

NHH

Norwegian School of Economics  
Bergen, Spring 2017



# Higher-moment portfolios with practical constraints based on Polynomial goal programming

By Jens Gautefall Carlenius and Vis Chen

Supervisor: Trond Døskeland

Master Thesis, Finance

NORWEGIAN SCHOOL OF ECONOMICS

This thesis was written as a part of the Master of Science in Economics and Business Administration at NHH. Please note that neither the institution nor the examiners are responsible - through the approval of this thesis - for the theories and methods used, or results and conclusions drawn in this work.

## Abstract

This thesis contributes to the field of portfolio selection by constructing and analyzing the impact of incorporating higher-moments by Polynomial goal programming. We construct the mean-variance-skewness and the mean-variance-skewness-kurtosis portfolio over a 20-year horizon using 29 stocks from the S&P Global 1200-index. We examine the performance of higher-moment portfolios in terms of return, risk and allocation, compared to two benchmark portfolios; the traditional Markowitz portfolio and the global minimum variance portfolio. Our findings suggest that an investor obtains a higher return and risk-adjusted return by incorporating skewness into the mean-variance allocation framework. The mean-variance-skewness portfolio can further be improved by a diversification constraint as a result of the portfolio's occasional concentrated allocations, while limiting turnover turns out to be relatively detrimental for its performance. The results are less clear when both skewness and kurtosis are incorporated into the asset allocation framework, as the mean-variance-skewness-kurtosis portfolio is outperformed by the benchmark portfolios unless a turnover or a strong diversification constraint is imposed. In general we find that higher-moment portfolios obtain more optimal out-of-sample higher-moments at the cost of higher out-of-sample variance. The differences between the out-of-sample moments are augmented by rebalancing the portfolios or by imposing the strong diversification constraint.

## Acknowledgements

We want to thank our supervisor Trond Døskeland for his guidance throughout the process, especially for his constructive comments and his patience with our time-consuming thesis.

In addition we want to express gratitude to Heine Didriksen at the IT department of NHH for providing us with the 2 GHz Intel(R) Xeon(R) CPU X7550 on a six-processor socket server with 24 GB RAM, enabling us to handle the computational requirements for constructing dynamic higher-moment portfolios.

We also want to thank Michael Wolf from the University of Zurich for elaborating the difference test of moments, and for the corresponding code in R.

Bergen, June 2017

Jens Gautefall Carlenius

Vis Chen

# Contents

<b>1</b>	<b>Introduction</b>	<b>7</b>
<b>2</b>	<b>Literature Review</b>	<b>10</b>
2.1	Modern Portfolio Theory . . . . .	10
2.2	Shortcomings of the mean-variance framework . . . . .	11
2.3	The role of higher moments in portfolio selection . . . . .	12
2.4	The implication of skewness . . . . .	13
2.5	The implication of kurtosis . . . . .	14
2.6	Estimation of higher-moments . . . . .	15
2.7	Asset allocation framework incorporating higher-moments . . . . .	16
2.7.1	Primal versus dual approach . . . . .	16
2.7.2	Polynomial goal programming . . . . .	16
2.8	Additional considerations for portfolio construction . . . . .	18
2.8.1	Buy-hold versus rebalancing . . . . .	18
2.8.2	Short-sale constraint . . . . .	19
2.8.3	Diversification constraint . . . . .	19
2.8.4	Turnover constraint . . . . .	19
<b>3</b>	<b>Methodology</b>	<b>20</b>
3.1	Assumptions and notation . . . . .	21
3.2	Estimation of mean, covariance, coskewness and cokurtosis . . . . .	22
3.2.1	Sample estimates of moments over a rolling window . . . . .	22
3.2.2	Single-index estimates of moments over a rolling window . . . . .	23
3.2.3	Bayesian shrinkage of mean over a rolling window . . . . .	24
3.2.4	Bayesian shrinkage estimate of covariance, coskewness and cokurtosis, over a rolling window . . . . .	25
3.3	Portfolio moments . . . . .	26
3.4	Asset allocation by Polynomial goal programming . . . . .	29
3.4.1	Polynomial goal programming . . . . .	29
3.4.2	Additional constraints used in Polynomial goal programming . . . . .	31
3.4.3	Convex optimization by the interior-point method . . . . .	32
3.4.4	Non-convex optimization by differential evolution . . . . .	34
3.5	Performance evaluation . . . . .	37
3.5.1	Traditional measures . . . . .	37

3.5.2	Hypothesis test of difference between portfolio moments . . . . .	38
<b>4</b>	<b>Data</b>	<b>40</b>
<b>5</b>	<b>Model</b>	<b>40</b>
5.1	Forecasts of mean, variance, skewness and kurtosis . . . . .	42
5.2	Portfolio allocation, diversification and turnover . . . . .	44
5.3	Portfolio aspired moments . . . . .	46
<b>6</b>	<b>Results</b>	<b>47</b>
6.1	Out-of-sample performance . . . . .	48
6.1.1	Portfolio wealth . . . . .	48
6.1.2	Performance measures . . . . .	52
6.2	Out-of-sample portfolio moments . . . . .	56
6.2.1	Test of out-of-sample portfolio mean . . . . .	56
6.2.2	Test of out-of-sample portfolio variance . . . . .	57
6.2.3	Test of out-of-sample portfolio skewness . . . . .	58
6.2.4	Test of out-of-sample portfolio kurtosis . . . . .	59
<b>7</b>	<b>Conclusion</b>	<b>62</b>
<b>8</b>	<b>Limitations</b>	<b>64</b>
	<b>Appendix</b>	<b>72</b>
<b>A</b>	<b>Components for shrinkage intensities</b>	<b>72</b>
<b>B</b>	<b>Modifying turnover constraint for convex set</b>	<b>77</b>
<b>C</b>	<b>Standard error of hypothesis test</b>	<b>77</b>
<b>D</b>	<b>Description of stocks in empirical data set</b>	<b>81</b>
D.1	Descriptive statistics of the empirical data set . . . . .	81
D.2	Encoded colors for defensive and cyclical stocks . . . . .	83
<b>E</b>	<b>Model</b>	<b>84</b>
E.1	Shrinkage intensities . . . . .	84
E.2	Portfolio weight plots . . . . .	85
E.3	Quarterly in-sample portfolio mean, variance skewness and kurtosis . . . . .	87

<b>F</b>	<b>Cumulative returns for benchmark portfolios</b>	<b>90</b>
<b>G</b>	<b>R-code Appendix</b>	<b>91</b>

## List of Figures

1	Mean-variance efficient frontier . . . . .	11
2	Mean-variance efficient frontier with higher moments . . . . .	12
3	Skewness . . . . .	13
4	Kurtosis . . . . .	14
5	Portfolio selection process . . . . .	20
6	Quarterly forecasts of mean, variance, skewness and kurtosis by rolling shrinkage	43
7	Quarterly allocations of portfolios with basic constraints . . . . .	44
8	Quarterly in-sample moments for portfolios with basic constraints . . . . .	47
9	Cumulative returns for portfolios with basic constraints . . . . .	48
10	Cumulative returns for portfolios with mild div. constraint . . . . .	49
11	Cumulative returns for portfolios with strong div. constraint . . . . .	49
12	Cumulative returns for portfolios with turn. constraint . . . . .	50
13	Cumulative returns for portfolios with strong div. and turn. constraint . . . . .	50
14	Wealth of mean-variance-skewness portfolios . . . . .	51
15	Wealth of mean-variance-skewness-kurtosis portfolios . . . . .	52

## List of Appendix Figures

16	Shrinkage intensities . . . . .	84
17	Quarterly allocation of portfolios with mild div. constraint . . . . .	85
18	Quarterly allocation of portfolios with strong div. constraint . . . . .	85
19	Quarterly allocation of portfolios with mild turn. constraint . . . . .	86
20	Quarterly allocation of portfolios with strong turn. constraint . . . . .	86
21	Quarterly allocation of portfolios with strong div.- and turn. constraint . . . . .	87
22	Quarterly aspired portfolio moments given mild div. constraint . . . . .	87
23	Quarterly aspired portfolio moments given strong div. constraint . . . . .	88
24	Quarterly aspired portfolio moments given mild turn. constraint . . . . .	88
25	Quarterly aspired portfolio moments given strong turn. constraint . . . . .	89
26	Quarterly aspired portfolio moments given strong div.- and turn. constraint . . . . .	89
27	Wealth of global minimum variance portfolios . . . . .	90
28	Wealth of mean-variance portfolios . . . . .	90

## List of Tables

1	Portfolio specifications . . . . .	41
2	Turnover- and diversification rates . . . . .	45
3	Annualized performance measures . . . . .	55
4	Difference test of monthly out-of-sample mean with GMVP as benchmark . . . .	56
5	Difference test of monthly out-of-sample mean with MV portfolio as benchmark	57
6	Difference test of monthly out-of-sample log variance with GMVP as benchmark	58
7	Difference test of monthly out-of-sample log variance with MV portfolio as benchmark	58
8	Difference test of monthly out-of-sample skewness with GMVP as benchmark . .	59
9	Difference test of monthly out-of-sample skewness with MV portfolio as benchmark	60
10	Difference test of monthly out-of-sample kurtosis with GMVP as benchmark . .	60
11	Difference test of monthly out-of-sample kurtosis with MV portfolio as benchmark	61

## List of Appendix Tables

12	Descriptive statistics for cyclical stocks . . . . .	81
13	Descriptive statistics for defensive stocks . . . . .	82
14	Encoded colors for stocks in empirical data set . . . . .	83



## List of Abbreviations

<i>GMVP</i>	Global minimum variance portfolio
<i>MV</i>	Mean-variance
<i>MVS</i>	Mean-variance-skewness
<i>MVSK</i>	Mean-variance-skewness-kurtosis
<i>PGP</i>	Polynomial goal programming
<i>BH</i>	Buy-hold
<i>RB</i>	Rebalancing

# 1 Introduction

The mean-variance framework outlined by Markowitz (1952) is regarded as the cornerstone of Modern Portfolio Theory and captures two essential concepts in finance. The investor may reduce the idiosyncratic risk of his portfolio by diversification, and faces a trade-off between expected return and risk, illustrated by the mean-variance efficient frontier. The mean-variance framework is built upon the assumption that the investor focuses solely on mean and variance, and by that implies quadratic utility and/or that asset returns are normally distributed. Both of these assumptions are inadequate and have been rejected both theoretically and empirically. Samuelson (1970) shows that higher-moments are relevant to portfolio selection in a finite-time interval, while the presence of skewness and excess kurtosis in asset returns has been extensively documented, for instance by Mandelbrot (1963), Cootner (1966), Fama (1965) and Officer (1972). Utility functions allowing for higher-moments are further motivated and supported by Brooks et al. (2011). Accompanied by investors' general preference for positive skewness (Arditti (1971), Arditti and Levy (1975)) and aversion towards kurtosis (Jurczenko and Maillet, 2006), this implies that higher-moments should be incorporated in the asset allocation framework in order to avoid suboptimal allocations.

