW and strategy i, $i = \text{MV}, \text{MinV}, \ldots, \text{RRT}$. The second number in parentheses represents the p-value of the corresponding two-sided LW bootstrap test. CEs that outperform the EW benchmark at a significance level of $\alpha = 0.1$ with respect to LW are highlighted in bold. The numbers indicated by -·- denote the relative ranking of the specific strategy within each data set based on the value of the CE. All results rely on a rolling-window approach with window length $M = 120$, risk-aversion parameter $\gamma = 1$, and tuning parameter $\eta = 1$ (only VT and RRT).

Strategy	PF6 USA $N = 6$	PF25 USA $N = 25$	IND10 USA $N = 10$	IND48 USA $N = 48$	PF6 EU $N = 6$	PF25 EU $N = 25$
EW	0.0102 -12-	0.0104 -13-	0.0101 -6-	0.0099 -4-	0.0060 -14-	0.0065 -11-
MV IS	0.0196 -1-	0.0281 -1-	0.0114 -1-	0.0173 -1-	0.0285 -1-	0.0900 -1-
MV	**0.0184** (0.00) -2- (0.00)	**0.0228** (0.00) -2- (0.00)	0.0093 (0.35) -13- (0.72)	0.0053 (0.20) -14- (0.47)	0.0106 (0.14) -2- (0.36)	0.0248 (0.12) -2- (0.20)
MinV	**0.0145** (0.00) -7- (0.02)	**0.0154** (0.00) -8- (0.01)	0.0105 (0.39) -4- (0.81)	0.0076 (0.13) -11- (0.30)	0.0094 (0.06) -7- (0.16)	0.0057 (0.38) -14- (0.75)
VW	0.0057 (0.00) -16- (0.00)	0.0057 (0.00) -17- (0.00)	0.0057 (0.00) -16- (0.00)	0.0057 (0.00) -13- (0.00)	0.0037 (0.00) -16- (0.00)	0.0037 (0.00) -17- (0.00)
BS	**0.0169** (0.00) -3- (0.01)	**0.0210** (0.00) -3- (0.00)	0.0099 (0.46) -9- (0.94)	0.0079 (0.26) -9- (0.59)	0.0102 (0.10) -3- (0.29)	0.0228 (0.06) -3- (0.12)
MVsc	**0.0120** (0.02) -10- (0.04)	0.0119 (0.05) -10- (0.17)	0.0082 (0.17) -15- (0.35)	0.0072 (0.14) -12- (0.36)	0.0079 (0.04) -10- (0.22)	0.0077 (0.12) -9- (0.30)
MinVsc	0.0104 (0.40) -11- (0.81)	0.0105 (0.45) -12- (0.91)	0.0101 (0.50) -7- (0.99)	0.0093 (0.33) -6- (0.67)	0.0078 (0.03) -11- (0.12)	0.0077 (0.15) -8- (0.41)
BSsc	**0.0121** (0.01) -9- (0.03)	0.0116 (0.10) -11- (0.29)	0.0090 (0.24) -14- (0.46)	0.0086 (0.23) -8- (0.54)	0.0079 (0.03) -9- (0.19)	0.0074 (0.18) -10- (0.42)
MVlsc	**0.0146** (0.00) -6- (0.01)	0.0167 (0.06) -7- (0.14)	0.0045 (0.07) -17- (0.13)	−0.0109 (0.02) -16- (0.07)	0.0096 (0.03) -4- (0.22)	0.0149 (0.05) -7- (0.25)
MP	0.0100 (0.19) -13- (0.39)	0.0102 (0.09) -14- (0.19)	0.0094 (0.09) -12- (0.20)	0.0098 (0.24) -5- (0.54)	0.0058 (0.00) -15- (0.03)	0.0063 (0.00) -12- (0.01)
MV/MinV	**0.0165** (0.00) -4- (0.00)	**0.0198** (0.00) -4- (0.00)	0.0095 (0.37) -11- (0.76)	−0.0012 (0.02) -15- (0.03)	0.0095 (0.13) -6- (0.38)	0.0182 (0.06) -5- (0.13)
EW/MinV	**0.0138** (0.00) -8- (0.01)	**0.0139** (0.00) -9- (0.01)	0.0104 (0.38) -5- (0.80)	0.0092 (0.25) -7- (0.54)	0.0084 (0.08) -8- (0.22)	0.0057 (0.29) -13- (0.56)
EW/MV	**0.0163** (0.00) -5- (0.00)	**0.0183** (0.00) -5- (0.00)	0.0098 (0.33) -10- (0.67)	0.0076 (0.25) -10- (0.41)	0.0096 (0.08) -5- (0.27)	**0.0171** (0.03) -6- (0.07)
EW/MV/MinV	−0.0582 (0.00) -17- (0.45)	**0.0178** (0.00) -6- (0.00)	0.0099 (0.46) -8- (0.93)	−0.0152 (0.00) -17- (0.38)	−0.1815 (0.00) -17- (0.41)	0.0203 (0.13) -4- (0.24)
VT	0.0100 (0.25) -14- (0.55)	0.0097 (0.08) -15- (0.20)	0.0107 (0.07) -2- (0.19)	0.0108 (0.14) -3- (0.29)	0.0062 (0.27) -12- (0.57)	0.0055 (0.02) -15- (0.04)
RRT	0.0098 (0.18) -15- (0.35)	0.0092 (0.05) -16- (0.09)	0.0106 (0.07) -3- (0.16)	0.0110 (0.09) -2- (0.18)	0.0062 (0.31) -13- (0.66)	0.0051 (0.00) -16- (0.01)

Table 6.6: Performance Evaluation Criteria Rank Correlations. This table lists the Spearman's rank correlation coefficients for all performance evaluation criteria with respect to the data sets PF6 USA, PF25 USA, PF6 EU, and PF25 EU. Hence, this table contains a total of four correlation matrices. The lower left triangular matrix of the first group column denoted by PF6 / PF25 USA containing plain numbers represents the correlation matrix of the PF6 USA data set. The upper right triangular matrix of the first group column denoted by PF6 / PF25 USA containing italic numbers represents the correlation matrix of the PF25 USA data set. The second group column is defined accordingly. All listed rank correlations are significantly different from zero for a given significance level of $\alpha = 0.01$.

Performance	**PF6 / PF25 USA**						**PF6 / PF25 EU**					
Measure	SR	SoR	OR	CR	RoVaR	CE	SR	SoR	OR	CR	RoVaR	CE
SR	**1.00**	*1.00*	*1.00*	*0.97*	*1.00*	*0.89*	**1.00**	*0.99*	*0.98*	*0.97*	*1.00*	*0.97*
SoR	0.99	**1.00**	*1.00*	*0.97*	*1.00*	*0.89*	0.99	**1.00**	*0.99*	*0.99*	*0.99*	*0.99*
OR	0.99	0.98	**1.00**	*0.97*	*1.00*	*0.89*	0.99	1.00	**1.00**	*0.98*	*0.98*	*0.99*
CR	0.95	0.96	0.96	**1.00**	*0.97*	*0.95*	0.93	0.95	0.95	**1.00**	*0.97*	*1.00*
RoVaR	1.00	0.99	0.99	0.95	**1.00**	*0.89*	1.00	0.99	0.99	0.93	**1.00**	*0.97*
CE	0.96	0.97	0.98	0.97	0.96	**1.00**	0.89	0.90	0.90	0.97	0.89	**1.00**

USA, PF6 EU, and PF25 EU. As a result, the choice of a performance evaluation criterion does not excessively alter the relative performance ranking of the portfolio strategies under consideration if at all. Therefore, the Sharpe ratio appears to be a valid performance evaluation criterion despite its reliance on normally distributed return data.

The rank correlations listed in Table 6.7 focus on the relations between the six data sets under consideration. This time, rank correlations exhibit very heterogeneous values. In case of the Sharpe ratio, the correlation between PF6 USA and PF6 EU is high and amounts to 0.94. However, even the correlations between PF6 USA and PF25 USA but also between PF25 USA and PF25 EU decrease to 0.74 and 0.77, respectively, and hence are already lower than the lowest performance evaluation criteria rank correlation of 0.89.

Table 6.7: Data Sets Rank Correlations. This table lists the Spearman's rank correlation coefficients for all data sets with respect to the Sharpe ratio and certainty equivalent returns. Hence, this table contains a total of two correlation matrices. The lower left triangular matrix containing plain numbers represents the correlation matrix of the Sharpe ratios. The upper right triangular matrix containing italic numbers represents the correlation matrix of the certainty equivalent returns. The numbers in parentheses denote the p-values of a Student's t-distributed test statistic checking for a rank correlation significantly different from zero.

Performance	**Sharpe Ratio / Certainty Equivalent Returns**					
Measure	PF6 USA	PF25 USA	IND10 USA	IND48 USA	PF6 EU	PF25 EU
PF6 USA	**1.00**	*0.84* *(0.00)*	*−0.03* *(0.91)*	*−0.10* *(0.69)*	*0.97* *(0.00)*	*0.68* *(0.00)*
PF25 USA	0.74 (0.00)	**1.00**	*0.00* *(1.00)*	*−0.36* *(0.16)*	*0.80* *(0.00)*	*0.90* *(0.00)*
IND10 USA	0.23 (0.38)	0.43 (0.09)	**1.00**	*0.68* *(0.00)*	*0.00* *(1.00)*	*−0.11* *(0.67)*
IND48 USA	−0.04 (0.87)	−0.26 (0.32)	0.61 (0.01)	**1.00**	*−0.09* *(0.74)*	*−0.30* *(0.24)*
PF6 EU	0.94 (0.00)	0.66 (0.00)	0.38 (0.13)	0.11 (0.69)	**1.00**	*0.63* *(0.01)*
PF25 EU	0.64 (0.01)	0.77 (0.00)	0.01 (0.97)	−0.38 (0.13)	0.49 (0.05)	**1.00**

The two industry portfolios achieve an intermediate correlation of 0.61 between each other but have very low or even negative yet insignificant rank correlations to other data sets. Similar characteristics are obtained with respect to the certainty equivalent returns. The most striking performance differences between the data sets can be attributed to the VT and RRT strategies. Typically ranging at the end of the relative evaluation for the US and European PF6 and PF25 portfolios, the two timing strategies occupy top ranks for the two US industry data sets. Both strategies significantly outperform EW in case of the Sharpe ratio but marginally fail to do so for the certainty equivalent returns. These findings demonstrate that the relative performance of optimized portfolio strategies crucially depends on the underlying data set.

7. Robustness Checks

This section presents robustness checks for the most critical underlying assumptions with respect to the estimation methodology as well as for the most prominent parameters. Due to limitations regarding the scope of this analysis, there is no investigation about the effects of considering different subperiods on the out-of-sample portfolio performance. An interesting question in this context, for example, is the validity of results when excluding the global financial crisis in 2007 and 2008 and its adverse effects on returns. All findings of this section only rely on the Sharpe ratio and the certainty equivalent return as performance evaluation criteria and refer to the tables listed in Appendix D for the sake of readability.

7.1. Estimation Window

The robustness checks related to the estimation window are threefold. First, a rolling-window approach with a shorter window length of $M = 60$ months is considered. Second, a rolling-window approach with a longer window length of $M = 180$ months is examined. Third, the rolling-window approach is substituted with an expanding-window approach. The latter estimation methodology is similar to the rolling-window approach, but instead of always dropping the earliest return observation of the current window, all returns are kept such that the window expands through the data set.

In case of a shorter rolling-window estimation, results change both quantitatively and qualitatively for the Sharpe ratio. Only the MinV for PF6 USA and EW/MinV for PF6 USA and PF25 USA continue to outperform EW. The scenario for VT and RRT regarding the two US industry data sets is unchanged, i.e. these two strategies are still generally superior relative to EW. The most striking change can be seen for the two European data sets for which MVsc, MinVsc, BSsc, and MVlsc now significantly outperform EW. In general, the reduced estimation period increases the possibility of estimation errors which in turn might lead to more extreme optimal portfolio weight positions. Indeed, the most extreme position in this setting is the average minimum portfolio weight for EW/MV/MinV in case of PF25 USA which adopts a value of approximately $-63,000\%$ (not reported). Changes with respect to the certainty equivalent returns are in line with those reported for the Sharpe ratio.