Several approaches for incorporating higher-moments into the traditional Markowitz portfolio have been proposed, and the major conceptual difference is whether the allocation is based on solving an approximated utility function or a multi-objective function. In this paper we construct portfolios based on the latter approach by using the Polynomial goal programming (PGP) framework for portfolio selection, introduced by Lai (1991) for mean-variance-skewness efficient portfolios, and by Lai et al. (2006) for mean-variance-skewness-kurtosis efficient portfolios. As of today, most papers regarding higher-moment efficient portfolios based on PGP have been evaluated in terms of in-sample performance over one holding period (Škrinjarić, 2013; Lai et al., 2006; Harvey et al., 2010). Our goal is to provide a hybrid version of the higher-moment portfolios that is of more practical relevance to the investor according to desired frequency, horizon and constraints.

In this thesis we construct the mean-variance-skewness- (MVS) and mean-variance-skewness-kurtosis (MVSK) portfolio, with the long-only, diversification- and turnover constraint, from a buy-hold and a rebalancing perspective. Along with the higher-moment portfolios we also construct two benchmark portfolios; a mean-variance efficient portfolio and the global minimum variance portfolio (GMVP). In order to compare the higher-moment portfolios to the benchmark portfolios, we evaluate their performance over a 20-year horizon with an investment universe

consisting of 29 stocks on the S&P Global 1200-index. Each portfolio is revised quarterly with a quarterly holding period. To cope with the large estimation errors related to statistical moments, especially higher-moments, we use Bayesian shrinkage estimates over a one year rolling window, as input to the allocation framework.

We examine the following two objectives in this thesis:

- Given different sets of constraints and either a buy-hold or a rebalancing strategy, is the performance of the higher-moment portfolios better than the traditional Markowitz portfolio and the global minimum variance portfolio, in terms of return, risk and allocation?
- Do the higher-moment portfolios obtain more optimal out-of-sample higher-moments, and less optimal out-of-sample mean and/or variance, compared to the mean-variance portfolio and the global minimum variance portfolio, given different sets of constraints?

The first objective is evaluated by comparing the terminal portfolio wealth and the risk-adjusted return by the Sharpe-ratio and the adjusted Sharpe-ratio, of the higher-moment portfolios and the benchmark portfolios. In order to capture the risk profile of the portfolios we analyze the standard deviation and the expected shortfall of each portfolio. We also address the allocation of each portfolio by comparing the turnover and diversification ratio.

The second objective is evaluated using a difference test adjusted for heteroscedastic and autocorrelated standard errors, where the adjustment of standard errors is due to time-varying moments. The motivation behind this objective is the fact that numerous papers, e.g. Lai et al. (2006) and Davies et al. (2009), find that the higher-moment portfolios obtain worse in-sample mean and/or variance compared to the traditional Markowitz portfolio and the global minimum variance portfolio, but that the higher-moment portfolios are compensated by a higher in-sample skewness and/or a lower in-sample kurtosis. Yet, whether the higher-moment portfolios constructed by PGP obtain more optimal out-of-sample moments have not been addressed, to our knowledge, and can be relevant to the investor in practice. If the better in-sample higher-moments do not persist out-of-sample, an investor incorporating higher-moments may be misled and construct portfolios on false premises. Since we construct portfolios with different sets of constraint, this enables us to observe the implication of imposing more practically oriented constraints on the portfolios' out-of-sample moments as well.

First, we present relevant literature regarding portfolio selection and provide a discussion of higher-moments and their relevance to the investor. Further on we outline the methodology for moment-based asset allocation by PGP, followed by an evaluation of the in-sample moments,

allocations and out-of-sample performance of the empirical portfolios constructed. Finally, we present the results of the statistical difference test of the portfolios' out-of-sample mean, variance, skewness and kurtosis.

## 2 Literature Review

### 2.1 Modern Portfolio Theory

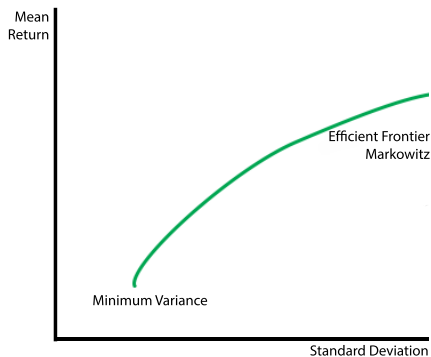
Modern Portfolio Theory (MPT) emerged by Markowitz (1952) work on the risk-return framework for investment decision making. Traditionally, portfolio theory has been centered around maximizing discounted returns, but Markowitz argued mathematically that investors should also focus on minimizing the risk. In order to quantify portfolio risk Markowitz used the variances and covariances of each asset in the portfolio, and derived how portfolio risk can be reduced by combining assets that are not perfectly correlated; a concept known as diversification. The implication of diversification is a focus on investing in several assets in order to reduce the impact of the risk of each asset, and choose assets from dissimilar industries to offset the losses of some by the gains of others.

The mean-variance framework is a normative theory based on the assumption that investors are risk-averse and only concerned about the expected return of a portfolio, measured by the mean, and its risk, measured by variance. This implies that an investor prefers the portfolio with the lowest variance for a given expected return, alternatively the portfolio with the highest expected return given the same level of variance. Any portfolio with such characteristics is mean-variance efficient and an analytical formulation of the optimization problem was given by Merton (1972) for the  $N$ -asset case;

$$\begin{aligned} & \underset{x}{\text{minimize}} && \frac{1}{2} \sqrt{\sum_{i=1}^N \sum_{j=1}^N x_i x_j \sigma_{ij}} \\ & \text{subject to} && \\ & && \sum_{i=1}^N x_i E[r_i] = E^* \\ & && \sum_{i=1}^N x_i = 1 \end{aligned} \tag{1}$$

where  $x_i$  is the percentage wealth allocated to asset  $i$ ,  $\sigma_{ij}$  denotes the covariance between asset  $i$  and  $j$ ,  $E[r_i]$  is the expected return of asset  $i$  and  $E^*$  is the desired expected portfolio return. sigma is the portfolio variance,  $w$  is the weight of asset  $i$  and  $\sigma_{ij}$  is the covariance between asset  $i$  and  $j$ . Markowitz (1952) illustrated all the mean-variance portfolios by an efficient frontier, illustrated in Figure 1. In fact, the efficient frontier illustrates the fundamental risk-reward trade-off in finance: the investor must be willing to take on higher risk in order to improve the

expected return.



**Figure 1:** Mean-variance efficient frontier

The origin of the efficient frontier is the portfolio with the lowest risk, also known as the global minimum variance portfolio (GMVP) Merton (1972). GMVP is obtained by solving the quadratic optimization problem in (1) without the constraint regarding a fixed expected return, and it is regarded as one of the most popular risk-optimization techniques (Clarke et al. (2006), Scherer (2010), Clarke et al. (2011)). In practice, the inputs of the mean-variance efficient portfolios suffer from estimation error of the expected return and risk, where Chopra and Ziemba (1993) show that the former contributes to the largest errors. Since the GMVP-portfolio only relies on the forecast of the variances and covariances of the assets, it is more robust than the other mean-variance efficient portfolios. Nevertheless, Stoyan Stoyanov (2011) argues that GMVP tends to be heavily concentrated in the assets with the lowest volatility and is only valuable for investors who seek to lower their portfolio’s volatility. His theoretical model shows that correlations are exploited only to increase concentration in further low volatility stocks and this leads to poorly diversified portfolios.

## 2.2 Shortcomings of the mean-variance framework

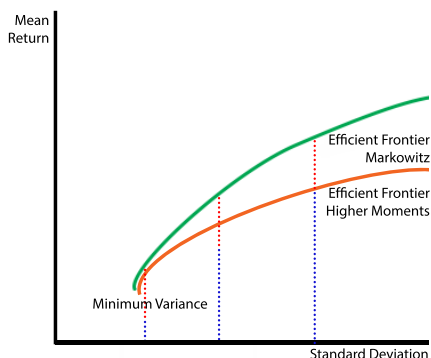
The mean-variance framework assumes a quadratic utility function and/or normally-distributed returns. While the latter might be correct according to Ito’s rule in a continuous-time model when asset prices follow a diffusion process, Samuelson (1970) shows that higher moments are relevant for portfolio selection over a finite time interval (Lai, 1991). Several empirical studies also confirm that stock return distributions deviate from normality and are characterized by skewness and kurtosis. For instance, Melnick and Everitt (2008) and Malmsten and Teräsvirta (2004) find that stocks on the S&P500-index exhibit a skewness and leptokurtosis, and Hwang and Satchell, E. (2001) find the same characteristics for the emerging markets. By neglecting

the empirical findings, an investor constructing a mean-variance efficient portfolio may neglect other moments of importance such as the skewness and kurtosis.

The standard assumption of the mean-variance framework is deterministic inputs, i.e. the expected return and variance are assumed to be true values. Yet, in practice the moments have to be estimated and as a result of not adjusting for the stochastic of the inputs, the framework is said to be an estimation error maximizer (Michaud, 1989). This is a result of the tendency of overweighing securities with large estimated returns, negative correlations and low risk, and vice versa. Due to the large estimation errors related to mean, are the unconstrained mean-variance portfolios found to be unreliable in practice (Chopra and Ziemba, 1993; Jobson and Korkie, 1980).

### 2.3 The role of higher moments in portfolio selection

Including higher moments into the traditional mean-variance framework is supported by several authors. For instance, Beardsley et al. (2012) find that even though the Markowitz solutions sometimes are close to the portfolio choice including higher moments, the compositions are never equal. In fact, an inefficient portfolio in the mean-variance framework may be efficient in the mean-variance-skewness or the mean-variance-skewness-kurtosis framework Lai (1991); Lai et al. (2006). Figure 2 demonstrates the implication of this: the efficient frontier including estimates of higher moments is never above the estimated mean-variance efficient frontier in the mean-variance space. In addition, the higher moments of a portfolio are not necessarily improved by diversification (Walther, 2014). In fact, Walter observes that diversification might reduce portfolio skewness and increase kurtosis. Thus the investor cannot rely on the higher-moments being appropriate with only a focus on the two first moments of the return distribution.

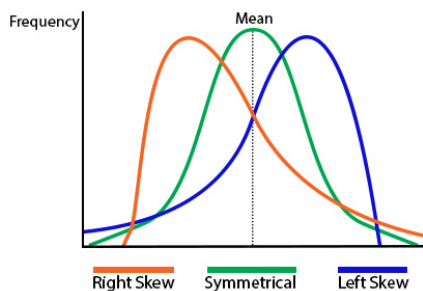


**Figure 2:** Mean-variance efficient frontier with higher moments

The no-free lunch concept of expected return and variance also applies to higher moments. The investor must be willing to forgo expected return or increase variance, if he wants to increase the skewness or lower the kurtosis of the portfolio, and vice versa. The willingness to do so depends on the investor’s utility function, yet in general risk averse investors often have a preference for odd moments as they are related to increasing gains, while even moments are disliked as they represent the risk of (large) losses (Athayde and Flores, 2004). Such preferences lead to a complex portfolio selection process, as maximizing expected return, minimizing variance, maximizing skewness and minimizing kurtosis, simultaneously, are competing and conflicting objectives (Lai et al., 2006).

## 2.4 The implication of skewness

Harvey and Siddique (2000) argue that investors should be compensated for bearing assets with systematic skewness. A risk-averse investor with non-increasing risk aversion prefers a positively skewed distribution to a negatively skewed distribution, because the former distribution, even though it is more likely to yield relatively lower returns, is less likely to yield extreme losses (Arditti, 1967). Kane (1982) derives how the mean-variance framework can be improved by including skewness, and Lai (1991) and Prakash et al. (2003) show how it affects the portfolio composition compared to the traditional Markowitz portfolios.



**Figure 3:** Skewness

Coskewness is the third standardized moment of a probability distribution and measures the joint degree of asymmetry around the mean. In this paper we use Pearson’s formula for coskewness, given by (2) for the return of asset  $x, y$  and  $z$  (Miller, 2014).

$$s(r_x, r_y, r_z) = \frac{E[(r_x - \mu_x)]E[(r_y - \mu_y)]E[(r_z - \mu_z)]}{\sigma_x \sigma_y \sigma_z} \quad (2)$$

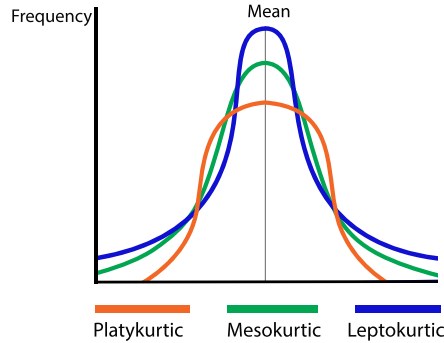
where  $\mu_i$  and  $\sigma_i$  is the expected value and standard deviation of asset  $i = x, y, z$ , respectively.



When  $r_x = r_y = r_z$ , we obtain the special case of coskewness known as skewness. Figure 3 contains three univariate distributions with different presence of skewness; one symmetrical distribution, i.e. no presence of skewness, together with a left- and right skewed distribution.

## 2.5 The implication of kurtosis

Signer and Favre (2002) suggest that negative excess kurtosis is preferred over positive excess kurtosis for an investor with decreasing absolute prudence, since the investor is more concerned about potential extreme losses than he favors potential extreme gains. A platykurtic distribution is preferred by a risk-averse investor, *ceteris paribus*, due to lower tail risk. The purpose of minimizing kurtosis is to protect the investors from extreme losses, yet at the same time one also reduces extreme gains (Saranya and Prasanna, 2014). Thus, incorporating kurtosis values in the portfolio optimization should contribute to better risk-adjusted returns, as one would underestimate risk by undermining kurtosis. Cokurtosis is the fourth standardized moment and



**Figure 4:** Kurtosis

measures the extent to which the distributions tend to have jointly relatively large frequencies in the tails. In this paper we apply Pearson’s formula for cokurtosis given by (3) for the return of asset  $x, y, z$  and  $v$  (Miller, 2014).

$$k(r_x, r_y, r_z, r_v) = \frac{E[(r_x - \mu_x)]E[(r_y - \mu_y)]E[(r_z - \mu_z)]E[(r_v - \mu_v)]}{\sigma_x \sigma_y \sigma_z \sigma_v} \quad (3)$$

where  $\mu_i$  and  $\sigma_i$  is the expected return and standard deviation of asset  $i = x, y, z, v$ , respectively. In case of  $r_x = r_y = r_z = r_v$ , the cokurtosis of the returns is reduced to the kurtosis. In order to normalize kurtosis one subtracts (3) by 3, since the normal distribution is mesokurtic and has a kurtosis of 3. Kurtosis is a measure of the peakness or flatness of a distribution relative to the normal distribution, and can be characterised by three formats. Figure 4 shows the shape of a mesokurtic distribution together with a platy- and leptokurtic distribution. A platykurtic

distribution has thinner tails than a normal distribution and lower kurtosis compared to a leptokurtic or a mesokurtic distribution, and with that follows a lower likelihood for extreme events, *ceteris paribus*.

## 2.6 Estimation of higher-moments

Martellini and Ziemann (2010) find that in order for higher-moment portfolios to outperform the global minimum variance portfolio, out-of-sample, appropriate forecasts of the higher moments are needed. Sample estimates are often used due to simplicity and the appealing property of being the maximum likelihood estimators under the normality assumption (Britten-Jones, 1999). Yet there are two major drawbacks of using sample estimates as it fails to capture the fact that moments are time-varying and are severely affected by increasing dimensionality (French et al., 1987; Brooks et al., 2005). The latter becomes a large problem when skewness and kurtosis are incorporated into the asset allocation framework.

To cope with the large estimation errors related to the higher-moments, Martellini and Ziemann (2010) suggest using a Bayesian shrinkage approach to construct a robust estimator based on a weighted average of the sample estimate and a shrinkage target. The weight, also called the shrinkage intensity, is based on the optimal trade-off between estimation error and specification error. While sample estimates are asymptotically unbiased, it suffers from estimation error (Disatnik and Benninga, 2007); especially when higher-moments are incorporated. Martellini and Ziemann (2010) compare the performance of higher-moment portfolios using both the constant correlation estimator by Elton and Gruber (1973) and the single-index model by Sharpe (1963), as shrinkage targets. They find that the investor's welfare is significantly improved by using Bayesian shrinkage over the sample method. In addition, portfolios based on shrinkage towards the single-index estimate generally outperforms portfolios based on shrinkage towards the constant correlation model. To illustrate the difference of structure between the sample estimates and the single-index estimates for the higher-moments, the authors show that while the sample estimates of the coskewness- and cokurtosis matrix require 2,925 and 23,725 parameters to not be rank deficient, respectively, the single-index estimates only require 51 and 77 parameters, respectively, for a portfolio with 25 assets. The trade-off of using the single-index estimates is the introduction of specification error due to the model's underlying assumptions, presented in section 3.2.2, and the fact that the common market factor only to some extent explains asset returns.

## 2.7 Asset allocation framework incorporating higher-moments

### 2.7.1 Primal versus dual approach

There is considered two conceptually different ways of incorporating higher moments into portfolio selection; the primal and dual approach (Jurczenko et al., 2015). The latter method is based on approximated utility functions derived from a Taylor series expansion of expected utility (Jondeau and Rockinger, 2006). The problem with this approach is that the inclusion of skewness and kurtosis does not guarantee an improved approximation of the utility function (Brockett and Garven, 1998), nor does it guarantee a solution at all. In addition, the investor's subjective utility function is generally unknown or very complicated, hence the reliability of the optimal portfolios derived by the dual approach is questionable Lai (1991).

The primal approach is based on the multi-objective approach known as Polynomial goal programming (PGP) (Lai, 1991). PGP was introduced by Tay and Leonard to explicitly incorporate bank balance-sheet managers' conflicting objectives such as maximization of returns and minimization of risks (Kumar et al., 1988). Lai (1991) was the first to apply this framework regarding portfolio selection by constructing the mean-variance-skewness portfolio, while Lai et al. (2006) modified the framework to incorporate kurtosis as well. The framework has been empirically tested on stock-, index- and hedge fund portfolios (Sun and Yan, 2003; Chunchachinda et al., 1997; Davies et al., 2009). Compared to the dual approach, PGP is not related precisely to the expected utility function but requires the investor to specify a preference parameter for each moment. Thus PGP constructs portfolios based on arbitrarily chosen preference parameters rather than on a utility function. While this may be considered a shortcoming from an academic perspective it may also be argued that it has a higher practical value due to parsimony. A shortcoming of PGP is the fact that the framework in theory leads to a Pareto-optimal solution, but in practice does not guarantee an efficient portfolio due to the non-convex optimization problem (Jurczenko et al., 2015).

### 2.7.2 Polynomial goal programming

This paper is based on PGP-constructed portfolios due to the intuitive concept of the framework together with the simplicity of handling moment preferences. Another feature of applying the PGP-framework for portfolio selection is that there is always an existing optimal solution (Lai, 1991). PGP deals with the conflicting and competing nature of the optimization problem of each moment by stepwise optimization (Lai et al., 2006). In the first step, the optimal values

of the portfolio moments are found by maximizing the expected return ( $M^*$ ), minimizing the variance ( $V^*$ ), maximizing the skewness ( $S^*$ ) and minimizing the kurtosis ( $K^*$ ), separately.

The deviation of the portfolio moment  $i$  and the related optimal moment, found in step one, is denoted  $d_i$ . In the second step the four moments are then consolidated into the objective function  $Z$  as the normalized Minkowski distance:

$$Z(\lambda) = \sum_{i=1}^m d_i^{*\lambda_i} = \sum_{i=1}^m \left| \frac{d_i}{Y_i} \right|^{\lambda_i} \quad (4)$$

where  $\lambda_i$  is a preference parameter for moment  $i$  and  $Y_i$  is its optimal value derived from the first step. The investor's subjective preference of each moment is specified by  $\lambda_i$ , and the higher the preference for moment  $i$ , the higher value on  $\lambda_i$  (Lai et al., 2006). Solving PGP for  $\lambda_1 = 1$ ,  $\lambda_2 = 1$ ,  $\lambda_3 = 0$  and  $\lambda_4 = 0$ , gives a mean-variance efficient portfolio (Lai, 1991). To derive mean-variance-skewness efficient portfolios the preference parameters are set to  $\lambda_1 \geq 0$ ,  $\lambda_2 > 0$ ,  $\lambda_3 > 0$ , and  $\lambda_4 = 0$ ; while the mean-variance-skewness-kurtosis efficient portfolios are based on  $\lambda_1 \geq 0$ ,  $\lambda_2 > 0$ ,  $\lambda_3 > 0$ , and  $\lambda_4 > 0$ . There have been conducted empirical studies to identify reasonable preference parameters for different types of investors. Proelss and Schweizer (2014) identifies the preference parameters of a US pension fund and a US insurance fund, and find that the US pension fund has a relatively higher preference for skewness, i.e. higher value of  $\lambda_3$ , while the latter having a relatively higher preference for kurtosis, i.e. higher value of  $\lambda_4$ . Note that the authors do not specify the preference parameter of variance since they construct portfolios with unit variance.