Kritzman et al. (2010) argue that past empirical findings that speak in favor of the naive diversification approach heavily rely on the use of a short-term rolling-window estimation approach which might not be the appropriate technique to compare different optimized strategies. According to Kritzman et al. (2010), the reliance on a time series of only 60 or 120 months of return data as basis for estimating expected returns is not in accordance with the assumption of a rational investor since this short window does not suffice for a reliable estimation of the data generating parameters. Following their argument, an extension of the rolling-window by five years is considered such that a total of $M = 180$ months are included. In general, the results confirm the findings of the baseline scenario. Optimal portfolio strategies that significantly outperform EW are MV, MinV, BS, MinVsc, MV/MinV, EW/MinV, and EW/MV in case of PF6 and PF25 USA. The two US industry data sets are characterized by an outperformance of VT and RRT in three of the four cases. The European data sets do not include a superior strategy. The results for the certainty equivalent returns are comparable to the baseline setting. The robustness check regarding an expanding-window approach verifies the baseline results for both the Sharpe ratio and the certainty equivalent returns. In case of the latter performance criteria, also the PF6 EU data set exhibits now some outperforming optimized strategies that mainly coincide with the superior strategies of the PF6 USA data set.

7.2. Risk-Aversion Parameter

The robustness checks for the risk-aversion parameter are realized for two different scenarios with $\gamma = 2$ and $\gamma = 4$, respectively. The out-of-sample performances for the Sharpe ratios are independent of the specific risk-aversion parameter by the majority. Only EW/MV and EW/MV/MinV are affected by the risk-aversion parameter, however, all baseline results are still valid. In contrast, the certainty equivalent returns negatively depend on the corresponding risk-aversion parameter, i.e. the larger γ the lower the certainty equivalent returns. Still, the results are almost identical relative to the baseline scenario. The only major difference occurs in case of $\gamma = 4$ for which VT and RRT achieve to outperform EW for the two US industry data sets for the first time with respect to the certainty equivalent returns.

7.3. Tuning Parameter

The so-called tuning parameter refers only to VT and RRT and can be understood as a sensitivity of the resulting portfolio weights regarding changes in volatility of the return data. The robustness checks consider values of $\eta = 2$ and $\eta = 4$, respectively. All superior performances of the baseline setting remain valid. Furthermore, in case of $\eta = 4$, VT and RRT also outperform EW in case of PF6 USA and PF6 EU with respect to the Sharpe ratio. In the context of certainty equivalent returns, PF6 EU exhibits superior performances for VT and RRT for $\eta = 2$ as well as $\eta = 4$. These findings are suggestive of an increasing performance of these timing strategies with increasing values of the tuning parameter, i.e. a higher sensitivity to volatility changes of the underlying return data.

8. Conclusion

This analysis conducts a horse race between 15 optimized portfolio strategies based on six performance evaluation criteria across six data sets. Although no optimized portfolio strategy can be identified that consistently outperforms the naive diversification approach across all data sets, this finding does not automatically translate into a negation of the question of whether optimized portfolios can beat 1/N. In fact, several different optimized strategies achieve statistically superior results relative to the naive diversification approach such as the volatility and reward-to-risk timing strategies in case of the two US industry data sets. Whereas the specific choice of a performance evaluation criterion seems to have no effect on portfolio performance, results critically depend on the underlying data set. Not only do data sets that consist of a higher number of observations tend to facilitate outperforming behavior among optimized strategies, but also other characteristics inherent in the data sets seem to be differently preferred by portfolio strategies. For example, the timing strategies tend to be more successful relative to other optimized strategies the more volatile the underlying data. A direction to pursue for future research is to capture these portfolio strategy sensitivities to data set characteristics and develop a classification that links optimal strategies to data characteristics that favor a statistically significant outperforming behavior relative to the naive diversification approach.

Appendix A. Details of the Data Sets

The data used for this analysis are obtained from Kenneth R. French's website[5]. This appendix is a description of how the specific portfolios are created and further specifies the actual sources of the monthly return data. The idea of forming these portfolios as well as the actual creation of them is solely attributed and performed by Kenneth R. French. The following is an excerpt of the raw description of the data generating process provided on Kenneth R. French's website. Only minor changes have been made to the original wording so as to fit more to the background of this analysis. The creation of the different portfolios crucially depends on the following variables.

> Market equity (size) is price times shares outstanding. Price is from ... [the Center for Research and Security Prices (CRSP)], shares outstanding are from Compustat (if available) or CRSP.
>
> Book equity [(BE)] is constructed from Compustat data or collected from ... Moody's [...]. BE is the book value of stockholders' equity, plus balance sheet deferred taxes and investment tax credit (if available), minus the book value of preferred stock. Depending on availability, ... the redemption, liquidation, or par value (in that order) [is used] to estimate the book value of preferred stock. Stockholders' equity is the value reported by Moody's or Compustat, if it is available. If not, ... stockholders' equity [is measured] as the book value of common equity plus the par value of preferred stock, or the book value of assets minus total liabilities (in that order).

A.1. US MKT, PF6, PF25, IND10, and IND48

> ... [MKT], the excess return on the market, [is the value-weighted] ... return of all CRSP firms incorporated in the US and listed on the NYSE, AMEX, or NASDAQ that have a CRSP share code of 10 or 11 at the beginning of month t, good shares and price data at the beginning of t, and good return data for t minus the one-month Treasury bill rate (from Ibbotson Associates).
>
> The [six] portfolios, which are constructed at the end of each June, are the intersections of 2 portfolios formed on size (market equity, ME) and 3 portfolios formed on the ratio of book equity to market equity (BE/ME). The size breakpoint for year t is the median NYSE market equity at the end of June of year t. BE/ME for June of year t is the book equity for the last fiscal year end in $t-1$ divided by ME for December of $t-1$. The BE/ME breakpoints are the 30th and 70th NYSE percentiles.
>
> The [25] portfolios, which are constructed at the end of each June, are the intersections of 5 portfolios formed on size (market equity, ME) and 5 port-

[5] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

> folios formed on the ratio of book equity to market equity (BE/ME). The size breakpoints for year t are the NYSE market equity quintiles at the end of June of t. BE/ME for June of year t is the book equity for the last fiscal year end in $t-1$ divided by ME for December of $t-1$. The BE/ME breakpoints are NYSE quintiles.
>
> ... [Each] NYSE, AMEX, and NASDAQ stock [is assigned] to an industry portfolio at the end of June of year t based on its four-digit [Standard Industrial Classification (SIC)] ... code at that time. (... Compustat SIC codes for the fiscal year ending in calendar year $t-1$ [are used]. Whenever Compustat SIC codes are not available, ... CRSP SIC codes for June of year t [are used].) ... [Returns are then computed] from July of t to June of $t+1$.
>
> The portfolios for July of year t to June of $t+1$ include all NYSE, AMEX, and NASDAQ stocks for which ... [there is] market equity data for December of $t-1$ and June of t, and (positive) book equity data for $t-1$.

The ten industry sample consists of the following industries: Consumer Nondurables, Consumer Durables, Manufacturing, Energy, Business Equipment, Telecommunication, Wholesale and Retail, Health, Utilities, and Other.

The 48 industry sample consists of the following industries: Agriculture, Food Products, Candy and Soda, Beer and Liquor, Tobacco Products, Toys, Entertainment, Printing and Publishing, Consumer Goods, Clothes, Healthcare, Medical Equipment, Pharmaceutical Products, Chemicals, Rubber and Plastic Products, Textiles, Construction Material, Construction, Steel, Fabricated Products, Machinery, Electrical Equipment, Automobiles and Trucks, Aircraft, Shipbuilding and Railroad Equipment, Defense, Precious Metals, Nonmetallic and Industrial Metal Mining, Coal, Petroleum and Natural Gas, Utilities, Communication, Personal Services, Business Services, Computers, Electronic Equipment, Measuring and Control Equipment, Business Supplies, Shipping Containers, Transportation, Wholesale, Retail, Restaurants and Hotels and Motels, Banking, Insurance, Real Estate, Trading, and Other.

A.2. European MKT, PF6, and PF25

The returns on the European portfolio are constructed by averaging the returns on the country portfolios in accordance with their Europe, Australasia, and Far East weight. The raw data are from MSCI from 1991 to 2006 and from Bloomberg from 2007 to present.

> All returns are in U.S. dollars, include dividends and capital gains, and are not continuously compounded. ... [MKT] is the return ... [of Europe's value-weighted] market portfolio minus the U.S. one month T-bill rate.
>
> The [six] portfolios, which are constructed at the end of each June, are the intersections of 2 portfolios formed on size (market equity, ME) and 3 portfolios formed on the ratio of book equity to market equity (BE/ME). Big

> stocks are those in the top 90% of June market cap ..., and small stocks are those in the bottom 10%. The B/M breakpoints for big and small stocks ... are the 30th and 70th percentiles of B/M for the big stocks ....
>
> The [25] portfolios, which are constructed at the end of each June, are the intersections of 5 portfolios formed on size (market equity, ME) and 5 portfolios formed on the ratio of book equity to market equity (BE/ME). The size breakpoints ... are the 3rd, 7th, 13th, and 25th percentiles of ... [Europe's] aggregate market capitalization. The B/M breakpoints for all stocks ... are the 20th, 40th, 60th, and 80th percentiles of B/M for the big (top 90% of market cap) stocks ....
>
> The portfolios for July of year t to June of $t + 1$ include all stocks for which ... [there is] market equity data for December of $t - 1$ and June of t, and (positive) book equity data for $t - 1$.

Countries included in the European sample are: Austria, Belgium, Switzerland, Germany, Denmark, Spain, Finland, France, Great Britain, Greece, Ireland, Italy, Netherlands, Norway, Portugal, and Sweden.

Appendix B. Mean-Variance Optimization in Excess Returns

Let

$$\max_{x_t} x_{f,t} r_f + x_t^T \tilde{\mu}_t - \frac{\gamma}{2} x_t^T \Sigma_t x_t \quad \text{subject to} \quad x_{f,t} + \mathbf{1}_N^T x_t = 1 \tag{B.1}$$

denote the mean-variance optimization problem in terms of nonexcess returns, i.e. $\tilde{\mu}_t$ is a vector containing the expected nonexcess returns of the N available risky assets. It is possible to rewrite $x_{f,t} r_f + x_t^T \tilde{\mu}_t$ such that

$$\begin{aligned} x_{f,t} r_f + x_t^T \tilde{\mu}_t &= x_{f,t} r_f + x_t^T (\mu_t + \mathbf{1}_N r_f) \\ &= x_{f,t} r_f + x_t^T \mu_t + x_t^T \mathbf{1}_N r_f \\ &= (x_{f,t} + x_t^T \mathbf{1}_N) r_f + x_t^T \mu_t \\ &= r_f + x_t^T \mu_t. \end{aligned} \tag{B.2}$$

The first step utilizes the expression for the expected excess returns, i.e. $\mu_t = \tilde{\mu}_t - \mathbf{1}_N r_f$. Hence, Eq. (B.1) can be expressed in terms of excess returns μ_t such that the optimization problem transforms to

$$\max_{x_t} r_f + x_t^T \mu_t - \frac{\gamma}{2} x_t^T \Sigma_t x_t. \tag{B.3}$$

Since the risk-free rate r_f is a constant it does not alter the solution of the optimization problem and thus can be ignored yielding

$$\max_{x_t} x_t^T \mu_t - \frac{\gamma}{2} x_t^T \Sigma_t x_t. \tag{B.4}$$

Appendix C. Minimum-Variance Optimization

The minimum-variance strategy solves the optimization given by

$$\min_{\omega_t} \frac{1}{2} \omega_t^T \Sigma_t \omega_t \quad \text{subject to} \quad \mathbf{1}_N^T \omega_t = 1. \tag{C.5}$$

The Lagrangian is

$$\mathcal{L}(\omega_t, \lambda) = \frac{1}{2} \omega_t^T \Sigma_t \omega_t - \lambda(\mathbf{1}_N^T \omega_t - 1). \tag{C.6}$$

The first derivative with respect to the weight vector yields

$$\frac{\partial \mathcal{L}(\omega_t, \lambda)}{\partial \omega_t} = \Sigma_t \omega_t - \lambda \mathbf{1}_N = 0 \tag{C.7}$$

and the first derivative with respect to the Lagrange parameter yields

$$\frac{\partial \mathcal{L}(\omega_t, \lambda)}{\partial \lambda} = \mathbf{1}_N^T \omega_t - 1 = 0. \tag{C.8}$$

Solving Eq. (C.7) for ω_t gives $\omega_t = \lambda \Sigma_t^{-1} \mathbf{1}_N$. Inserting this result in Eq. (C.8) and solving for λ produces $\lambda = (\mathbf{1}_N^T \Sigma_t^{-1} \mathbf{1}_N)^{-1}$ with $(\Sigma_t^{-1})^T = \Sigma_t^{-1}$ due to symmetry of the variance-covariance matrix. Hence

$$\omega_t = \frac{\Sigma_t^{-1} \mathbf{1}_N}{\mathbf{1}_N^T \Sigma_t^{-1} \mathbf{1}_N}. \tag{C.9}$$

Appendix D. Robustness Checks Tables

Table D.1: Sharpe Ratios, Estimation Window $M = 60$. This table lists the out-of-sample Sharpe ratios (SRs) of the naive diversification (EW) approach, the in-sample SRs for the mean-variance (MV IS) approach, and the out-of-sample SRs of the 15 optimized portfolio strategies for all six data sets under consideration. The first number in parentheses represents the p-value of the one-sided JK test on the pairwise difference between the out-of-sample SR of EW and strategy i, $i = \text{MV}, \text{MinV}, \ldots, \text{RRT}$. The second number in parentheses represents the p-value of the corresponding two-sided LW bootstrap test. SRs that outperform the EW benchmark at a significance level of $\alpha = 0.1$ with respect to LW are highlighted in bold. The numbers indicated by - · - denote the relative ranking of the specific strategy within each data set based on the value of the SR. All results rely on a rolling-window approach with window length $M = 60$, risk-aversion parameter $\gamma = 1$, and tuning parameter $\eta = 1$ (only VT and RRT).