In PGP's second step, the following optimization problem is solved in order to find the optimal weights of the portfolio based on the moment preferences:

$$\min Z(\lambda) = \left| \frac{d_1}{M^*} \right|^{\lambda_1} + \left| \frac{d_2}{V^*} \right|^{\lambda_2} + \left| \frac{d_3}{S^*} \right|^{\lambda_3} + \left| \frac{d_4}{K^*} \right|^{\lambda_4} \quad (5)$$

The investor's preferences can be approximated through polynomial expressions (Kumar et al., 1988) and be expressed as the marginal rate of substitution (MRS) between two moments:

$$MRS_{ij} = \frac{\frac{\delta Z}{\delta(d_i^*)}}{\frac{\delta Z}{\delta(d_j^*)}} = \left[ \frac{\lambda_i}{\lambda_j} \right] * \left[ \frac{d_i^{*\lambda_i-1}}{d_j^{*\lambda_j-1}} \right] \quad (6)$$

The relationship between  $\lambda_i$  and  $\lambda_j$  forms a negative convex indifference curve, and the relative desirability of moment  $i$  can be approximated by varying  $\lambda_i$  in the objective function (5). Thus, a larger value of  $\lambda_i$ , ceteris paribus, indicates a greater importance of moment  $i$  for the

investor. The different combinations of  $\lambda$  enable the investor to specify the investors preferences simultaneously for higher moments in the objective function.

## **2.8 Additional considerations for portfolio construction**

The portfolio selection process requires the investor, in accordance with his leeway, to decide on his preferences, investment universe, investment horizon and frequency of revisions. An institutional investor, such as a pension fund or an insurance fund, often has a restriction for short-selling stocks, and often also requires a diversification and/or a turnover constraint to be followed (Haslem, 2003). Several studies also show that the inclusion of constraints can improve portfolio performance in practice by reducing the estimation error of the mean and variance (Frost and Savarino, 1988). Yet, a too severe constraint might come at the cost of not taking advantage of valuable information, and as a result lead to underperformance (Fabozzi et al., 2010).

### **2.8.1 Buy-hold versus rebalancing**

The asset allocation decision is the most important determinant for a portfolio's return and risk characteristics (Brinson et al., 1995; Davies et al., 2009). Yet, as asset prices change over time the allocation often drifts away from the target allocation, and the portfolio might end up with unwanted characteristics, for instance a too high concentration in one asset class or too high portfolio volatility. To maintain the initial allocation the investor can rebalance the portfolio at a given frequency, i.e. buy assets that have experienced a decline in price and vice versa, to ensure that the portfolio weights are according to the target allocation. The higher the frequency of rebalancing, the less the allocations between the rebalancing periods are allowed to deviate from the target allocation. The opposite of the rebalancing strategy is the buy-hold strategy where the investor passively holds the initial allocation till the end of the investment horizon. There are benefits and disadvantages with both strategies; rebalanced Markowitz portfolios have been found to outperform the buy-hold counterpart empirically, yet after adjusting for the higher costs of using the former the difference is close to negligible (Dayanandan and Lam, 2015).

### 2.8.2 Short-sale constraint

Practitioners often impose a long-only constraint that precludes short position in assets, because the constructed portfolios using sample moments often involve extreme long or short positions. (Ma and Jagannathan, 2001). Especially institutional investors such as pension funds and insurance funds, and nowadays most mutual funds, limit or restrict short positions (Chen et al., 2013). Ma and Jagannathan (2001) show that imposing the short-sale constraint on a portfolio is equivalent to using the sample covariance matrix after reducing its high sample co-estimates by e.g. shrinkage. The high row sums between assets, associated with high covariance, tend to be caused by estimation error and as a result the unconstrained portfolio often receive negative portfolio weights from the high covariance.

### 2.8.3 Diversification constraint

The investor can control the concentration of the portfolio by imposing a diversification constraint. Practitioners also include a diversification constraint because minimize variance does not guarantee appropriately diversified portfolios. The investor can specify the diversification target ( $D_{target}$ ) using the Herfindahl index (7) (Heinze, 2016). A high value of  $\mathcal{H}(x)$  indicates a more diversified portfolio, i.e., an equally-weighted portfolio gives the highest  $\mathcal{H}(x)$ , and a portfolio that is concentrated in one asset gives the value of  $\mathcal{H}(x) = 0$ . The diversification constraint is given by (Richard and Roncalli, 2015):

$$\mathcal{H}(x) = 1 - \sum_{i=1}^N (x_i^2) \quad (7)$$

### 2.8.4 Turnover constraint

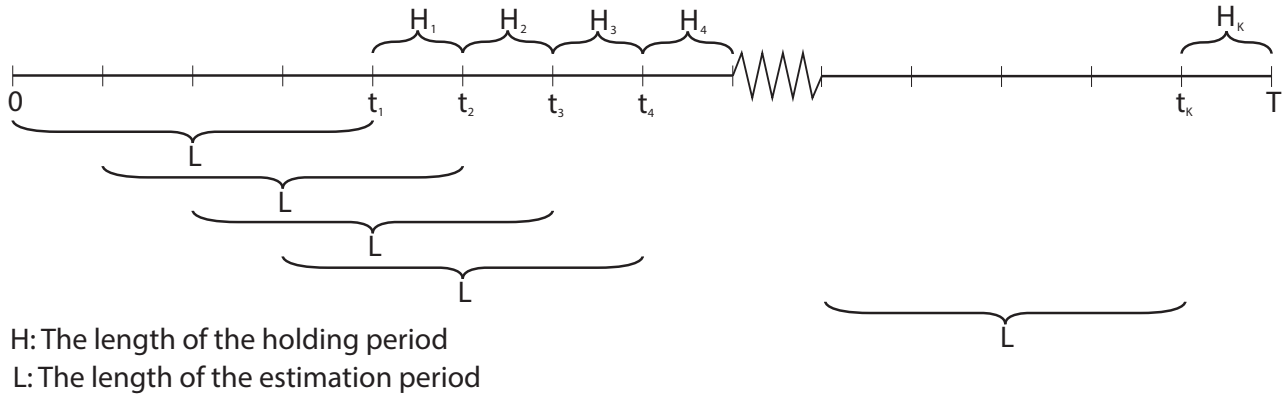
The turnover constraint allows the investor to specify a maximum turnover as churning the portfolio increases costs through brokerage commissions, illiquidity risks, and taxes. Lummer and Reipe (1994) point out that small input changes in the mean-variance framework can result in large changes in the optimized allocated weights. To cope with this, and to limit the costs of drastically changing allocation, we impose a turnover constraint, given by the turnover function of Schreiner (1980):

$$\mathcal{T}(x) = \frac{1}{N} \sum_{i=1}^N |x_{i,t} - x_{i,t-1}| \quad (8)$$

The turnover calculation in this paper abandons the cost approach proposed by DeMiguel et al. (2009), because of the difficulties assessing commission costs. For example, as the transaction size increases, commissions costs might decrease as it is cheaper to make larger orders of liquid stocks, while commission costs can increase if the stocks are illiquid. Furthermore, different tax regulations related to profit and loss of the equities might vary for international investors. From a practical standpoint, a turnover constraint benefits the investors by making selective changes from the existing portfolio that reduce the costs related to turnover for all investors, such as tax, commissions, and other costs related to turnover.

### 3 Methodology

In this section we outline the methodology for the moment-based asset allocation, illustrated in Figure 5. We use the PGP-framework to construct the MV-, MVS- and MVS<sub>K</sub> portfolio, in addition to GMVP. The portfolios' allocations are initially determined at time  $t_1$  and then revised at  $t_2, t_3, \dots, t_K$ , with a constant holding period of length  $H$  following each revision  $n$ . The inputs of the allocation framework for each portfolio revision  $n$  are the forecasted stock moments over the holding period, using Bayesian shrinkage estimates from a rolling window of length  $L$ .



**Figure 5:** Portfolio selection process

Conceptually each portfolio revision is similar and we simplify the notation by presenting a revision at time  $t$  with the corresponding estimation window starting at  $t - L$  and the corresponding holding period ending at  $t + H$ . The portfolio selection process follows the following four steps for the revision at time  $t$ :

1. Forecast moments of each stock in the investment universe for the end of the holding

- period  $t + H$  by Bayesian shrinkage estimation over the time interval  $[t - L, t)$ .
2. Determine the allocation of each portfolio for the holding period  $[t, t + H]$  by solving the optimization problem of PGP at time  $t$ , using the corresponding portfolio moments derived from step 1.
  3. Under the rebalancing strategy each portfolio is rebalanced during the holding period  $\langle t, t + H \rangle$ , at a given frequency, to the allocation determined in step 2, and static during the holding period under the buy-hold strategy.
  4. The portfolio wealth ( $W_t$ ) is computed for each portfolio based on the value of the investments at time  $t$ . The wealth at time  $t + H$  is then reinvested for the following revision, for the respective portfolio.

### 3.1 Assumptions and notation

We follow the standard assumptions made in portfolio theory according to Lai (1991). We assume that the market is perfect with no taxes and transaction costs, and with perfectly divisible assets. The latter implies that the investor can buy and sell any amount of any asset. Furthermore, we assume that all assets have limited liability so that the maximum loss is limited to the total investment. We also assume that short-sale is not allowed, motivated in section 2.8.3. Throughout the methodology section we consider an investment universe with  $N$  risky assets and the corresponding time  $t$  return vector  $R_t = (r_{1,t}, r_{2,t}, \dots, r_{N,t})^T$ . Since our empirical dataset consists of stocks, we compute the holding period return of asset  $i$  at time  $t$  by the dividend-adjusted logarithmic return:

$$r_{i,t} = \ln\left(\frac{p_{i,t} + d_{i,t}}{p_{i,t-1}}\right) \quad \forall i \quad (9)$$

where  $p_{i,t}$  is the price and  $d_{i,t}$  is the dividend, of asset  $i$  at time  $t$ .



### 3.2 Estimation of mean, covariance, coskewness and cokurtosis

The input to the PGP-framework at time  $t$  are based on the true statistical moments at time  $t - 1$ , since we do not have information about the future, given by:

$$\begin{aligned}
\mu_{i,t} &= E_{t-1}[r_{i,t}] && \forall i \\
\sigma_{ij,t} &= E_{t-1}[(r_{i,t} - E_{t-1}[r_{i,t}])(r_{j,t} - E_{t-1}[r_{j,t}])] && \forall i, j \\
s_{ijk,t} &= E_{t-1}[(r_{i,t} - E_{t-1}[r_{i,t}])(r_{j,t} - E_{t-1}[r_{j,t}])(r_{k,t} - E_{t-1}[r_{k,t}])] && \forall i, j, k \\
k_{ijkl,t} &= E_{t-1}[(r_{i,t} - E_{t-1}[r_{i,t}])(r_{j,t} - E_{t-1}[r_{j,t}])(r_{l,t} - E_{t-1}[r_{l,t}])(r_{k,t} - E_{t-1}[r_{k,t}])] && \forall i, j, k, l
\end{aligned} \tag{10}$$

The estimates of (10) are based on the information we have up till time  $t - 1$ , and hence our rolling-window approach used for the revision at time  $t$  are based on  $[t - L, t)$ . Each moment is forecasted by a rolling Bayesian shrinkage estimator based on the linear convex combination of the sample estimate and a shrinkage target. The mean is shrunk towards the mean of the global minimum variance portfolio, and the covariance, coskewness- and cokurtosis matrix are shrunk towards the respective single-index estimates. We apply shrinkage over a rolling window to capture, to some extent, the fact that moments are time-varying.