Strategy	PF6 USA $N = 6$	PF25 USA $N = 25$	IND10 USA $N = 10$	IND48 USA $N = 48$	PF6 EU $N = 6$	PF25 EU $N = 25$
EW	0.2616 -10-	0.2615 -6-	0.2744 -7-	0.2530 -7-	0.1619 -9-	0.1653 -8-
MV IS	0.4249 -1-	0.5860 -1-	0.3378 -1-	0.4357 -1-	0.3354 -1-	0.5592 -1-
MV	0.2456 (0.37) -12- (0.80)	−0.0039 (0.00) -17- (0.00)	0.1148 (0.00) -17- (0.01)	0.0071 (0.00) -13- (0.01)	−0.0554 (0.01) -16- (0.32)	−0.0580 (0.00) -16- (0.03)
MinV	**0.3725** (0.00) -2- (0.00)	0.3358 (0.05) -2- (0.12)	0.3165 (0.14) -3- (0.41)	0.0804 (0.00) -12- (0.02)	**0.2964** (0.00) -2- (0.01)	0.2694 (0.04) -3- (0.14)
VW	0.1593 (0.00) -16- (0.00)	0.1593 (0.00) -12- (0.00)	0.1593 (0.00) -15- (0.00)	0.1593 (0.00) -10- (0.00)	0.1147 (0.00) -11- (0.00)	0.1147 (0.00) -11- (0.00)
BS	0.2993 (0.19) -6- (0.49)	0.0019 (0.00) -16- (0.01)	0.2130 (0.10) -9- (0.26)	0.0040 (0.00) -14- (0.01)	−0.0501 (0.01) -14- (0.41)	−0.0579 (0.00) -15- (0.04)
MVsc	0.2683 (0.33) -9- (0.72)	0.2627 (0.47) -5- (0.97)	0.2083 (0.02) -11- (0.10)	0.2412 (0.35) -8- (0.73)	**0.2162** (0.00) -7- (0.05)	**0.2247** (0.01) -5- (0.02)
MinVsc	0.2900 (0.04) -7- (0.13)	0.2859 (0.10) -4- (0.21)	0.3149 (0.06) -4- (0.15)	**0.3038** (0.04) -2- (0.10)	**0.2400** (0.00) -5- (0.00)	**0.2558** (0.00) -4- (0.00)
BSsc	0.2731 (0.24) -8- (0.55)	0.2554 (0.36) -7- (0.79)	0.2223 (0.05) -8- (0.12)	0.2578 (0.43) -6- (0.86)	**0.2235** (0.00) -6- (0.03)	**0.2222** (0.01) -6- (0.01)
MVlsc	0.2409 (0.24) -14- (0.58)	0.2094 (0.12) -11- (0.34)	0.1374 (0.00) -16- (0.02)	0.1155 (0.01) -11- (0.07)	**0.2414** (0.02) -4- (0.08)	**0.2959** (0.02) -2- (0.09)
MP	0.2178 (0.04) -15- (0.15)	0.2111 (0.02) -10- (0.13)	0.2103 (0.01) -10- (0.09)	0.1928 (0.01) -9- (0.12)	0.0707 (0.01) -13- (0.14)	0.0701 (0.01) -12- (0.13)
MV/MinV	0.3020 (0.17) -5- (0.43)	0.0092 (0.00) -15- (0.01)	0.1977 (0.06) -12- (0.27)	−0.0179 (0.00) -16- (0.01)	−0.0505 (0.01) -15- (0.44)	−0.0575 (0.00) -14- (0.05)
EW/MinV	**0.3545** (0.00) -3- (0.00)	**0.3169** (0.00) -3- (0.02)	0.3195 (0.03) -2- (0.12)	0.2592 (0.19) -5- (0.38)	**0.2616** (0.00) -3- (0.01)	**0.2087** (0.02) -7- (0.05)
EW/MV	0.3151 (0.03) -4- (0.17)	0.0383 (0.00) -13- (0.02)	0.1632 (0.00) -14- (0.20)	−0.0201 (0.00) -17- (0.02)	−0.0601 (0.01) -17- (0.44)	−0.0591 (0.00) -17- (0.06)
EW/MV/MinV	0.0858 (0.00) -17- (0.04)	0.0351 (0.00) -14- (0.01)	0.1672 (0.02) -13- (0.35)	−0.0021 (0.00) -15- (0.01)	0.1086 (0.28) -12- (0.62)	−0.0541 (0.00) -13- (0.05)
VT	0.2550 (0.06) -11- (0.21)	0.2422 (0.01) -8- (0.02)	**0.3052** (0.00) -5- (0.02)	**0.2964** (0.00) -3- (0.01)	0.1663 (0.31) -8- (0.69)	0.1533 (0.15) -9- (0.36)
RRT	0.2433 (0.00) -13- (0.01)	0.2248 (0.00) -9- (0.00)	0.2903 (0.03) -6- (0.12)	**0.2880** (0.01) -4- (0.02)	0.1589 (0.37) -10- (0.78)	0.1382 (0.02) -10- (0.07)

Table D.2: Certainty Equivalent Returns, Estimation Window $M = 60$. This table lists the out-of-sample certainty equivalent returns (CEs) of the naive diversification (EW) approach, the in-sample CEs for the mean-variance (MV IS) approach, and the out-of-sample CEs of the 15 optimized portfolio strategies for all six data sets under consideration. The first number in parentheses represents the p-value of the one-sided JK test on the pairwise difference between the out-of-sample CE of EW and strategy i, $i = \text{MV}, \text{MinV}, \ldots, \text{RRT}$. The second number in parentheses represents the p-value of the corresponding two-sided LW bootstrap test. CEs that outperform the EW benchmark at a significance level of $\alpha = 0.1$ with respect to LW are highlighted in bold. The numbers indicated by -·- denote the relative ranking of the specific strategy within each data set based on the value of the CE. All results rely on a rolling-window approach with window length $M = 60$, risk-aversion parameter $\gamma = 1$, and tuning parameter $\eta = 1$ (only VT and RRT).

Strategy	PF6 USA $N = 6$		PF25 USA $N = 25$		IND10 USA $N = 10$		IND48 USA $N = 48$		PF6 EU $N = 6$		PF25 EU $N = 25$	
EW	0.0114	-12-	0.0118	-6-	0.0105	-9-	0.0108	-5-	0.0069	-9-	0.0071	-8-
MV IS	0.0187	-1-	0.0260	-1-	0.0114	-1-	0.0187	-1-	0.0179	-1-	0.0502	-1-
MV	0.0149 (0.12) (0.29)	-2-	−2.5287 (0.00) (0.41)	-16-	0.0060 (0.11) (0.18)	-17-	−0.5676 (0.00) (0.19)	-14-	−0.3742 (0.00) (0.46)	-16-	−332.9523 (0.00) (0.46)	-16-
MinV	0.0134 (0.09) (0.22)	-5-	0.0134 (0.21) (0.43)	-3-	0.0105 (0.49) (0.99)	-8-	0.0031 (0.01) (0.04)	-12-	**0.0115** (0.02) (0.07)	-3-	0.0119 (0.05) (0.15)	-3-
VW	0.0060 (0.00) (0.00)	-16-	0.0060 (0.00) (0.00)	-12-	0.0060 (0.00) (0.00)	-16-	0.0060 (0.00) (0.00)	-10-	0.0045 (0.00) (0.00)	-11-	0.0045 (0.00) (0.00)	-11-
BS	0.0143 (0.09) (0.24)	-3-	−0.5961 (0.00) (0.45)	-14-	0.0094 (0.32) (0.68)	-13-	−0.0898 (0.00) (0.37)	-13-	−0.1984 (0.00) (0.45)	-14-	−191.8042 (0.00) (0.44)	-15-
MVsc	0.0123 (0.13) (0.32)	-8-	0.0124 (0.25) (0.54)	-4-	0.0104 (0.49) (0.98)	-10-	0.0114 (0.35) (0.71)	-2-	**0.0099** (0.00) (0.02)	-6-	**0.0108** (0.00) (0.01)	-4-
MinVsc	0.0112 (0.39) (0.79)	-14-	0.0109 (0.19) (0.40)	-10-	0.0101 (0.37) (0.78)	-12-	0.0097 (0.20) (0.42)	-8-	**0.0096** (0.00) (0.03)	-7-	**0.0100** (0.00) (0.04)	-6-
BSsc	0.0122 (0.15) (0.34)	-9-	0.0116 (0.44) (0.90)	-7-	0.0101 (0.41) (0.79)	-11-	0.0098 (0.22) (0.46)	-7-	**0.0101** (0.00) (0.01)	-5-	**0.0105** (0.00) (0.01)	-5-
MVlsc	0.0122 (0.31) (0.66)	-10-	0.0151 (0.18) (0.37)	-2-	0.0085 (0.30) (0.59)	-14-	0.0057 (0.24) (0.49)	-11-	**0.0124** (0.00) (0.03)	-2-	**0.0231** (0.00) (0.00)	-2-
MP	0.0097 (0.09) (0.16)	-15-	0.0097 (0.05) (0.17)	-11-	0.0085 (0.04) (0.16)	-15-	0.0085 (0.04) (0.17)	-9-	0.0023 (0.01) (0.13)	-12-	0.0023 (0.01) (0.12)	-12-
MV/MinV	0.0133 (0.16) (0.36)	-6-	−1.2765 (0.00) (0.20)	-15-	0.0120 (0.31) (0.61)	-1-	−27.1123 (0.00) (0.23)	-17-	−0.1784 (0.00) (0.47)	-13-	−77.7802 (0.00) (0.45)	-13-
EW/MinV	0.0128 (0.10) (0.21)	-7-	0.0122 (0.34) (0.67)	-5-	0.0105 (0.47) (0.96)	-7-	0.0105 (0.23) (0.48)	-6-	**0.0101** (0.02) (0.08)	-4-	0.0084 (0.10) (0.28)	-7-
EW/MV	**0.0140** (0.03) (0.10)	-4-	−0.5287 (0.00) (0.41)	-13-	0.0109 (0.45) (0.82)	-5-	−1.0319 (0.00) (0.35)	-15-	−4.7530 (0.00) (0.46)	-17-	−85.9413 (0.00) (0.41)	-14-
EW/MV/MinV	−0.0057 (0.05) (0.20)	-17-	−676.3086 (0.00) (0.45)	-17-	0.0111 (0.44) (0.70)	-4-	−19.9794 (0.00) (0.21)	-16-	−0.3472 (0.00) (0.20)	-15-	−643.8187 (0.00) (0.44)	-17-
VT	0.0114 (0.38) (0.80)	-11-	0.0116 (0.34) (0.73)	-8-	0.0111 (0.06) (0.16)	-3-	0.0113 (0.26) (0.52)	-3-	0.0071 (0.33) (0.70)	-8-	0.0064 (0.13) (0.30)	-9-
RRT	0.0112 (0.32) (0.61)	-13-	0.0111 (0.15) (0.34)	-9-	0.0109 (0.10) (0.26)	-6-	0.0112 (0.27) (0.57)	-4-	0.0067 (0.38) (0.80)	-10-	0.0058 (0.02) (0.06)	-10-

Table D.3: Sharpe Ratios, Estimation Window $M = 180$. This table lists the out-of-sample Sharpe ratios (SRs) of the naive diversification (EW) approach, the in-sample SRs for the mean-variance (MV IS) approach, and the out-of-sample SRs of the 15 optimized portfolio strategies for all six data sets under consideration. The first number in parentheses represents the p-value of the one-sided JK test on the pairwise difference between the out-of-sample SR of EW and strategy i, $i = \text{MV}, \text{MinV}, \dots, \text{RRT}$. The second number in parentheses represents the p-value of the corresponding two-sided LW bootstrap test. SRs that outperform the EW benchmark at a significance level of $\alpha = 0.1$ with respect to LW are highlighted in bold. The numbers indicated by -··- denote the relative ranking of the specific strategy within each data set based on the value of the SR. All results rely on a rolling-window approach with window length $M = 180$, risk-aversion parameter $\gamma = 1$, and tuning parameter $\eta = 1$ (only VT and RRT).

Strategy	PF6 USA $N = 6$	PF25 USA $N = 25$	IND10 USA $N = 10$	IND48 USA $N = 48$	PF6 EU $N = 6$	PF25 EU $N = 25$
EW	0.2276 -10-	0.2259 -10-	0.2548 -11-	0.2315 -6-	0.1260 -11-	0.1289 -6-
MV IS	0.4299 -1-	0.6327 -1-	0.3278 -1-	0.4496 -1-	0.2702 -1-	0.6752 -1-
MV	**0.3702** (0.00) (0.03) -6-	**0.4374** (0.00) (0.01) -5-	0.2239 (0.25) (0.58) -14-	0.1303 (0.06) (0.21) -16-	0.1295 (0.48) (0.97) -9-	0.2252 (0.16) (0.37) -3-
MinV	**0.3750** (0.00) (0.00) -5-	**0.4241** (0.00) (0.00) -6-	0.3098 (0.08) (0.25) -3-	0.2241 (0.44) (0.90) -8-	0.1725 (0.15) (0.38) -2-	0.1006 (0.31) (0.59) -13-
VW	0.1640 (0.00) (0.01) -17-	0.1640 (0.00) (0.00) -16-	0.1640 (0.00) (0.00) -16-	0.1640 (0.00) (0.01) -13-	0.0984 (0.01) (0.07) -16-	0.0984 (0.02) (0.12) -14-
BS	**0.3849** (0.00) (0.01) -3-	**0.4577** (0.00) (0.00) -2-	0.2807 (0.26) (0.61) -7-	0.1854 (0.21) (0.52) -12-	0.1391 (0.40) (0.85) -4-	0.2227 (0.15) (0.33) -4-
MVsc	0.2203 (0.34) (0.72) -12-	0.2126 (0.27) (0.61) -13-	0.1950 (0.04) (0.12) -15-	0.1407 (0.01) (0.05) -14-	0.1260 (0.50) (1.00) -12-	0.1070 (0.12) (0.35) -12-
MinVsc	**0.2604** (0.04) (0.09) -9-	**0.2708** (0.02) (0.02) -9-	0.2940 (0.10) (0.25) -5-	0.2799 (0.07) (0.16) -2-	0.1276 (0.46) (0.93) -10-	0.1237 (0.38) (0.78) -8-
BSsc	0.2229 (0.39) (0.81) -11-	0.2129 (0.26) (0.58) -12-	0.2367 (0.27) (0.58) -13-	0.1874 (0.10) (0.29) -11-	0.1254 (0.49) (0.98) -13-	0.0941 (0.03) (0.12) -15-
MVlsc	0.2030 (0.21) (0.49) -15-	0.1537 (0.09) (0.26) -17-	0.0840 (0.00) (0.02) -17-	0.0298 (0.00) (0.02) -17-	0.1173 (0.37) (0.73) -15-	0.0391 (0.05) (0.29) -17-
MP	0.2190 (0.00) (0.01) -13-	0.2160 (0.00) (0.01) -11-	0.2382 (0.00) (0.01) -12-	0.2183 (0.00) (0.00) -9-	0.1250 (0.12) (0.37) -14-	0.1282 (0.11) (0.32) -7-
MV/MinV	**0.3851** (0.00) (0.01) -2-	**0.4539** (0.00) (0.00) -3-	0.2874 (0.21) (0.52) -6-	0.1359 (0.04) (0.12) -15-	0.1374 (0.41) (0.87) -5-	0.2157 (0.15) (0.31) -5-
EW/MinV	**0.3647** (0.00) (0.00) -7-	**0.3979** (0.00) (0.00) -8-	0.3103 (0.04) (0.15) -2-	0.2578 (0.19) (0.42) -5-	0.1639 (0.15) (0.40) -3-	0.1194 (0.39) (0.76) -9-
EW/MV	**0.3770** (0.00) (0.00) -4-	**0.4168** (0.00) (0.00) -7-	0.2583 (0.44) (0.91) -10-	0.1934 (0.03) (0.10) -10-	0.1346 (0.42) (0.88) -6-	0.2288 (0.07) (0.17) -2-
EW/MV/MinV	0.3456 (0.03) (0.11) -8-	**0.4469** (0.00) (0.00) -4-	0.2967 (0.09) (0.32) -4-	0.2242 (0.36) (0.72) -7-	−0.0897 (0.03) (0.11) -17-	0.0735 (0.31) (0.77) -16-
VT	0.2164 (0.02) (0.17) -14-	0.1974 (0.00) (0.03) -14-	**0.2792** (0.02) (0.08) -8-	**0.2736** (0.01) (0.05) -3-	0.1329 (0.19) (0.47) -7-	0.1142 (0.08) (0.13) -10-
RRT	0.2019 (0.00) (0.02) -16-	0.1766 (0.00) (0.01) -15-	0.2684 (0.08) (0.23) -9-	**0.2732** (0.02) (0.06) -4-	0.1325 (0.20) (0.47) -8-	0.1108 (0.04) (0.08) -11-

Table D.4: Certainty Equivalent Returns, Estimation Window $M = 180$. This table lists the out-of-sample certainty equivalent returns (CEs) of the naive diversification (EW) approach, the in-sample CEs for the mean-variance (MV IS) approach, and the out-of-sample CEs of the 15 optimized portfolio strategies for all six data sets under consideration. The first number in parentheses represents the p-value of the one-sided JK test on the pairwise difference between the out-of-sample CE of EW and strategy i, i = MV, MinV, . . . , RRT. The second number in parentheses represents the p-value of the corresponding two-sided LW bootstrap test. CEs that outperform the EW benchmark at a significance level of $\alpha = 0.1$ with respect to LW are highlighted in bold. The numbers indicated by -·- denote the relative ranking of the specific strategy within each data set based on the value of the CE. All results rely on a rolling-window approach with window length $M = 180$, risk-aversion parameter $\gamma = 1$, and tuning parameter $\eta = 1$ (only VT and RRT).

Strategy	PF6 USA $N = 6$	PF25 USA $N = 25$	IND10 USA $N = 10$	IND48 USA $N = 48$	PF6 EU $N = 6$	PF25 EU $N = 25$
EW	0.0096 -13-	0.0098 -11-	0.0096 -9-	0.0095 -5-	0.0055 -13-	0.0057 -6-
MV IS	0.0189 -1-	0.0287 -1-	0.0112 -1-	0.0195 -1-	0.0136 -1-	0.1526 -1-
MV	**0.0164** (0.00) -3- (0.02)	**0.0219** (0.00) -2- (0.00)	0.0084 (0.26) -14- (0.58)	0.0071 (0.28) -13- (0.60)	0.0064 (0.41) -5- (0.86)	0.0190 (0.08) -2- (0.17)
MinV	**0.0138** (0.00) -7- (0.03)	**0.0148** (0.00) -7- (0.02)	0.0102 (0.34) -3- (0.71)	0.0073 (0.14) -12- (0.32)	0.0072 (0.24) -2- (0.56)	0.0037 (0.25) -15- (0.47)
VW	0.0062 (0.00) -17- (0.00)	0.0062 (0.00) -17- (0.00)	0.0062 (0.00) -16- (0.00)	0.0062 (0.00) -15- (0.00)	0.0039 (0.01) -16- (0.03)	0.0039 (0.02) -13- (0.10)
BS	**0.0155** (0.00) -4- (0.02)	**0.0195** (0.00) -3- (0.00)	0.0095 (0.47) -10- (0.96)	0.0078 (0.26) -11- (0.57)	0.0064 (0.38) -4- (0.80)	0.0154 (0.08) -3- (0.21)
MVsc	0.0102 (0.26) -11- (0.54)	0.0095 (0.37) -13- (0.73)	0.0079 (0.13) -15- (0.25)	0.0070 (0.14) -14- (0.33)	0.0058 (0.38) -10- (0.79)	0.0048 (0.23) -10- (0.54)
MinVsc	0.0100 (0.30) -12- (0.59)	0.0103 (0.31) -9- (0.58)	0.0097 (0.48) -8- (0.97)	0.0092 (0.40) -7- (0.82)	0.0054 (0.44) -15- (0.89)	0.0050 (0.25) -8- (0.49)
BSsc	0.0103 (0.23) -10- (0.51)	0.0094 (0.35) -14- (0.70)	0.0089 (0.27) -13- (0.56)	0.0085 (0.26) -8- (0.57)	0.0058 (0.40) -11- (0.82)	0.0039 (0.06) -14- (0.20)
MVlsc	0.0108 (0.26) -9- (0.55)	0.0099 (0.49) -10- (0.99)	0.0035 (0.04) -17- (0.10)	−0.0135 (0.01) -17- (0.04)	0.0055 (0.49) -12- (0.97)	−0.0015 (0.11) -16- (0.34)
MP	0.0094 (0.14) -14- (0.23)	0.0096 (0.06) -12- (0.11)	0.0095 (0.29) -11- (0.55)	0.0094 (0.20) -6- (0.42)	0.0055 (0.24) -14- (0.55)	0.0057 (0.24) -7- (0.56)
MV/MinV	**0.0154** (0.00) -5- (0.02)	**0.0178** (0.00) -5- (0.00)	0.0103 (0.33) -2- (0.70)	0.0057 (0.08) -16- (0.17)	0.0063 (0.40) -6- (0.82)	0.0131 (0.09) -5- (0.23)
EW/MinV	**0.0133** (0.00) -8- (0.02)	**0.0137** (0.00) -8- (0.02)	0.0101 (0.34) -6- (0.74)	0.0083 (0.18) -9- (0.39)	0.0068 (0.25) -3- (0.56)	0.0046 (0.29) -11- (0.57)
EW/MV	**0.0153** (0.00) -6- (0.02)	**0.0172** (0.00) -6- (0.00)	0.0090 (0.25) -12- (0.56)	0.0096 (0.49) -4- (0.99)	0.0063 (0.39) -7- (0.84)	0.0135 (0.04) -4- (0.12)
EW/MV/MinV	**0.0170** (0.01) -2- (0.03)	**0.0182** (0.00) -4- (0.00)	0.0101 (0.32) -4- (0.70)	0.0083 (0.08) -10- (0.17)	−5.5939 (0.00) -17- (0.29)	−50.6224 (0.00) -17- (0.43)
VT	0.0094 (0.28) -15- (0.63)	0.0091 (0.10) -15- (0.24)	0.0101 (0.15) -5- (0.33)	0.0103 (0.20) -3- (0.38)	0.0059 (0.21) -8- (0.43)	0.0048 (0.06) -9- (0.09)
RRT	0.0090 (0.11) -16- (0.20)	0.0084 (0.05) -16- (0.12)	0.0100 (0.16) -7- (0.41)	0.0105 (0.15) -2- (0.34)	0.0059 (0.20) -9- (0.39)	0.0046 (0.03) -12- (0.05)

Table D.5: Sharpe Ratios, Expanding Window. This table lists the out-of-sample Sharpe ratios (SRs) of the naive diversification (EW) approach, the in-sample SRs for the mean-variance (MV IS) approach, and the out-of-sample SRs of the 15 optimized portfolio strategies for all six data sets under consideration. The first number in parentheses represents the p-value of the one-sided JK test on the pairwise difference between the out-of-sample SR of EW and strategy i, $i = \text{MV}, \text{MinV}, \ldots, \text{RRT}$. The second number in parentheses represents the p-value of the corresponding two-sided LW bootstrap test. SRs that outperform the EW benchmark at a significance level of $\alpha = 0.1$ with respect to LW are highlighted in bold. The numbers indicated by -·- denote the relative ranking of the specific strategy within each data set based on the value of the SR. All results rely on an expanding-window approach with initial length $M = 120$, risk-aversion parameter $\gamma = 1$, and tuning parameter $\eta = 1$ (only VT and RRT).