#### 3.2.1 Sample estimates of moments over a rolling window

The sample estimate is the base case of the Bayesian shrinkage estimator for each moment, i.e. when the shrinkage intensity is zero the full weight is on this estimate. Following the method outlined by Ledoit and Wolf (2003) and Martellini and Ziemann (2010) we use the sample estimates of covariance, coskewness and cokurtosis without Bessel's correction. The rolling sample estimates of the mean, covariance, coskewness and cokurtosis at time  $t$  are given by:

$$\begin{aligned}
\hat{\mu}_{i,t} &= \bar{r}_{i,t} = \frac{1}{L} \sum_{x=t-L}^{t-1} r_{i,x} && \forall i \\
\hat{\sigma}_{ij,t} &= \frac{1}{L} \sum_{x=t-L}^{t-1} (r_{i,x} - \bar{r}_{i,t})(r_{j,x} - \bar{r}_{j,t}) && \forall i, j \\
\hat{s}_{ijk,t} &= \frac{1}{L} \sum_{x=t-L}^{t-1} (r_{i,x} - \bar{r}_{i,t})(r_{j,x} - \bar{r}_{j,t})(r_{k,x} - \bar{r}_{k,t}) && \forall i, j, k \\
\hat{k}_{ijkl,t} &= \frac{1}{L} \sum_{x=t-L}^{t-1} (r_{i,x} - \bar{r}_{i,t})(r_{j,x} - \bar{r}_{j,t})(r_{k,x} - \bar{r}_{k,t})(r_{l,x} - \bar{r}_{l,t}) && \forall i, j, k, l
\end{aligned} \tag{11}$$

where  $r_{i,x}$  is the return of asset  $i$  at time  $x$  and  $L$  denotes the length of the estimation window.

### 3.2.2 Single-index estimates of moments over a rolling window

The single-index model by Sharpe (1963) models returns of an asset by regressing the returns as a dependent variable onto the returns of a broad market index:

$$r_{i,t} = \alpha_i + \beta_i r_{M,t} + \varepsilon_{i,t} \quad \forall i \quad (12)$$

where  $r_{i,t}$  is the return of asset  $i$ ,  $\alpha_i$  and  $\beta_i$  are constants,  $r_{M,t}$  is the return of the market and  $\varepsilon_{i,t}$  is a normally distributed error term with a mean of zero and variance of  $\sigma_{\varepsilon_i}^2$ , of time  $t$ . The single-index model assumes that returns are determined by a systematic part, related to the market, and an unsystematic part related to the specific industry or asset. Each asset's sensitivity towards the market is reflected by  $\beta$  and by our rolling-window approach it is given for asset  $i$  at time  $t$ :

$$\beta_{i,t} = \frac{\sigma_{iM,t}}{\sigma_{M,t}^2} \quad \forall i \quad (13)$$

where  $\sigma_{iM,t}$  is the rolling sample covariance between the returns of asset  $i$  and the market index, and  $\sigma_{M,t}^2$  is the rolling sample variance of the market returns. The explanatory variable  $r_M$  is by construction assumed to be independent of the error term, i.e.  $Cov(R_M, \varepsilon_i) = 0$ . In addition, the key assumption of the single-index model is that the error terms between the assets are independent, i.e.  $E[\varepsilon_i \varepsilon_j] = 0$  for  $i \neq j$ ,  $E[\varepsilon_i \varepsilon_j \varepsilon_k] = 0$  for  $i \neq j \neq k$  and  $E[\varepsilon_i \varepsilon_j \varepsilon_k \varepsilon_l] = 0$  for  $i \neq j \neq k \neq l$ . This assumption implies that assets vary together only through co-movement with the market, and that other industry or firm specific factors only affect the assets individually. As a result of the aforementioned assumptions, we only need the  $\beta$  of each stock and the second, third and fourth moment of the market in order to forecast the systematic covariance-, coskewness- and cokurtosis matrix. To compute the covariance-, coskewness- and cokurtosis matrix, i.e. the sum of the systematic and the corresponding unsystematic matrices, one also needs the unsystematic risk, skewness and kurtosis. We use the single-index estimates computed by Martellini and Ziemann (2010), and adjust them according to the rolling window we apply. The estimates of the moments at time  $t$  are given as:

$$\hat{\sigma}_{ij,t} = \begin{cases} \beta_{i,t}^2 \sigma_{M,t}^2 + e_{i,t}^2 & \forall i = j \\ \beta_{i,t} \beta_{j,t} \sigma_{M,t}^2 & \forall i \neq j \end{cases} \quad (14)$$

$$\widehat{s}_{ijk,t} = \begin{cases} \beta_{i,t}^3 s_{M,t}^3 + e_{i,t}^3 & \forall i = j = k \\ \beta_{i,t}^2 \beta_{k,t} s_{M,t}^3 & \forall i = j \neq k \\ \beta_{i,t} \beta_{j,t} \beta_{k,t} s_{M,t}^3 & \forall i \neq j \neq k \end{cases} \quad (15)$$

$$\widehat{k}_{ijkl,t} = \begin{cases} \beta_{i,t}^4 k_{M,t}^4 + e_{i,t}^4 & \forall i = j = k = l \\ \beta_{i,t}^3 \beta_{l,t} k_{M,t}^4 + 3\beta_{i,t} \beta_{l,t} \sigma_{M,t}^2 e_{i,t}^2 & \forall i = j = k \neq l \\ \beta_{i,t}^2 \beta_{l,t}^2 k_{M,t}^4 + \beta_{i,t}^2 \sigma_{M,t}^2 e_{l,t}^2 + \beta_{l,t}^2 \sigma_{M,t}^2 e_{i,t}^2 + e_{i,t}^2 e_{l,t}^2 & \forall i = j \neq k = l \\ \beta_{i,t}^2 \beta_{k,t} \beta_{l,t} k_{M,t}^4 + \beta_{k,t} \beta_{l,t} \sigma_{M,t}^2 e_{i,t}^2 & \forall i = j \neq k \neq l \\ \beta_{i,t} \beta_{j,t} \beta_{k,t} \beta_{l,t} k_{M,t}^4 & \forall i \neq j \neq k \neq l \end{cases} \quad (16)$$

where  $e_{i,t}^n = \frac{1}{L} \sum_{x=t-L}^{t-1} \varepsilon_{i,x}^n$  denotes the centered  $n$ th-moment of the error term for asset  $x = i, j, k, l$ , and  $\sigma_M^2$ ,  $s_M^3$ , and  $k_M^4$  denote the variance, skewness and kurtosis of the market index, respectively.

### 3.2.3 Bayesian shrinkage of mean over a rolling window

Jorion (1986) proposes the empirical Bayes-Stein estimator as a robust estimate of the mean. The method is based on shrinking the sample estimate of mean, with high estimation error and low bias, towards a grand mean. The GMVP serves as the grand mean, i.e. the shrinkage target, since it does not rely on mean forecasts and is thus less vulnerable to estimation error. The rolling shrinkage estimate of mean is given by the weighted average of the sample estimate and the GMVP-estimate:

$$\widehat{\mu}_t = \omega_{1,t} \mu_{mvp,t} \mathbf{1}_N + (1 - \omega_{1,t}) \mu_{S,t} \quad \omega_{1,t} \in [0, 1] \quad (17)$$

where  $\omega_{1,t}$  is the shrinkage intensity at time  $t$ ,  $\mu_{mvp,t}$  is the mean of the minimum variance portfolio at time  $t$ ,  $\mathbf{1}_N$  is a  $N \times 1$  vector of ones and  $\mu_{S,t}$  is the  $N \times 1$  sample mean vector from (11). Note that the weight vector of GMVP is obtained by solving (38) with only the full-investment constraint imposed, and the mean of the portfolio at time  $t$  is the weighted average of the weight vector and the return vector, shown in formula (27). Adjusting the optimal shrinkage intensity, derived by Jorion (1986), according to our notation and rolling-window

approach, the shrinkage intensity for a portfolio of  $N$  assets is given by:

$$\widehat{\omega}_{1,t} = \frac{N + 2}{(N + 2) + (\mu_{S,t} - \mu_{mvp,t}1)' L \Sigma_t^{-1} (\mu_{S,t} - \mu_{mvp,t}1)} \quad (18)$$

where  $\Sigma$  is the true covariance matrix. Since  $\Sigma$  is unknown Jorion (1986) proposes the following estimate:

$$\widehat{\Sigma}_t = \frac{L}{L - N - 2} \Sigma_t^S \quad (19)$$

where  $\Sigma_t^S$  is given by the covariance matrix from (11). The expression of the shrinkage intensity (18) provides the following insight of the trade-off between the sample estimate and the shrinkage target estimate; for a longer estimation window, i.e.  $L$  increases, more weight is put on the sample estimate.

### 3.2.4 Bayesian shrinkage estimate of covariance, coskewness and cokurtosis, over a rolling window

The Bayesian shrinkage approach proposed by Ledoit and Wolf (2003), is based on a robust estimator that is the weighted average of the the sample- and single-index estimate of the covariance matrix. Martellini and Ziemann (2010) extended the Bayesian shrinkage estimator to the coskewness- and cokurtosis matrix, and we have used the authors' framework and only modified the notation as we implemented the rolling version. The implementation in R can be found in Listing 2. The rolling shrinkage estimator for the covariance-, coskewness- and cokurtosis matrix, is given by, respectively;

$$\widehat{\Sigma}_t = \omega_{2,t} \Sigma_t^{SI} + (1 - \omega_{2,t}) \Sigma_t^S \quad \omega_{2,t} \in [0, 1] \quad (20)$$

$$\widehat{\Phi}_t = \omega_{3,t} \Phi_t^{SI} + (1 - \omega_{3,t}) \Phi_t^S \quad \omega_{3,t} \in [0, 1] \quad (21)$$

$$\widehat{\Psi}_t = \omega_{4,t} \Psi_t^{SI} + (1 - \omega_{4,t}) \Psi_t^S \quad \omega_{4,t} \in [0, 1] \quad (22)$$

where  $\omega_{n,t}$  is the time  $t$  shrinkage intensity for moment  $n$ , for  $n = 2, 3, 4$ , S denotes a matrix containing the sample estimates from section 3.2.1 and SI denotes a matrix with the single-index estimates from section 3.2.2. Ledoit and Wolf (2003) derives the optimal shrinkage intensity for the covariance matrix by minimizing a squared loss function, and (Martellini and Ziemann, 2010) analogously do the similar procedure to derive the shrinkage intensity for the coskewness-

and cokurtosis matrix. The optimal shrinkage intensity for the covariance-, coskewness- and cokurtosis matrix, at time  $t$ , is given by:

$$\omega_{n,t}^* = \max\{0, \min\{\frac{\hat{\pi}_{n,t} - \hat{\rho}_{n,t}}{L\gamma_{n,t}}, 1\}\} \quad \text{for } n = 2, 3, 4 \quad (23)$$

where  $\hat{\pi}_{n,t}$  is the sum of the asymptotic variance of the sample estimates for moment  $n$ ,  $\hat{\rho}_{n,t}$  is the sum of the asymptotic covariance of the single-index estimates and sample estimates for moment  $n$ , and  $\hat{\gamma}_{n,t}$  is a measure of the misspecification of the single-index estimate. Note that the shrinkage intensity has a floor value of 0 and a cap of 1 because the shrinkage estimator is constructed as a weighted average between the sample estimate and the shrinkage target, but in practice minimizing the aforementioned loss function might result in intensities outside  $[0, 1]$ . The closed-form expressions of  $\hat{\pi}_{n,t}$ ,  $\hat{\rho}_{n,t}$  and  $\hat{\gamma}_{n,t}$  are given in appendix A.