Strategy	PF6 USA $N = 6$	PF25 USA $N = 25$	IND10 USA $N = 10$	IND48 USA $N = 48$	PF6 EU $N = 6$	PF25 EU $N = 25$
EW	0.2396 -13-	0.2386 -10-	0.2663 -9-	0.2379 -8-	0.1379 -14-	0.1461 -12-
MV IS	0.4460 -1-	0.6351 -1-	0.3399 -1-	0.4350 -1-	0.3896 -1-	0.6324 -1-
MV	**0.3777** (0.00) -4- (0.04)	**0.3911** (0.00) -3- (0.04)	0.2456 (0.32) -13- (0.70)	0.1949 (0.23) -11- (0.55)	0.2013 (0.13) -6- (0.39)	0.2100 (0.22) -3- (0.51)
MinV	**0.3601** (0.00) -6- (0.04)	**0.3843** (0.00) -4- (0.04)	0.3096 (0.12) -3- (0.70)	0.2416 (0.47) -7- (0.55)	0.2390 (0.00) -3- (0.39)	0.1702 (0.29) -6- (0.51)
VW	0.1525 (0.00) -17- (0.00)	0.1525 (0.00) -17- (0.00)	0.1525 (0.00) -16- (0.00)	0.1525 (0.00) -14- (0.00)	0.0964 (0.00) -16- (0.01)	0.0964 (0.00) -17- (0.01)
BS	**0.3871** (0.00) -2- (0.02)	**0.4322** (0.00) -2- (0.00)	0.3000 (0.18) -4- (0.43)	0.2469 (0.43) -6- (0.86)	**0.2342** (0.02) -4- (0.08)	0.2266 (0.12) -2- (0.30)
MVsc	0.2624 (0.08) -10- (0.18)	0.2404 (0.46) -9- (0.93)	0.1948 (0.01) -15- (0.07)	0.1861 (0.09) -12- (0.27)	0.1653 (0.05) -10- (0.26)	0.1373 (0.31) -14- (0.67)
MinVsc	0.2436 (0.42) -12- (0.83)	0.2544 (0.25) -6- (0.51)	0.2959 (0.13) -5- (0.30)	**0.3007** (0.02) -2- (0.08)	**0.1738** (0.00) -8- (0.03)	**0.1767** (0.01) -5- (0.07)
BSsc	0.2579 (0.12) -11- (0.27)	0.2481 (0.27) -7- (0.59)	0.2512 (0.29) -12- (0.57)	0.2295 (0.40) -9- (0.82)	0.1666 (0.04) -9- (0.24)	0.1483 (0.45) -11- (0.92)
MVlsc	0.2648 (0.21) -8- (0.48)	0.2224 (0.38) -13- (0.79)	0.1359 (0.01) -17- (0.05)	0.0667 (0.00) -15- (0.03)	0.1584 (0.21) -11- (0.49)	0.1604 (0.40) -9- (0.87)
MP	0.2320 (0.00) -14- (0.01)	0.2300 (0.00) -12- (0.03)	0.2563 (0.00) -10- (0.00)	0.2284 (0.00) -10- (0.00)	0.1350 (0.00) -15- (0.01)	0.1432 (0.00) -13- (0.03)
MV/MinV	**0.3780** (0.00) -3- (0.01)	0.2026 (0.21) -15- (0.46)	0.2135 (0.11) -14- (0.26)	−0.0387 (0.00) -17- (0.00)	**0.2467** (0.01) -2- (0.04)	0.1625 (0.38) -8- (0.74)
EW/MinV	**0.3499** (0.00) -7- (0.00)	**0.3547** (0.00) -5- (0.00)	**0.3147** (0.02) -2- (0.09)	**0.2712** (0.01) -5- (0.03)	**0.1997** (0.00) -7- (0.02)	0.1537 (0.23) -10- (0.49)
EW/MV	**0.3761** (0.00) -5- (0.00)	0.2371 (0.46) -11- (0.93)	0.2529 (0.06) -11- (0.22)	0.0615 (0.00) -16- (0.00)	0.2016 (0.03) -5- (0.12)	0.1683 (0.18) -7- (0.50)
EW/MV/MinV	0.2647 (0.28) -9- (0.76)	0.2410 (0.46) -8- (0.92)	0.2680 (0.47) -8- (0.95)	0.1761 (0.00) -13- (0.01)	0.0770 (0.24) -17- (0.63)	0.2055 (0.12) -4- (0.24)
VT	0.2246 (0.00) -15- (0.02)	0.2059 (0.00) -14- (0.01)	**0.2865** (0.01) -6- (0.07)	**0.2879** (0.00) -3- (0.01)	0.1401 (0.38) -12- (0.79)	0.1268 (0.02) -15- (0.03)
RRT	0.2223 (0.00) -16- (0.03)	0.1981 (0.00) -16- (0.01)	0.2824 (0.03) -7- (0.12)	**0.2858** (0.00) -4- (0.01)	0.1395 (0.41) -13- (0.82)	0.1181 (0.00) -16- (0.01)

Table D.6: Certainty Equivalent Returns, Expanding Window. This table lists the out-of-sample certainty equivalent returns (CEs) of the naive diversification (EW) approach, the in-sample CEs for the mean-variance (MV IS) approach, and the out-of-sample CEs of the 15 optimized portfolio strategies for all six data sets under consideration. The first number in parentheses represents the p-value of the one-sided JK test on the pairwise difference between the out-of-sample CE of EW and strategy i, i = MV, MinV, . . . , RRT. The second number in parentheses represents the p-value of the corresponding two-sided LW bootstrap test. CEs that outperform the EW benchmark at a significance level of $\alpha = 0.1$ with respect to LW are highlighted in bold. The numbers indicated by -·- denote the relative ranking of the specific strategy within each data set based on the value of the CE. All results rely on an expanding-window approach with initial length $M = 120$, risk-aversion parameter $\gamma = 1$, and tuning parameter $\eta = 1$ (only VT and RRT).

Strategy	PF6 USA $N = 6$	PF25 USA $N = 25$	IND10 USA $N = 10$	IND48 USA $N = 48$	PF6 EU $N = 6$	PF25 EU $N = 25$
EW	0.0102 -12-	0.0104 -10-	0.0101 -9-	0.0099 -6-	0.0060 -14-	0.0065 -12-
MV IS	0.0196 -1-	0.0281 -1-	0.0114 -1-	0.0173 -1-	0.0285 -1-	0.0900 -1-
MV	**0.0175** (0.00) -2- (0.01)	**0.0198** (0.00) -2- (0.00)	0.0099 (0.46) -13- (0.95)	0.0112 (0.36) -2- (0.75)	0.0111 (0.07) -4- (0.21)	0.0164 (0.08) -2- (0.21)
MinV	**0.0136** (0.01) -7- (0.05)	**0.0136** (0.04) -5- (0.09)	0.0105 (0.37) -6- (0.77)	0.0083 (0.21) -12- (0.46)	**0.0100** (0.01) -5- (0.05)	0.0068 (0.45) -9- (0.91)
VW	0.0057 (0.00) -17- (0.00)	0.0057 (0.00) -17- (0.00)	0.0057 (0.00) -17- (0.00)	0.0057 (0.00) -14- (0.00)	0.0037 (0.00) -16- (0.00)	0.0037 (0.00) -17- (0.00)
BS	**0.0155** (0.00) -3- (0.02)	**0.0163** (0.00) -4- (0.00)	0.0105 (0.39) -8- (0.82)	0.0093 (0.39) -10- (0.81)	**0.0113** (0.02) -3- (0.09)	0.0132 (0.06) -3- (0.20)
MVsc	**0.0123** (0.00) -9- (0.02)	0.0115 (0.13) -7- (0.33)	0.0082 (0.10) -15- (0.23)	0.0090 (0.33) -11- (0.67)	0.0078 (0.03) -10- (0.19)	0.0063 (0.42) -14- (0.85)
MinVsc	0.0095 (0.23) -16- (0.45)	0.0099 (0.31) -12- (0.59)	0.0099 (0.45) -12- (0.89)	0.0101 (0.45) -5- (0.91)	**0.0075** (0.01) -11- (0.05)	0.0075 (0.11) -7- (0.24)
BSsc	**0.0119** (0.01) -11- (0.05)	0.0113 (0.13) -8- (0.30)	0.0096 (0.35) -14- (0.69)	0.0093 (0.34) -9- (0.71)	0.0079 (0.02) -9- (0.18)	0.0067 (0.41) -10- (0.84)
MVlsc	**0.0138** (0.02) -6- (0.04)	0.0180 (0.05) -3- (0.19)	0.0066 (0.13) -16- (0.29)	−0.0028 (0.06) -16- (0.13)	0.0079 (0.11) -8- (0.34)	0.0112 (0.20) -4- (0.52)
MP	0.0101 (0.24) -13- (0.55)	0.0103 (0.13) -11- (0.32)	0.0100 (0.40) -10- (0.80)	0.0098 (0.12) -8- (0.30)	0.0059 (0.00) -15- (0.01)	0.0064 (0.00) -13- (0.03)
MV/MinV	**0.0145** (-0.01) -5- (0.03)	0.0085 (0.18) -16- (0.37)	0.0107 (0.39) -2- (0.78)	−0.0242 (0.00) -17- (0.00)	**0.0115** (0.01) -2- (0.05)	0.0070 (0.43) -8- (0.85)
EW/MinV	**0.0130** (0.01) -8- (0.05)	**0.0126** (0.03) -6- (0.08)	0.0105 (0.33) -7- (0.71)	0.0099 (0.48) -7- (0.97)	**0.0084** (0.02) -7- (0.06)	0.0065 (0.50) -11- (0.99)
EW/MV	**0.0149** (0.00) -4- (0.00)	0.0110 (0.21) -9- (0.39)	0.0100 (0.44) -11- (0.88)	−0.0001 (0.02) -15- (0.03)	**0.0096** (0.03) -6- (0.09)	0.0080 (0.14) -6- (0.41)
EW/MV/MinV	0.0122 (0.17) -10- (0.43)	0.0092 (0.13) -15- (0.26)	0.0107 (0.29) -3- (0.59)	0.0073 (0.00) -13- (0.01)	−0.3003 (0.00) -17- (0.42)	0.0097 (0.12) -5- (0.29)
VT	0.0099 (0.11) -15- (0.25)	0.0096 (0.05) -13- (0.14)	0.0106 (0.08) -4- (0.21)	0.0109 (0.10) -4- (0.20)	0.0061 (0.37) -12- (0.74)	0.0054 (0.01) -15- (0.02)
RRT	0.0099 (0.25) -14- (0.50)	0.0094 (0.05) -14- (0.11)	0.0105 (0.08) -5- (0.21)	0.0110 (0.08) -3- (0.16)	0.0061 (0.39) -13- (0.80)	0.0049 (0.00) -16- (0.00)

Table D.7: Sharpe Ratios, Risk-Aversion Parameter $\gamma = 2$. This table lists the out-of-sample Sharpe ratios (SRs) of the naive diversification (EW) approach, the in-sample SRs for the mean-variance (MV IS) approach, and the out-of-sample SRs of the 15 optimized portfolio strategies for all six data sets under consideration. The first number in parentheses represents the p-value of the one-sided JK test on the pairwise difference between the out-of-sample SR of EW and strategy i, i = MV, MinV, ..., RRT. The second number in parentheses represents the p-value of the corresponding two-sided LW bootstrap test. SRs that outperform the EW benchmark at a significance level of $\alpha = 0.1$ with respect to LW are highlighted in bold. The numbers indicated by -··- denote the relative ranking of the specific strategy within each data set based on the value of the SR. All results rely on a rolling-window approach with window length $M = 120$, risk-aversion parameter $\gamma = 2$, and tuning parameter $\eta = 1$ (only VT and RRT).

Strategy	PF6 USA $N = 6$	PF25 USA $N = 25$	IND10 USA $N = 10$	IND48 USA $N = 48$	PF6 EU $N = 6$	PF25 EU $N = 25$
EW	0.2396 -12-	0.2386 -12-	0.2663 -9-	0.2379 -6-	0.1379 -14-	0.1461 -11-
MV IS	0.4460 -1-	0.6351 -1-	0.3399 -1-	0.4350 -1-	0.3896 -1-	0.6324 -1-
MV	**0.3912** (0.00) -6- (0.01)	**0.3942** (0.00) -7- (0.04)	0.2208 (0.17) -13- (0.46)	0.1037 (0.02) -15- (0.09)	0.1861 (0.23) -8- (0.58)	0.2227 (0.20) -4- (0.48)
MinV	**0.3964** (0.00) -4- (0.00)	**0.4302** (0.00) -5- (0.00)	0.3194 (0.08) -3- (0.23)	0.2078 (0.27) -8- (0.53)	**0.2219** (0.02) -2- (0.10)	0.1383 (0.44) -14- (0.87)
VW	0.1525 (0.00) -16- (0.00)	0.1525 (0.00) -17- (0.00)	0.1525 (0.00) -16- (0.00)	0.1525 (0.00) -11- (0.00)	0.0964 (0.00) -17- (0.00)	0.0964 (0.00) -17- (0.01)
BS	**0.4122** (0.00) -2- (0.00)	**0.4682** (0.00) -3- (0.00)	0.2827 (0.34) -8- (0.73)	0.1564 (0.08) -10- (0.24)	0.2040 (0.13) -3- (0.36)	0.2261 (0.18) -3- (0.43)
MVsc	0.2640 (0.08) -11- (0.20)	0.2658 (0.08) -10- (0.24)	0.1653 (0.00) -15- (0.02)	0.1342 (0.00) -12- (0.03)	0.1688 (0.06) -11- (0.26)	0.1627 (0.18) -9- (0.41)
MinVsc	0.2665 (0.07) -10- (0.15)	**0.2756** (0.03) -9- (0.04)	0.3091 (0.07) -4- (0.18)	0.2823 (0.08) -4- (0.18)	**0.1802** (0.01) -9- (0.10)	0.1803 (0.05) -8- (0.22)
BSsc	0.2683 (0.05) -9- (0.14)	0.2605 (0.11) -11- (0.32)	0.2010 (0.02) -14- (0.06)	0.1780 (0.04) -9- (0.13)	0.1700 (0.04) -10- (0.26)	0.1595 (0.23) -10- (0.52)
MVlsc	0.2793 (0.09) -8- (0.21)	0.2278 (0.42) -13- (0.85)	0.0953 (0.00) -17- (0.00)	0.0556 (0.00) -16- (0.02)	0.1874 (0.06) -7- (0.30)	0.2033 (0.16) -7- (0.46)
MP	0.2302 (0.00) -13- (0.03)	0.2270 (0.00) -14- (0.03)	0.2351 (0.00) -12- (0.04)	0.2233 (0.00) -7- (0.00)	0.1343 (0.00) -15- (0.02)	0.1420 (0.00) -12- (0.01)
MV/MinV	**0.4122** (0.00) -3- (0.00)	**0.4712** (0.00) -2- (0.00)	0.2630 (0.47) -10- (0.95)	0.0476 (0.00) -17- (0.01)	0.1946 (0.15) -5- (0.43)	0.2201 (0.18) -5- (0.44)
EW/MinV	**0.3826** (0.00) -7- (0.00)	**0.3938** (0.00) -8- (0.00)	**0.3215** (0.03) -2- (0.09)	0.2648 (0.13) -5- (0.27)	0.1988 (0.03) -4- (0.13)	0.1404 (0.42) -13- (0.83)
EW/MV	**0.3915** (0.00) -5- (0.00)	**0.3977** (0.00) -6- (0.00)	0.2567 (0.27) -11- (0.57)	0.1232 (0.00) -13- (0.00)	0.1932 (0.11) -6- (0.35)	0.2321 (0.10) -2- (0.24)
EW/MV/MinV	0.0506 (0.00) -17- (0.03)	**0.4475** (0.00) -4- (0.00)	0.2964 (0.15) -6- (0.37)	0.1061 (0.00) -14- (0.05)	0.0982 (0.29) -16- (0.61)	0.2120 (0.23) -6- (0.52)
VT	0.2282 (0.01) -14- (0.10)	0.2102 (0.00) -15- (0.00)	**0.2972** (0.00) -5- (0.03)	**0.2886** (0.00) -2- (0.01)	0.1422 (0.27) -12- (0.56)	0.1287 (0.03) -15- (0.05)
RRT	0.2178 (0.00) -15- (0.01)	0.1934 (0.00) -16- (0.01)	**0.2862** (0.02) -7- (0.09)	**0.2879** (0.00) -3- (0.01)	0.1412 (0.32) -13- (0.65)	0.1211 (0.01) -16- (0.01)

Table D.8: Certainty Equivalent Returns, Risk-Aversion Parameter $\gamma = 2$. This table lists the out-of-sample certainty equivalent returns (CEs) of the naive diversification (EW) approach, the in-sample CEs for the mean-variance (MV IS) approach, and the out-of-sample CEs of the 15 optimized portfolio strategies for all six data sets under consideration. The first number in parentheses represents the p-value of the one-sided JK test on the pairwise difference between the out-of-sample CE of EW and strategy i, $i = \text{MV}, \text{MinV}, \ldots, \text{RRT}$. The second number in parentheses represents the p-value of the corresponding two-sided LW bootstrap test. CEs that outperform the EW benchmark at a significance level of $\alpha = 0.1$ with respect to LW are highlighted in bold. The numbers indicated by - · - denote the relative ranking of the specific strategy within each data set based on the value of the CE. All results rely on an expanding-window approach with initial length $M = 120$, risk-aversion parameter $\gamma = 2$, and tuning parameter $\eta = 1$ (only VT and RRT).

Strategy	PF6 USA $N = 6$	PF25 USA $N = 25$	IND10 USA $N = 10$	IND48 USA $N = 48$	PF6 EU $N = 6$	PF25 EU $N = 25$
EW	0.0090 -12-	0.0092 -13-	0.0092 -8-	0.0088 -4-	0.0045 -14-	0.0050 -10-
MV IS	0.0185 -1-	0.0271 -1-	0.0107 -1-	0.0165 -1-	0.0253 -1-	0.0788 -1-
MV	**0.0171** (0.00) -2- (0.00)	**0.0209** (0.00) -2- (0.00)	0.0082 (0.31) -13- (0.67)	−0.0010 (0.04) -15- (0.12)	0.0081 (0.20) -4- (0.49)	0.0019 (0.43) -17- (0.81)
MinV	**0.0137** (0.00) -6- (0.00)	**0.0147** (0.00) -7- (0.01)	0.0099 (0.33) -3- (0.69)	0.0067 (0.16) -9- (0.34)	0.0083 (0.04) -3- (0.15)	0.0044 (0.41) -13- (0.82)
VW	0.0048 (0.00) -16- (0.00)	0.0048 (0.00) -17- (0.00)	0.0048 (0.00) -16- (0.00)	0.0048 (0.00) -11- (0.00)	0.0023 (0.00) -16- (0.00)	0.0023 (0.00) -16- (0.00)
BS	**0.0160** (0.00) -3- (0.00)	**0.0199** (0.00) -3- (0.00)	0.0092 (0.49) -9- (0.99)	0.0059 (0.18) -10- (0.43)	0.0085 (0.11) -2- (0.33)	0.0113 (0.29) -4- (0.51)
MVsc	**0.0107** (0.03) -10- (0.08)	0.0107 (0.06) -10- (0.20)	0.0063 (0.07) -15- (0.16)	0.0045 (0.04) -12- (0.14)	0.0063 (0.05) -11- (0.24)	0.0060 (0.16) -8- (0.37)
MinVsc	0.0095 (0.31) -11- (0.63)	0.0097 (0.32) -12- (0.65)	0.0095 (0.42) -6- (0.84)	0.0087 (0.45) -5- (0.88)	**0.0066** (0.02) -9- (0.10)	0.0065 (0.10) -7- (0.30)
BSsc	**0.0109** (0.02) -9- (0.06)	0.0104 (0.10) -11- (0.26)	0.0077 (0.16) -14- (0.29)	0.0069 (0.15) -8- (0.34)	0.0064 (0.04) -10- (0.21)	0.0058 (0.21) -9- (0.47)
MVlsc	**0.0129** (0.01) -8- (0.03)	0.0125 (0.21) -9- (0.45)	0.0005 (0.01) -17- (0.02)	−0.0337 (0.00) -17- (0.00)	0.0077 (0.05) -7- (0.27)	0.0103 (0.16) -5- (0.46)
MP	0.0088 (0.09) -13- (0.16)	0.0089 (0.04) -14- (0.08)	0.0084 (0.05) -12- (0.16)	0.0086 (0.09) -6- (0.22)	0.0043 (0.00) -15- (0.02)	0.0048 (0.00) -11- (0.01)
MV/MinV	**0.0156** (0.00) -4- (0.00)	**0.0188** (0.00) -4- (0.00)	0.0087 (0.39) -11- (0.81)	−0.0080 (0.00) -16- (0.00)	0.0078 (0.15) -5- (0.39)	0.0121 (0.18) -3- (0.34)
EW/MinV	**0.0131** (0.00) -7- (0.00)	**0.0132** (0.00) -8- (0.00)	0.0098 (0.30) -5- (0.64)	0.0085 (0.37) -7- (0.72)	0.0072 (0.06) -8- (0.19)	0.0045 (0.37) -12- (0.73)
EW/MV	**0.0152** (0.00) -5- (0.00)	**0.0170** (0.00) -5- (0.00)	0.0089 (0.29) -10- (0.62)	0.0026 (0.04) -13- (0.02)	0.0078 (0.10) -6- (0.28)	0.0129 (0.08) -2- (0.17)
EW/MV/MinV	−7.5672 (0.00) -17- (0.47)	**0.0168** (0.00) -6- (0.00)	0.0093 (0.49) -7- (0.97)	0.0015 (0.03) -14- (0.16)	−0.0026 (0.17) -17- (0.33)	0.0092 (0.36) -6- (0.70)
VT	0.0088 (0.15) -14- (0.34)	0.0083 (0.03) -15- (0.09)	0.0100 (0.05) -2- (0.15)	0.0100 (0.08) -3- (0.17)	0.0048 (0.27) -12- (0.59)	0.0040 (0.03) -14- (0.05)
RRT	0.0085 (0.08) -15- (0.15)	0.0077 (0.02) -16- (0.03)	0.0098 (0.06) -4- (0.16)	0.0101 (0.06) -2- (0.12)	0.0047 (0.32) -13- (0.63)	0.0036 (0.01) -15- (0.01)

Table D.9: Sharpe Ratios, Risk-Aversion Parameter $\gamma = 4$. This table lists the out-of-sample Sharpe ratios (SRs) of the naive diversification (EW) approach, the in-sample SRs for the mean-variance (MV IS) approach, and the out-of-sample SRs of the 15 optimized portfolio strategies for all six data sets under consideration. The first number in parentheses represents the p-value of the one-sided JK test on the pair-wise difference between the out-of-sample SR of EW and strategy i, $i = \text{MV}, \text{MinV}, \ldots, \text{RRT}$. The second number in parentheses represents the p-value of the corresponding two-sided LW bootstrap test. SRs that outperform the EW benchmark at a significance level of $\alpha = 0.1$ with respect to LW are highlighted in bold. The numbers indicated by -··- denote the relative ranking of the specific strategy within each data set based on the value of the SR. All results rely on a rolling-window approach with window length $M = 120$, risk-aversion parameter $\gamma = 4$, and tuning parameter $\eta = 1$ (only VT and RRT).

Strategy	PF6 USA $N = 6$	PF25 USA $N = 25$	IND10 USA $N = 10$	IND48 USA $N = 48$	PF6 EU $N = 6$	PF25 EU $N = 25$
EW	0.2396 -12-	0.2386 -12-	0.2663 -9-	0.2379 -6-	0.1379 -14-	0.1461 -11-
MV IS	0.4460 -1-	0.6351 -1-	0.3399 -1-	0.4350 -1-	0.3896 -1-	0.6324 -1-
MV	**0.3912** (0.00) (0.01) -5-	**0.3942** (0.00) (0.04) -6-	0.2208 (0.17) (0.44) -13-	0.1037 (0.02) (0.08) -15-	0.1861 (0.23) (0.57) -8-	0.2227 (0.20) (0.48) -4-
MinV	**0.3964** (0.00) (0.00) -4-	**0.4302** (0.00) (0.00) -5-	0.3194 (0.08) (0.23) -3-	0.2078 (0.27) (0.58) -9-	**0.2219** (0.02) (0.09) -2-	0.1383 (0.44) (0.89) -14-
VW	0.1525 (0.00) (0.00) -16-	0.1525 (0.00) (0.00) -17-	0.1525 (0.00) (0.00) -16-	0.1525 (0.00) (0.00) -12-	0.0964 (0.00) (0.00) -16-	0.0964 (0.00) (0.01) -17-
BS	**0.4122** (0.00) (0.00) -2-	**0.4682** (0.00) (0.00) -3-	0.2827 (0.34) (0.75) -8-	0.1564 (0.08) (0.24) -11-	0.2040 (0.13) (0.34) -3-	0.2261 (0.18) (0.41) -3-
MVsc	0.2640 (0.08) (0.19) -11-	0.2658 (0.08) (0.26) -10-	0.1653 (0.00) (0.02) -15-	0.1342 (0.00) (0.03) -13-	0.1688 (0.06) (0.26) -11-	0.1627 (0.18) (0.37) -9-
MinVsc	0.2665 (0.07) (0.16) -10-	**0.2756** (0.03) (0.05) -9-	0.3091 (0.07) (0.20) -4-	0.2823 (0.08) (0.16) -4-	**0.1802** (0.01) (0.09) -9-	0.1803 (0.05) (0.21) -8-
BSsc	0.2683 (0.05) (0.14) -9-	0.2605 (0.11) (0.33) -11-	0.2010 (0.02) (0.07) -14-	0.1780 (0.04) (0.16) -10-	0.1700 (0.04) (0.25) -10-	0.1595 (0.23) (0.53) -10-
MVlsc	0.2793 (0.09) (0.21) -8-	0.2278 (0.42) (0.86) -13-	0.0953 (0.00) (0.00) -17-	0.0556 (0.00) (0.03) -16-	0.1874 (0.06) (0.30) -7-	0.2033 (0.16) (0.46) -7-
MP	0.2302 (0.00) (0.03) -13-	0.2270 (0.00) (0.02) -14-	0.2351 (0.00) (0.04) -12-	0.2233 (0.00) (0.00) -7-	0.1343 (0.00) (0.03) -15-	0.1420 (0.00) (0.02) -12-
MV/MinV	**0.4122** (0.00) (0.00) -3-	**0.4712** (0.00) (0.00) -2-	0.2630 (0.47) (0.95) -10-	0.0476 (0.00) (0.00) -17-	0.1946 (0.15) (0.42) -5-	0.2201 (0.18) (0.41) -5-
EW/MinV	**0.3826** (0.00) (0.00) -7-	**0.3938** (0.00) (0.00) -7-	**0.3215** (0.03) (0.07) -2-	0.2648 (0.13) (0.26) -5-	0.1988 (0.03) (0.14) -4-	0.1404 (0.42) (0.84) -13-
EW/MV	**0.3861** (0.00) (0.00) -6-	**0.3932** (0.00) (0.00) -8-	0.2458 (0.13) (0.29) -11-	0.1248 (0.00) (0.00) -14-	0.1883 (0.14) (0.36) -6-	0.2293 (0.12) (0.30) -2-
EW/MV/MinV	0.0785 (0.01) (0.02) -17-	**0.4591** (0.00) (0.00) -4-	0.2947 (0.16) (0.42) -6-	0.2087 (0.14) (0.31) -8-	−0.0714 (0.01) (0.12) -17-	0.2191 (0.20) (0.48) -6-
VT	0.2282 (0.01) (0.10) -14-	0.2102 (0.00) (0.00) -15-	**0.2972** (0.00) (0.03) -5-	**0.2886** (0.00) (0.01) -2-	0.1422 (0.27) (0.60) -12-	0.1287 (0.03) (0.07) -15-
RRT	0.2178 (0.00) (0.01) -15-	0.1934 (0.00) (0.00) -16-	**0.2862** (0.02) (0.09) -7-	**0.2879** (0.00) (0.01) -3-	0.1412 (0.32) (0.67) -13-	0.1211 (0.01) (0.01) -16-