### 3.3 Portfolio moments

We adopt the notation of Xu et al. (2008) for the notation of portfolio moments. The estimates of the portfolio moments at time  $t$ , at time  $t - 1$ , are based on the expected portfolio mean, -variance, -skewness and -kurtosis, respectively given by:

$$\begin{aligned} M_{P,t} &= E_{t-1}[R_{P,t}] \\ V_{P,t} &= E_{t-1} [(R_P - E_{t-1}[R_P])^2] \\ S_{P,t} &= E_{t-1} [(R_P - E_{t-1}[R_P])^3 / (V_{P,t}^{3/2})] \\ K_{P,t} &= E_{t-1} [(R_P - E_{t-1}[R_P])^3 / (V_{P,t}^2)] \end{aligned} \quad (24)$$

The percentage wealth invested in asset  $i$  at time  $t - 1$  is denoted by  $x_{i,t-1}$ , and the portfolio weights of time  $t - 1$  are stored in the transposed weight vector  $X_{t-1}^T = (x_{1,t-1}, x_{2,t-1}, \dots, x_{N,t-1})$ . Each asset contributes to the portfolio return at time  $t$  by the product of the return of the asset at time  $t$  and the weight allocated to the asset at time  $t - 1$ . The return of portfolio  $P$  at time  $t$  is then a linear combination of the weighted returns of the  $N$  assets, given by:

$$R_{P,t} = \sum_{i=1}^N x_{i,t-1} r_{i,t} = X_{t-1}^T R_t \quad (25)$$

The mean of portfolio  $P$  at time  $t$  is a weighted linear combination of the mean of each asset  $i$ ;

$$M_{P,t} = \sum_{i=1}^N x_{i,t-1} \mu_{i,t} = X_{t-1}^T \mu_t \quad (26)$$

where  $\mu_{i,t}$  is estimated by the rolling shrinkage method from section 3.2.3 and scaled by  $H$ , due to the additivity property of logarithmic returns, in order to match the frequency of the holding period. The  $N \times 1$ -vector  $\mu_t$  contains the forecasts of mean of time  $t$  for all the stocks.

The variance of portfolio  $P$  consists of the variance of each asset  $i$ ,  $\sigma_i^2 = \sigma_{ii}$ , and the covariance between each combination of asset  $i$  and  $j$ ,  $\sigma_{ij}$ ;

$$V_{P,t} = \sum_{i=1}^N \sum_{j=1}^N x_{i,t-1} x_{j,t-1} \sigma_{ij,t} = X_{t-1}^T \Sigma_t X_{t-1} \quad (27)$$

where  $\sigma_{ij,t}$  is estimated by the rolling-shrinkage method from section 3.2.4 and scaled by  $H$  in order to match frequency of the estimated risk to the frequency of the holding period. This is equal to summing the daily forecasts of the covariance over  $H$  days assuming each day is similar. The  $\Sigma_t$  is a  $N \times N$ -matrix containing the variances diagonally and the covariances elsewhere;

$$\Sigma_t = \begin{bmatrix} \sigma_{11,t} & \sigma_{12,t} & \dots & \sigma_{1N,t} \\ \sigma_{21,t} & \sigma_{22,t} & \dots & \sigma_{2N,t} \\ \dots & \dots & \dots & \dots \\ \sigma_{N1,t} & \sigma_{N2,t} & \dots & \sigma_{NN,t} \end{bmatrix} \quad (28)$$

We store the coskewness- and cokurtosis elements in a  $N \times N^2$ - and  $N \times N^3$ -matrix, respectively, following the work of Athayde and Flores (2004). By stacking the higher-moment elements column-wise, we work with two dimensional matrices as shown in (31) and (34). We do not scale the higher-moments because there is no systematic pattern between skewness and kurtosis of daily and annual returns (Jondeau et al., 2007). For instance, Jondeau et al. (2007) shows empirically that daily skewness of S&P500 is larger than the annual skewness, while the opposite is true for skewness of FT-SE.

The unstandardized skewness of portfolio  $P$  consists of the sum of the skewness of each asset  $i$ ,  $s_i^3 = s_{iii}$ , the coskewness between each permutation of asset  $i$  and  $j$ ,  $s_{ijj}$ , and the coskewness

between each combination of asset  $i$ ,  $j$  and  $k$ ,  $s_{ijk}$ ;

$$S'_{P,t} = \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N x_{i,t-1} x_{j,t-1} x_{k,t-1} s_{ijk,t} = X_{t-1}^T \Phi_t (X_{t-1} \otimes X_{t-1}) \quad (29)$$

where  $\otimes$  is the kronecker product and  $\Phi$  is a  $N \times N^2$ -matrix:

$$\Phi_t = \begin{bmatrix} S_{1,t} & | & S_{2,t} & | & \dots & | & S_{N,t} \end{bmatrix} \quad (30)$$

with

$$S_{m,t} = \begin{bmatrix} s_{m11,t} & s_{m12,t} & \dots & s_{m1N,t} \\ s_{m21,t} & s_{m22,t} & \dots & s_{m2N,t} \\ \dots & \dots & \dots & \dots \\ s_{mN1,t} & s_{mN2,t} & \dots & s_{mNN,t} \end{bmatrix} \quad (31)$$

The unstandardized kurtosis of portfolio P consists of the weighted sum of the kurtosis of each asset  $i$ ,  $k_i^4 = k_{iiii}$ , the cokurtosis between each permutation of asset  $i$  and  $j$ ,  $k_{iiij}$  and  $k_{iijj}$ , each permutation of asset  $i$ ,  $j$  and  $k$ ,  $k_{iijk}$ , and each combination of asset  $i$ ,  $j$ ,  $k$  and  $l$ ,  $k_{ijkl}$ ;

$$K'_{P,t} = \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \sum_{l=1}^N x_{i,t-1} x_{j,t-1} x_{k,t-1} x_{l,t-1} k_{ijkl,t} = X_{t-1}^T \Psi_t (X_{t-1} \otimes X_{t-1} \otimes X_{t-1}) \quad (32)$$

where  $\otimes$  is the kronecker product and  $\Psi_t$  is a  $N \times N^3$ -matrix:

$$\Psi_t = \begin{bmatrix} K_{11,t} & K_{12,t} & \dots & K_{1N,t} & | & K_{21,t} & K_{22,t} & \dots & K_{2N,t} & | & \dots & | & K_{N1,t} & K_{N2,t} \dots & K_{NN,t} \end{bmatrix} \quad (33)$$

with

$$K_{mg,t} = \begin{bmatrix} k_{mg11,t} & k_{mg12,t} & \dots & k_{mg1N,t} \\ k_{mg21,t} & k_{mg22,t} & \dots & k_{mg2N,t} \\ \dots & \dots & \dots & \dots \\ k_{mgN1,t} & k_{mgN2,t} & \dots & k_{mgNN,t} \end{bmatrix} \quad (34)$$

While the rolling-shrinkage forecasts are based on unstandardized coskewness and cokurtosis, i.e. the third and fourth moment, we apply the standardized moments in the optimization framework. We standardize portfolio skewness and -kurtosis by dividing the respective moments by the portfolio variance to the power of 1.5 and 2, respectively. The standardized portfolio

skewness and -kurtosis at time  $t$  are then given by:

$$S_{P,t} = \frac{X_{t-1}^T \Phi_t(X_{t-1} \otimes X_{t-1})}{(X_{t-1}^T \Sigma_t X_{t-1})^{3/2}} \quad (35)$$

$$K_{P,t} = \frac{X_{t-1}^T \Psi_t(X_{t-1} \otimes X_{t-1} \otimes X_{t-1})}{(X_{t-1}^T \Sigma_t X_{t-1})^2} \quad (36)$$

### 3.4 Asset allocation by Polynomial goal programming

This section describes the PGP-framework that we use to construct the MVS- and MVSK portfolio. Following the discussion of PGP in the literature review we also construct the MV portfolio by this framework. We also describe the optimization techniques we use to solve the convex and the non-convex parts of PGP.

#### 3.4.1 Polynomial goal programming

PGP consists of two subsequent steps, G1 and G2. In the first step, G1, the optimal values of the portfolio moments are found by solving the optimization problems separately. Step G1a is the maximization problem of mean, step G1b is the minimization problem of variance, step G1c is the maximization problem of skewness and G1d is the minimization problem of kurtosis. Hence an investor solely focusing on one moment would determine his allocation by solving the corresponding optimization procedure among G1a-G1d. The optimization problems are presented in (37)-(40), and each includes the full-investment constraint and the long-only constraint, and an additional set of constraints represented by  $C$ .  $C$  contains the diversification constraint and the turnover constraint, and the format of the constraints are shown in section 3.4.2

$$G1a = \left\{ \begin{array}{l} \underset{X_{t-1}}{\text{maximize}} \quad M_t^* = X_{t-1}^T \mu_t \\ \text{subject to} \\ \\ X_{t-1}^T I = 1 \\ X_{t-1} \succeq 0 \\ X_{t-1} \in C \end{array} \right. \quad (37)$$



$$G1b = \begin{cases} \text{minimize}_{X_{t-1}} & V_t^* = X_{t-1}^T \Sigma_t X_{t-1} \\ \text{subject to} & \\ & X_{t-1}^T I = 1 \\ & X_{t-1} \succeq 0 \\ & X_{t-1} \in \mathcal{C} \end{cases} \quad (38)$$

$$G1c = \begin{cases} \text{maximize}_{X_{t-1}} & S_t^* = \frac{X_{t-1}^T \Phi_t(X_{t-1} \otimes X_{t-1})}{(X_{t-1}^T \Sigma_t X_{t-1})^{3/2}} \\ \text{subject to} & \\ & X_{t-1}^T I = 1 \\ & X_{t-1} \succeq 0 \\ & X_{t-1} \in \mathcal{C} \end{cases} \quad (39)$$

$$G1d = \begin{cases} \text{minimize}_{X_{t-1}} & K_t^* = \frac{X_{t-1}^T \Psi_t(X_{t-1} \otimes X_{t-1} \otimes X_{t-1})}{(X_{t-1}^T \Sigma_t X_{t-1})^2} \\ \text{subject to} & \\ & X_{t-1}^T I = 1 \\ & X_{t-1} \succeq 0 \\ & X_{t-1} \in \mathcal{C} \end{cases} \quad (40)$$

The maximization problem of portfolio mean (37) and the minimization problem of portfolio variance (38), are convex problems with non-linear constraints without any additional constraints. Since the objective function of mean is linear, the problem can be solved by spending the budget on the asset with the highest mean. To solve the maximization problem of mean with additional constraints, and the minimization problem of variance with and without additional constraints, we use the interior-point algorithm, described in section 3.4.3.

The maximization problem of portfolio skewness (39) and the minimization problem of portfolio kurtosis (40), involve cubic and quadratic objective functions. Thus step G1c and G1d may have several local maxima and several local minima, and the problems are non-convex. To avoid suboptimal values of skewness and kurtosis, we apply a global optimization procedure, outlined in section 3.4.4.

The optimal moments of time  $t$  are found by solving G1a, G1b, G1c and G1d, and then used in the second step of PGP, G2. The four independent subproblems are consolidated into the

objective function  $Z_t(\lambda) = Z_t(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ , where  $\lambda_i$  is the preference parameter for moment  $i$ , where  $i = 1$  is the portfolio mean,  $i = 2$  is the portfolio variance,  $i = 3$  is the portfolio skewness and  $i = 4$  is the portfolio kurtosis. We denote the difference between the actual and optimal moment  $i$ , of time  $t$ , by  $d_{i,t}$ . The objective function  $Z_t(\lambda)$  is minimized in step G2, and the optimization problem is given by:

$$\begin{aligned}
& \underset{X_{t-1}}{\text{minimize}} && Z_t(\lambda) = \left| \frac{d_{1,t}}{M_t^*} \right|^{\lambda_1} + \left| \frac{d_{2,t}}{V_t^*} \right|^{\lambda_2} + \left| \frac{d_{3,t}}{S_t^*} \right|^{\lambda_3} + \left| \frac{d_{4,t}}{K_t^*} \right|^{\lambda_4} \\
& \text{subject to} && \\
& && X_{t-1}^T \mu_t + d_{1,t} = M_t^* \\
& && X_{t-1}^T \Sigma_t W_{t-1} - d_{2,t} = V_t^* \\
& && \frac{X_{t-1}^T \Phi_t (X_{t-1} \otimes X_{t-1})}{(X_{t-1}^T \Sigma_t X_{t-1})^{3/2}} + d_{3,t} = S_t^* \\
& && \frac{X_{t-1}^T \Psi_t (X_{t-1} \otimes X_{t-1} \otimes X_{t-1})}{(X_{t-1}^T \Sigma_t X_{t-1})^2} + d_{4,t} = K_t^* \\
& && d_t \succeq 0 \\
& && X_{t-1}^T I = 1 \\
& && X_{t-1} \succeq 0 \\
& && X_{t-1} \in \mathcal{C}
\end{aligned} \tag{41}$$

where  $d_t = (d_{1,t}, d_{2,t}, d_{3,t}, d_{4,t})$ . By solving  $Z_t(\lambda)$  we obtain the time-varying weight vector  $X_{t-1}^T$ , according to the specified preference parameters. In order to solve (41) we apply the global optimization algorithm DEoptim, described in section 3.4.6.

We obtain the allocation of GMVP by solving step G1b as the portfolio is only based on minimization of variance. As pointed out in the literature review regarding PGP, is the traditional mean-variance efficient portfolio obtained by minimizing  $Z_t(1, 1, 0, 0)$ . The mean-variance-skewness efficient portfolio is obtained by minimizing  $Z_t(\lambda_1, \lambda_2, \lambda_3, 0)$  for any  $\lambda_1 \geq 0$ ,  $\lambda_2 > 0$  and  $\lambda_3 > 0$ . The mean-variance-skewness-kurtosis efficient portfolio is obtained by minimizing  $Z_t(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$  for any  $\lambda_1 \geq 0$ ,  $\lambda_2 > 0$ ,  $\lambda_3 > 0$  and  $\lambda_4 > 0$ .

### 3.4.2 Additional constraints used in Polynomial goal programming

All the portfolios we construct have the full-investment constraint that forces the full budget to be used on at least one of the  $N$  risky assets, and the long-only constraint that precludes short-

positions, i.e. negative weights. To obtain more practical portfolios we also construct portfolios with the diversification constraint, the turnover constraint and both. The implementation of the optimization problems with these constraints in R is based on the matrix notation of the constraints formulas presented in the literature review, in section 2.8.3 and 2.8.4. The turnover constraint for the portfolio optimization problems at time  $t$  is given by:

$$\left(\frac{1}{N}\right)I^T|X_{t-1} - X_{t-2}| \leq c_1 \quad (42)$$

where  $c_1$  denotes the maximum allowed turnover and  $I^T$  is a transposed  $N \times 1$ -vector of 1s. Note that initial portfolios, i.e. portfolios constructed at  $t_1$ , are not affected by this constraint. The diversification constraint for the portfolio optimization problems at time  $t$  is given by

$$N - X_{t-1}^T X_{t-1} \geq c_2 \quad (43)$$

where  $c_2$  denotes the minimum required portfolio diversification.

### 3.4.3 Convex optimization by the interior-point method

We utilize the interior-point-algorithm to solve the convex nonlinear problems of maximization of portfolio mean (37), step G1a, and minimization of portfolio variance (38), step G1b. The implementation in R is based on the package Rmosek by Friberg (2014). The package solves the following primal problem by minimization:

$$\begin{aligned} & \underset{X_{t-1}}{\text{minimize}} && \frac{1}{2}X_{t-1}^T Q X_{t-1} + c^T X_{t-1} \\ & \text{subject to} && l_i^c \leq a_i^T X_{t-1} + z_i(X_{t-1}) \leq u_i^c \\ & && l_x^c \leq X_{t-1}^T \leq u_x^c \end{aligned} \quad (44)$$

where  $c_t$ ,  $a_i$ ,  $l_x^c$  and  $u_x^c$  are  $N \times 1$ -vectors,  $Q_t$  is a  $N \times N$  positive semi-definite matrix, and  $z_i(X_{t-1})$  a function given by:

$$z_i(X_{t-1}) = \sum_{k=1}^N f(x_{k,t} + h)^g \quad (45)$$

Since the format is minimization we specify the objective function of mean as  $c = -\mu_t$ , and  $Q = 0$ , where  $\mu_t$  is the forecasted time  $t$   $N \times 1$ -mean vector from section X. When minimizing portfolio variance the objective function is  $Q = 2\Sigma_t$  and  $c = 0$ , where  $\Sigma_t$  is the forecasted time  $t$   $N \times N$ -covariance matrix from section Y. The same constraints apply to the optimization

problem of mean and of variance. The base case constraints are the full-investment- and long-only constraint. In addition we apply the turnover constraint (42) and the diversification constraint (43), separately and simultaneously. Note that the turnover constraint is modified in order to keep the optimization problems convex, as shown in appendix B.

Listing 4 and Listing 5 in appendix G show the implementation in R for solving the mean- and variance problem, respectively, for all the different sets of constraints considered. We present the implementation of the turnover- and diversification constraint, simultaneously, by matrix notation. The weights for each asset  $x_{i,t-1}$  is stored together with the percent-wise amount bought,  $b_{i,t-1}$ , and sold,  $s_{i,t-1}$ , of asset  $i$ ,  $\forall i$ , in a  $3Nx1$ -vector given by:

$$X_{t-1}^T = \begin{bmatrix} x_{1,t-1} & \dots & x_{N,t-1} & b_{1,t-1} & \dots & b_{N,t-1} & s_{1,t-1} & \dots & s_{N,t-1} \end{bmatrix}$$

The objective function of mean and variance only depends on asset weights, i.e. the mean and covariances of  $b_i$  and  $s_i$  are 0, implemented by matrix partitioning. Thus, the objective function when maximizing mean is based on the following vector  $c$  while  $Q = 0$ , and the objective function when minimizing variance is based on the following matrix  $Q$  while  $c = 0$ :

$$c = \begin{bmatrix} -\mu_t \\ \mathbf{0}_{2N} \end{bmatrix} \quad Q = \begin{bmatrix} \mathbf{2}\Sigma_t & \mathbf{0}_{(Nx2N)} \\ \mathbf{0}_{(2NxN)} & \mathbf{0}_{(2Nx2N)} \end{bmatrix} \quad (46)$$

where  $\mathbf{0}_{2N}$  is a  $2Nx1$ -null vector and  $\mathbf{0}_{(2Nx2N)}$  is a  $2Nx2N$ -null matrix. The constraint matrix, denoted  $A$ , is together with the lower- and upper box constraint vectors,  $l^c$  and  $u^c$ , adjusted to incorporate the turnover and diversification constraints in addition to the full investment constraint:

$$A = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \mathbf{a}_3^T \\ \mathbf{a}_4^T \end{bmatrix} \quad l^c = \begin{bmatrix} \mathbf{l}_1 \\ \mathbf{l}_2 \\ \mathbf{l}_3 \\ \mathbf{l}_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \mathbf{X}_N^* \\ -\mathbf{Inf} \end{bmatrix} \quad u^c = \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \mathbf{u}_3 \\ \mathbf{u}_4 \end{bmatrix} = \begin{bmatrix} 1 \\ N \times c_1 \\ \mathbf{X}_N^* \\ 1 - c_2 \end{bmatrix} \quad (47)$$

where  $\mathbf{X}_N^*$  is a  $N \times 1$ -vector of last period's weights,  $N \times c_1$  is the maximum allowed portfolio turnover,  $c_2$  is the minimum allowed diversification,  $\mathbf{a}_1^T$  is a transposed  $3N \times 1$ -vector for the full investment constraint:

$$\mathbf{a}_1^T = \begin{bmatrix} \mathbf{1}_N & \mathbf{0}_{2N} \end{bmatrix},$$

The first part of the turnover constraint,  $\mathbf{a}_2^T$ , is a transposed  $3N \times 1$ -vector representing the left-hand side of formula (76) in appendix:

$$\mathbf{a}_2^T = \begin{bmatrix} \mathbf{0}_N & \mathbf{1}_{2N} \end{bmatrix},$$

The second part of the turnover constraint,  $\mathbf{a}_3^T$ , is a transposed  $3N \times N$ -matrix representing the left-hand side of formula (75) in appendix, for all  $N$  assets:

$$\mathbf{a}_3^T = \begin{bmatrix} \mathbf{I}_N & -\mathbf{I}_N & \mathbf{I}_N \end{bmatrix}$$

where  $\mathbf{I}_N$  is a  $N \times N$ -identity matrix. We denote the  $3N \times 1$ -null vector  $\mathbf{a}_4$ , and let  $f = 1$ ,  $h = 0$  and  $g = 2$ , so that the left-hand side of the diversification constraint (43) is given by:

$$z_4(X_{t-1}) = x_{1,t-1}^2 + x_{2,t-1}^2 + \dots + x_{N,t-1}^2 \quad (48)$$

Finally we specify the short-sale constraint and the lower limit of the amounts bought and sold, i.e.  $b_{i,t-1}$  and  $s_{i,t-1}$ , by setting the lower box constraint as the  $3N \times 1$ -null vector  $l_x^c$ .