Table D.10: Certainty Equivalent Returns, Risk-Aversion Parameter $\gamma = 4$. This table lists the out-of-sample certainty equivalent returns (CEs) of the naive diversification (EW) approach, the in-sample CEs for the mean-variance (MV IS) approach, and the out-of-sample CEs of the 15 optimized portfolio strategies for all six data sets under consideration. The first number in parentheses represents the p-value of the one-sided JK test on the pairwise difference between the out-of-sample CE of EW and strategy i, $i = \text{MV}, \text{MinV}, \ldots, \text{RRT}$. The second number in parentheses represents the p-value of the corresponding two-sided LW bootstrap test. CEs that outperform the EW benchmark at a significance level of $\alpha = 0.1$ with respect to LW are highlighted in bold. The numbers indicated by -··- denote the relative ranking of the specific strategy within each data set based on the value of the CE. All results rely on an expanding-window approach with initial length $M = 120$, risk-aversion parameter $\gamma = 4$, and tuning parameter $\eta = 1$ (only VT and RRT).

Strategy	PF6 USA $N = 6$	PF25 USA $N = 25$	IND10 USA $N = 10$	IND48 USA $N = 48$	PF6 EU $N = 6$	PF25 EU $N = 25$
EW	0.0068 -12-	0.0069 -12-	0.0075 -9-	0.0067 -6-	0.0016 -14-	0.0020 -7-
MV IS	0.0164 -1-	0.0251 -1-	0.0095 -1-	0.0150 -1-	0.0189 -1-	0.0565 -1-
MV	**0.0146** (0.00) (0.00) -2-	**0.0169** (0.00) (0.01) -3-	0.0060 (0.23) (0.51) -13-	−0.0136 (0.00) (0.01) -15-	0.0032 (0.36) (0.77) -10-	−0.0438 (0.01) (0.02) -17-
MinV	**0.0123** (0.00) (0.00) -6-	**0.0133** (0.00) (0.00) -7-	0.0087 (0.23) (0.50) -2-	0.0051 (0.23) (0.46) -9-	0.0060 (0.02) (0.11) -2-	0.0019 (0.49) (0.97) -8-
VW	0.0028 (0.00) (0.00) -16-	0.0028 (0.00) (0.00) -17-	0.0028 (0.00) (0.00) -15-	0.0028 (0.00) (0.00) -11-	−0.0006 (0.00) (0.00) -16-	−0.0006 (0.00) (0.01) -14-
BS	**0.0141** (0.00) (0.00) -3-	**0.0177** (0.00) (0.00) -2-	0.0078 (0.44) (0.89) -8-	0.0019 (0.07) (0.22) -12-	0.0051 (0.16) (0.39) -3-	−0.0115 (0.14) (0.21) -16-
MVsc	0.0082 (0.06) (0.15) -10-	0.0082 (0.08) (0.21) -9-	0.0027 (0.01) (0.04) -16-	−0.0010 (0.00) (0.02) -13-	0.0032 (0.09) (0.35) -11-	0.0028 (0.24) (0.54) -4-
MinVsc	0.0077 (0.16) (0.34) -11-	0.0080 (0.14) (0.28) -10-	0.0083 (0.26) (0.57) -5-	0.0074 (0.31) (0.62) -4-	**0.0040** (0.01) (0.04) -7-	0.0040 (0.05) (0.18) -2-
BSsc	**0.0084** (0.04) (0.09) -9-	0.0080 (0.12) (0.30) -11-	0.0050 (0.06) (0.10) -14-	0.0036 (0.05) (0.15) -10-	0.0033 (0.07) (0.31) -9-	0.0026 (0.29) (0.63) -5-
MVlsc	0.0095 (0.07) (0.13) -8-	0.0040 (0.26) (0.52) -16-	−0.0075 (0.00) (0.00) -17-	−0.0792 (0.00) (0.00) -17-	0.0040 (0.12) (0.44) -8-	0.0011 (0.44) (0.91) -10-
MP	0.0065 (0.01) (0.04) -13-	0.0064 (0.00) (0.02) -13-	0.0065 (0.01) (0.08) -12-	0.0061 (0.01) (0.05) -7-	0.0014 (0.00) (0.01) -15-	0.0018 (0.00) (0.00) -9-
MV/MinV	**0.0138** (0.00) (0.00) -4-	**0.0169** (0.00) (0.00) -4-	0.0072 (0.42) (0.87) -10-	−0.0216 (0.00) (0.00) -16-	0.0046 (0.18) (0.44) -5-	0.0000 (0.41) (0.79) -13-
EW/MinV	**0.0116** (0.00) (0.00) -7-	**0.0119** (0.00) (0.00) -8-	0.0086 (0.17) (0.41) -3-	0.0071 (0.37) (0.76) -5-	**0.0049** (0.03) (0.10) -4-	0.0021 (0.49) (0.97) -6-
EW/MV	**0.0131** (0.00) (0.00) -5-	**0.0143** (0.00) (0.00) -6-	0.0069 (0.19) (0.39) -11-	−0.0046 (0.00) (0.00) -14-	0.0042 (0.17) (0.45) -6-	0.0034 (0.42) (0.80) -3-
EW/MV/MinV	−0.2084 (0.00) (0.34) -17-	**0.0154** (0.00) (0.00) -5-	0.0079 (0.37) (0.76) -7-	0.0054 (0.17) (0.33) -8-	−1.1525 (0.00) (0.16) -17-	−0.0051 (0.25) (0.43) -15-
VT	0.0064 (0.04) (0.13) -14-	0.0055 (0.00) (0.01) -14-	**0.0085** (0.03) (0.07) -4-	**0.0084** (0.03) (0.05) -3-	0.0018 (0.28) (0.61) -12-	0.0011 (0.05) (0.09) -11-
RRT	0.0059 (0.01) (0.02) -15-	0.0046 (0.00) (0.00) -15-	**0.0082** (0.04) (0.07) -6-	0.0084 (0.02) (0.14) -2-	0.0018 (0.33) (0.71) -13-	0.0007 (0.01) (0.02) -12-

Table D.11: Sharpe Ratios and Certainty Equivalent Returns, Tuning Parameter $\eta = 2$ and $\eta = 4$. This table lists the out-of-sample Sharpe ratios (SRs) and certainty equivalent returns (CEs) of the naive diversification (EW) approach and the out-of-sample SRs and CEs of the VT and RRT strategy for tuning parameter values of $\eta = 1$, $\eta = 2$, and $\eta = 4$ for all six data sets under consideration, respectively. The first number in parentheses represents the p-value of the one-sided JK test on the pairwise difference between the out-of-sample SR or CE of EW and strategy i, $i = \text{VT}, \text{RRT}$. The second number in parentheses represents the p-value of the corresponding two-sided LW bootstrap test. SRs and CEs that outperform the EW benchmark at a significance level of $\alpha = 0.1$ with respect to LW are highlighted in bold. All results rely on a rolling-window approach with window length $M = 120$ and risk-aversion parameter $\gamma = 1$.

Sharpe Ratios

Strategy	PF6 USA $N = 6$	PF25 USA $N = 25$	IND10 USA $N = 10$	IND48 USA $N = 48$	PF6 EU $N = 6$	PF25 EU $N = 25$
EW	0.2396	0.2386	0.2663	0.2379	0.1379	0.1461
$\text{VT}_{\eta=1}$	0.2282	0.2102	**0.2972**	**0.2886**	0.1422	0.1287
	(0.01)	(0.00)	(0.00)	(0.00)	(0.27)	(0.03)
	(0.10)	(0.00)	(0.03)	(0.01)	(0.56)	(0.06)
$\text{RTT}_{\eta=1}$	0.2178	0.1934	**0.2862**	**0.2879**	0.1412	0.1211
	(0.00)	(0.00)	(0.02)	(0.00)	(0.32)	(0.01)
	(0.02)	(0.00)	(0.10)	(0.01)	(0.68)	(0.02)
$\text{VT}_{\eta=2}$	0.2459	0.2353	**0.3133**	**0.3030**	0.1525	0.1431
	(0.04)	(0.30)	(0.00)	(0.00)	(0.02)	(0.37)
	(0.21)	(0.66)	(0.02)	(0.00)	(0.12)	(0.75)
$\text{RTT}_{\eta=2}$	0.2308	0.2092	**0.2976**	**0.3021**	**0.1520**	0.1310
	(0.07)	(0.01)	(0.00)	(0.00)	(0.02)	(0.05)
	(0.23)	(0.04)	(0.04)	(0.00)	(0.09)	(0.11)
$\text{VT}_{\eta=4}$	**0.2626**	**0.2627**	**0.3257**	**0.3201**	**0.1683**	0.1628
	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.05)
	(0.01)	(0.02)	(0.02)	(0.00)	(0.05)	(0.22)
$\text{RRT}_{\eta=4}$	**0.2516**	0.2346	**0.3125**	**0.3202**	**0.1733**	0.1538
	(0.01)	(0.37)	(0.00)	(0.00)	(0.00)	(0.21)
	(0.08)	(0.82)	(0.03)	(0.00)	(0.02)	(0.49)

Certainty Equivalent Returns

Strategy	PF6 USA $N = 6$	PF25 USA $N = 25$	IND10 USA $N = 10$	IND48 USA $N = 48$	PF6 EU $N = 6$	PF25 EU $N = 25$
EW	0.0102	0.0104	0.0101	0.0099	0.0060	0.0065
$\text{VT}_{\eta=1}$	0.0100	0.0097	0.0107	0.0108	0.0062	0.0055
	(0.25)	(0.08)	(0.07)	(0.14)	(0.27)	(0.02)
	(0.55)	(0.20)	(0.19)	(0.29)	(0.57)	(0.04)
$\text{RRT}_{\eta=1}$	0.0098	0.0092	0.0106	0.0110	0.0062	0.0051
	(0.18)	(0.05)	(0.07)	(0.09)	(0.31)	(0.00)
	(0.35)	(0.09)	(0.16)	(0.18)	(0.66)	(0.01)
$\text{VT}_{\eta=2}$	0.0103	0.0104	0.0109	0.0109	**0.0067**	0.0062
	(0.16)	(0.48)	(0.09)	(0.15)	(0.03)	(0.30)
	(0.42)	(0.97)	(0.19)	(0.31)	(0.09)	(0.57)
$\text{RRT}_{\eta=2}$	0.0102	0.0098	0.0108	0.0113	**0.0067**	0.0056
	(0.50)	(0.19)	(0.06)	(0.07)	(0.03)	(0.04)
	(0.99)	(0.39)	(0.17)	(0.16)	(0.05)	(0.07)
$\text{VT}_{\eta=4}$	0.0105	0.0109	0.0109	0.0110	**0.0075**	0.0072
	(0.23)	(0.08)	(0.17)	(0.19)	(0.00)	(0.09)
	(0.46)	(0.26)	(0.36)	(0.37)	(0.02)	(0.21)
$\text{RRT}_{\eta=4}$	**0.0107**	0.0107	0.0110	0.0117	**0.0078**	0.0068
	(0.02)	(0.31)	(0.07)	(0.06)	(0.00)	(0.29)
	(0.07)	(0.68)	(0.21)	(0.11)	(0.00)	(0.57)

References

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath, 1999, Coherent measures of risk, *Mathematical Finance* 9, 203–228.

Barry, Christopher B., 1974, Portfolio analysis under uncertain means, variances, and covariances, *Journal of Finance* 29, 515–522.

Bawa, Vijay S., 1975, Optimal rules for ordering uncertain prospects, *Journal of Financial Economics* 2, 95–121.

Bawa, Vijay S., 1978, Safety-first, stochastic dominance, and optimal portfolio choice, *Journal of Financial and Quantitative Analysis* 13, 255–271.

Benartzi, Shlomo, and Richard Thaler, 2007, Heuristics and biases in retirement savings behavior, *Journal of Economic Perspectives* 21, 81–104.

Benartzi, Shlomo, and Richard H. Thaler, 2001, Naive diversification strategies in defined contribution saving plans, *American Economic Review* 91, 79–98.

Bessler, Wolfgang, Heiko Opfer, and Dominik Wolff, 2017, Multi-asset portfolio optimization and out-of-sample performance: An evaluation of black-litterman, mean-variance, and naive diversification approaches, *European Journal of Finance* 23, 1–30.

Black, Fischer, and Robert Litterman, 1992, Global portfolio optimization, *Financial Analysts Journal* 48, 28–43.

Brandt, Michael W., 2009, Portfolio choice problems, in Yacine Aït-Sahalia, and Lars P. Hansen, eds., *Handbook of Financial Econometrics*, volume 1, 269–336, Amsterdam: North-Holland, Elsevier.

Broadie, Mark, 1993, Computing efficient frontiers using estimated parameters, *Annals of Operations Research* 45, 21–58.

Brown, Stephen, 1979, The effect of estimation risk on capital market equilibrium, *Journal of Financial and Quantitative Analysis* 14, 215–220.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Chopra, Vijay K., and William T. Ziemba, 1993, The effect of errors in means, variances, and covariances on optimal portfolio choice, *Journal of Portfolio Management* 19, 6–11.

Christensen, Peter O., and Gerald A. Feltham, 2003, *Economics of accounting - volume I, information in markets*, first edition, New York: Springer Science+Business Media, LLC.

Cohn, Richard A., Wilbur G. Lewellen, Ronald C. Lease, and Gary G. Schlarbaum, 1975, Individual investor risk aversion and investment portfolio composition, *Journal of Finance* 30, 605–620.

DeMiguel, Victor, Lorenzo Garlappi, Francisco J. Nogales, and Raman Uppal, 2009a, A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms, *Management Science* 55, 798–812.

DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal, 2009b, Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?, *Review of Financial Studies* 22, 1915–1953.

Dowd, Kevin, 2000, Adjusting for risk: An improved sharpe ratio, *International Review of Economics and Finance* 9, 209–222.

Efron, Bradley, and Carl Morris, 1977, Stein's paradox in statistics, *Scientific American* 236, 119–127.

Eling, Martin, 2008, Does the measure matter in the mutual fund industry?, *Financial Analysts Journal* 64, 54–66.

Fabozzi, Frank J., Dashan Huang, and Guofu Zhou, 2010, Robust portfolios: Contributions from operations research and finance, *Annals of Operations Research* 176, 191–220.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Fishburn, Peter C., 1977, Mean-risk analysis with risk associated with below-target returns, *American Economic Review* 67, 116–126.

Frahm, Gabriel, and Christoph Memmel, 2010, Dominating estimators for minimum-variance portfolios, *Journal of Econometrics* 159, 289–302.

Frahm, Gabriel, Tobias Wickern, and Christof Wiechers, 2012, Multiple tests for the performance of different investment strategies, *Advances in Statistical Analysis* 96, 343–383.

Frost, Peter A., and James E. Savarino, 1986, An empirical bayes approach to efficient portfolio selection, *Journal of Financial and Quantitative Analysis* 21, 293–305.

Golosnoy, Vasyl, and Yarema Okhrin, 2007, Multivariate shrinkage for optimal portfolio weights, *European Journal of Finance* 13, 441–458.

Harlow, W. Van, 1991, Asset allocation in a downside-risk framework, *Financial Analysts Journal* 47, 28–40.

Huberman, Gur, and Wei Jiang, 2006, Offering versus choice in 401 (k) plans: Equity exposure and number of funds, *Journal of Finance* 61, 763–801.

Jacobs, Heiko, Sebastian Müller, and Martin Weber, 2014, How should individual investors diversify? an empirical evaluation of alternative asset allocation policies, *Journal of Financial Markets* 19, 62–85.

Jagannathan, Ravi, and Tongshu Ma, 2003, Risk reduction in large portfolios: Why imposing the wrong constraints helps, *Journal of Finance* 58, 1651–1683.

James, William, and Charles Stein, 1961, Estimation with quadratic loss, in Jerzy Neyman, ed., *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, 361–379, Berkeley: University of California Press.

Jobson, Dave J., and Bob M. Korkie, 1980, Estimation for markowitz efficient portfolios, *Journal of the American Statistical Association* 75, 544–554.

Jobson, Dave J., and Bob M. Korkie, 1981, Performance hypothesis testing with the sharpe and treynor measures, *Journal of Finance* 36, 889–908.

Jorion, Philippe, 1985, International portfolio diversification with estimation risk, *Journal of Business* 259–278.

Jorion, Philippe, 1986, Bayes-stein estimation for portfolio analysis, *Journal of Financial and Quantitative Analysis* 21, 279–292.

Jorion, Philippe, 1992, Portfolio optimization in practice, *Financial Analysts Journal* 48, 68–74.

Kan, Raymond, and Guofu Zhou, 2007, Optimal portfolio choice with parameter uncertainty, *Journal of Financial and Quantitative Analysis* 42, 621–656.

Kaplan, Paul D., and James A. Knowles, 2004, Kappa: A generalized downsidde risk-adjusted performance measure, *Journal of Performance Measurement* 8, 42–54.

Keating, Con, and William F. Shadwick, 2002, A universal performance measure, *Journal of Performance Measurement* 6, 59–84.

Kirby, Chris, and Barbara Ostdiek, 2012, It's all in the timing: Simple active portfolio strategies that outperform naive diversification, *Journal of Financial and Quantitative Analysis* 47, 437–467.

Klein, Roger W., and Vijay S. Bawa, 1976, The effect of estimation risk on optimal portfolio choice, *Journal of Financial Economics* 3, 215–231.

Kourtis, Apostolos, George Dotsis, and Raphael N. Markellos, 2012, Parameter uncertainty in portfolio selection: Shrinking the inverse covariance matrix, *Journal of Banking and Finance* 36, 2522–2531.

Kritzman, Mark, Sébastian Page, and David Turkington, 2010, In defense of optimization: The fallacy of 1/n, *Financial Analysts Journal* 66, 31–39.

Ledoit, Oliver, and Michael Wolf, 2008, Robust performance hypothesis testing with the sharpe ratio, *Journal of Empirical Finance* 15, 850–859.

Ledoit, Olivier, and Michael Wolf, 2003, Improved estimation of the covariance matrix of stock returns with an application to portfolio selection, *Journal of Empirical Finance* 10, 603–621.

Ledoit, Olivier, and Michael Wolf, 2004, Honey, i shrunk the sample covariance matrix: Problems in mean-variance optimization, *Journal of Portfolio Management* 30, 110–119.

Lintner, John, 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, 13–37.

MacKinlay, Craig A., and Luboš Pástor, 2000, Asset pricing models: Implications for expected returns and portfolio selection, *Review of Financial Studies* 13, 883–916.

Markowitz, Harry, 1952, Portfolio selection, *Journal of Finance* 7, 77–91.

Memmel, Christoph, 2003, Performance hypothesis testing with the sharpe ratio, *Finance Letters* 1, 21–23.

Merton, Robert C., 1980, On estimating the expected return on the market: An exploratory investigation, *Journal of Financial Economics* 8, 323–361.

Michaud, Richard O., 1989, The markowith optimization enigma: Is “optimized” optimal?, *Financial Analysts Journal* 45, 31–42.

Michaud, Richard O., 1993, Are long-short equity strategies superior?, *Financial Analysts Journal* 49, 44–49.

Morin, Roger-A., and A. Fernandez Suarez, 1983, Risk aversion revisited, *Journal of Finance* 38, 1201–1216.

Mossin, Jan, 1966, Equilibrium in a capital asset market, *Econometrica: Journal of the Econometric Society* 34, 768–783.

Pennacchi, George Gaetano, 2008, *Theory of asset pricing*, first edition, Boston: Pearson/Addison-Wesley.

Pflug, Georg C., Alois Pichler, and David Wozabal, 2012, The 1/n investment strategy is optimal under high model ambiguity, *Journal of Banking and Finance* 36, 410–417.

Pástor, Luboš, 2000, Portfolio selection and asset pricing models, *Journal of Finance* 55, 179–223.

Pástor, Luboš, and Robert F. Stambaugh, 2000, Comparing asset pricing models: An investment perspective, *Journal of Financial Economics* 56, 335–381.

Samuelson, Paul A., 1967, General proof that diversification pays, *Journal of Financial and Quantitative Analysis* 2, 1–13.

Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–442.

Sharpe, William F., 1966, Mutual fund performance, *Journal of Business* 39, 119–138.

Sharpe, William F., 1994, The sharpe ratio, *Journal of Portfolio Management* 21, 49–58.

Sorensen, Eric H., Ronald Hua, and Edward Qian, 2007, Aspects of constrained long-short equity portfolios, *Journal of Portfolio Management* 33, 12–22.

Sortino, Frank A., and Robert Van Der Meer, 1991, Downside risk, *Journal of Portfolio Management* 17, 27–31.

Statman, Meir, 1987, How many stocks make a diversified portfolio?, *Journal of Financial and Quantitative Analysis* 22, 353–363.

Stein, Charles, 1956, Inadmissibility of the usual estimator for the mean of a multivariate normal distribution, in Jerzy Neyman, ed., *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, 197–206, Berkeley: University of California Press.

Tobin, James, 1958, Liquidity preference as behavior towards risk, *Review of Economic Studies* 25, 65–86.

Tu, Jun, and Guofu Zhou, 2011, Markowitz meets talmud: A combination of sophisticated and naive diversification strategies, *Journal of Financial Economics* 99, 204–215.

Young, Terry W., 1991, Calmar ratio: A smoother tool, *Futures* 20, 40.

Declaration

1. I hereby declare that this thesis, entitled:

   Can Optimized Portfolios Beat 1/N?

   is a result of my own work and that no other than the indicated aids have been used for its completion. Material borrowed directly or indirectly from the works of others is indicated in each individual case by acknowledgment of the source and also the secondary literature used.

   This work has not previously been submitted to any other examining authority and has not yet been published.

2. After completion of the examining process, this work will be given to the library of the University of Konstanz, where it will be accessible to the public for viewing and borrowing. As author of this work, I ~~agree~~ / do not agree*) to this procedure.

Konstanz, August 21, 2018
(Date) (Signature)

Erklärung

1. Ich versichere hiermit, dass ich die vorliegende Arbeit mit dem Thema:

   Can Optimized Portfolios Beat 1/N?

   selbständig verfasst und keine anderen Hilfsmittel als die angegebenen benutzt habe. Die Stellen, die anderen Werken dem Wortlaut oder dem Sinne nach entnommen sind, habe ich in jedem einzelnen Falle durch Angaben der Quelle, auch der benutzten Sekundärliteratur, als Entlehnung kenntlich gemacht. Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

2. Diese Arbeit wird nach Abschluss des Prüfungsverfahrens der Universitätsbibliothek Konstanz übergeben und ist durch Einsicht und Ausleihe somit der Öffentlichkeit zugänglich. Als Urheber der anliegenden Arbeit stimme ich diesem Verfahren ~~zu~~ / nicht zu*).

Konstanz, den 21. August 2018
(Datum) (Unterschrift)

*)Please delete as applicable / Nichtzutreffendes bitte streichen.