### 3.4.4 Non-convex optimization by differential evolution

Maximization of skewness (39), minimization of kurtosis (40) and step 2 of PGP (41) are non-convex optimization problems with several local optima, hence a global optimizer must be applied to avoid allocations based on sub-optimal solutions.. In order to solve the aforementioned optimization problems we apply the global optimizer DEoptim by Ardia et al. (2011). DEoptim implements the stochastic, evolutionary algorithm known as Differential Evolution (DE) proposed by Price et al. (2005). The differential evolution algorithm is developed for solving single-objective unconstrained optimization problems that are not necessarily continuous nor differentiable. The process of DEoptim follows the stages of the differential evolutionary algorithm; the mutation stage, the crossover stage and the selection stage. Since we use an index to describe both the position in population and what generation each each vector belongs to, we remove the time-index to simplify the notation. The concept of solving the non-convex

problems are still the same for each  $t$ .

The current population can be describe by  $P_{z,g}(X_{z,g}^T)$  and consists of NP parameter vectors,  $X_{z,g}^T$ , each containing  $N$  asset weights, where  $z = 0, \dots, (NP - 1)$  and  $g = 0, \dots, g_{max}$ . The parameter vector  $X_{z,g}$  is the  $N \times 1$  transposed weight vector  $X_{z,g}^T = (x_{1,z,g}, x_{2,z,g}, \dots, x_{N,z,g})$ , where  $z$  indicates the index of the vector within the population of generation  $g$ . We set the population size by  $NP = 15N$  and amount of generations by  $g_{max} = 20.000$ . Initially, i.e.  $g = 0$ , random weights for each parameter vector are generated by  $w_{i,z,g} = rand_i(0, 1) * (upper - lower) + lower$  for each asset  $i$ ,  $i = 1, 2, \dots, N$ . The lower and upper boundaries are set to 0 and 1, respectively, as we do not allow for short-sale, and  $rand_i(0, 1)$  denotes a value drawn from the uniform distribution between 0 and 1, for asset  $i$ .

In the mutation stage a mutant vector is created by the sum of three parameter vectors, based on randomly chosen vectors from the current population. A mutation strategy is then applied to maintain the diversity of one generation of portfolio weights to the next generation, and DEoptim allows the user to select among 6 different strategies for spawning the mutant vector. According to Ardia et al. (2014), and our own experimenting, we utilize strategy 6 with adaptive parameter control, known as the "DE/current-to-pbest/1" strategy. The strategy gives the following mutant vector Sanderson and Zhang (2009):

$$v_{z,g}^T \doteq X_{z,g}^T + F_z(X_{best,g}^T - X_{z,g}^T) + F_z(X_{r1,g}^T - X_{r2,g}^T) \quad (49)$$

where  $X_{z,g}^T$ ,  $X_{r1,g}^T$  and  $X_{r2,g}^T$ , are three parameter vectors from the population selected at random, and  $X_{best,g}^T$  is the adaptive parameter control and  $F_z = F$  is a fixed mutation factor that defaults to 0.5. Note that the adaptive parameter control introduces learning by mutating one of the randomly chosen top [100p%] best solutions and uses information of the average mutation factor of all successful mutations to the next generation  $g$ . The adaptive parameter in the mutation operation increases the models reliability because it alleviate the consequence of fast premature convergence from only using the information from the best solution.

In the crossover stage trial vectors are created by combining generated asset weights from different vectors; a crossover between asset weights from a parameter vector and a mutant vector. The trial vector is given by:

$$X_{z,g}^{*T} = (x_{1,z,g}^*, x_{2,z,g}^*, \dots, x_{N,z,g}^*) \quad (50)$$

where  $x_{i,z,g}^*$  denotes the weight of asset  $i$  from parameter vector  $z$  for generation  $g$ .

$$x_{i,z,g}^* = \begin{cases} v_{i,z,g} & \text{if } \{rand(0,1) \leq CR \wedge i = i_{rand}\} \\ x_{i,z,g} & \text{otherwise} \end{cases} \quad \forall i \quad (51)$$

where  $CR$ , by default  $CR = 0.5$ , is applied to each mutation vector  $v_{z,g}$  in the population for each asset weight and each parameter vector, and controls the fraction of the parameter values that are copied from the mutant. Note that  $i_{rand} = randint(1, N)$ , ie. we select a stock based on the random index.

After the mutation stage and crossover stage we ensure that each asset weight is feasible according to the long-only constraint, by:

$$x_{i,z,g} = \begin{cases} 1 - rand_i(0,1) & \text{for } x_{i,z,g} > 1 \\ 0 + rand_i(0,1) & \text{for } x_{i,z,g} < 0 \end{cases} \quad \forall i \quad (52)$$

In the selection stage we compare the target vector  $X_{z,g}^T$  and trial vector  $X_{z,g}^{*T}$ , and the vector that minimizes the objective function  $f$  takes place in the next generation  $g + 1$ , for each  $z = 0, 1, \dots, (NP - 1)$ ;

$$X_{z,g+1}^T = \begin{cases} X_{z,g}^{*T} & \text{if } f(X_{z,g}^{*T}) < f(X_{z,g}^T) \\ X_{z,g}^T & \text{otherwise} \end{cases} \quad (53)$$

where  $f(\cdot)$  is a static penalty function incorporating the full-investment constraint, and the diversification- and/or turnover constraint if imposed. When all NP vectors have been evaluated for  $g + 1$  in (53), the mutation-, crossover- and selection stage, are repeated for  $g + 2$ ,  $g + 3$ , etc., until the optimum is located or maximum generations have been reached.

In (54) we use penalty functions to incorporate the full-investment constraint, and the additional diversification- and turnover constraint, by adding penalty terms for each constraint considered to the cost function  $f$ . Thus a constrained optimization problem is transformed to an unconstrained one. Note that there is a trade-off when setting the value of the penalty  $p$ , since a too large penalty may lead to neglecting the objective function, while a too low penalty may lead to an infeasible portfolio. We use the default penalty value  $p = 1000$  and have confirmed that it is adequate for our purpose. The penalty function incorporating a constraint is given by:

$$f(X^T) = Y + p * \text{constraint}(s) \quad (54)$$

where  $Y$  represents the objective function for portfolio skewness when solving step G1c of

PGP, portfolio kurtosis when solving step G1d of PGP, and  $Z_t(\lambda)$  when solving step G2 of PGP. We denote the penalty of the full investment constraint by *constraint*<sub>1</sub>, the penalty of the diversification constraint by *constraint*<sub>2</sub> and the penalty of the turnover constraint by *constraint*<sub>3</sub>;

$$\begin{aligned}
\text{constraint}_1 &= \max \left\{ \left[ \sum_{i=1}^N x_{i,t-1} - 1.01 \right], \left[ 0.99 - \sum_{i=1}^N x_{i,t-1} \right] \right\} \\
\text{constraint}_2 &= \max \left\{ \left[ c_2 - \left| 1 - \sum_{i=1}^N x_{i,t-1}^2 \right| \right], 0 \right\} \\
\text{constraint}_3 &= \max \left\{ \left[ \frac{1}{N} \sum_{i=1}^N |x_{i,t-1} - x_{i,t-2}| - c_1 \right], 0 \right\}
\end{aligned} \tag{55}$$

for  $t = 1, 2, \dots, T$ . Note that we add slack to the full investment constraint according to the recommendation of Peterson et al. (2015), in order to enable DEoptim to generate a sufficient number of feasible portfolios. In fact, a strict constraint of 1 may reduce the number of feasible portfolios by 1/3. Afterwards we normalize the weight vector to ensure the budget constraint is met.

### 3.5 Performance evaluation

We evaluate the performance of the portfolios constructed from the first allocation to the end of the last holding period, i.e. over the period  $[t_1, T]$ , where  $T = t_K + H$ . The performance evaluation is twofold as we use traditional measures such as accumulated returns and the Sharpe-ratio, and in addition compare the out-of-sample moments of each portfolio to the benchmark portfolios, by conducting a difference test using heteroscedastic and autocorrelation robust (HAC) kernel estimation.

#### 3.5.1 Traditional measures

The first evaluation criterium of the portfolios are based on portfolio wealth. For a \$1 investment in portfolio P, the wealth of portfolio  $P$  at time

$$W_{P,T} = \prod_{t=1}^T (1 + R_{P,t}) \tag{56}$$



where the return of the portfolio at time  $t$  is:

$$R_{P,t} = \left( \frac{p_t q_t - p_{t-1} q_{t-1}}{p_{t-1} q_{t-1}} \right) \quad (57)$$

where  $p_t$  and  $q_t$  are the  $N \times 1$ -price vector and quantity vector containing each asset price and number of shares, respectively, at time  $t$ . Note that for the buy-hold portfolio we have a constant number of shares during the holding period, i.e.  $q_t = q$  for a given holding period, and that it  $q$  is changing between each revision. For the daily rebalanced portfolio the relative wealth allocated to each stock is the same during the holding period, but this changes during each revision as the target allocation changes.

We also evaluate the performance of the constructed portfolios by traditional criteria, extensively explained in the finance literature, in order to capture the portfolios risk-adjusted return, risk profile and practicability. We cover the risk profile of each portfolio by computing the annualized standard deviation and expected shortfall. The expected shortfall formula, also known as conditional Value-at-risk (VaR), is based on (Tsay, 2013) formula:

$$ES_{1-p} = E(R_{P,t} | R_{P,t} > VaR) = \frac{\int_{VaR}^{\infty} r f(r) dx}{Pr(R_{P,t} > VaR)} \quad (58)$$

The expected shortfall is the expected loss of  $R_{P,t}$  given that  $R_{P,t}$  exceeds its VaR (Tsay, 2013) and is the average value of all the values exceeding the VaR.

The risk-adjusted return of the portfolio is measured by the annualized daily Sharpe-ratio:

$$SR_P = \frac{R_P - r_f}{\sigma_P} \quad (59)$$

where  $r_f$  is the risk-free rate of return. The adjusted Sharpe-ratio incorporates skewness and kurtosis, and is based on Pezier and White's (2006) formula:

$$ASR_P = SR_P \times \left[ 1 + \left( \frac{S_P}{6} \right) \times SR_P - \left( \frac{K_P - 3}{24} \right) \times SR_P^2 \right] \quad (60)$$

### 3.5.2 Hypothesis test of difference between portfolio moments

We conduct a hypothesis test of the difference in the out-of-sample mean, variance, skewness and kurtosis, between each portfolio constructed and the two benchmark portfolios, by using heteroscedastic-autocorrelated consistent (HAC) standard errors. The method has been used

by Ledoit and Wolf (2008) and Ledoit and Wolf (2011) to compare the Sharpe-ratio and the variance between portfolios, respectively. By adjusting the method to compare out-of-sample Sharpe-ratios and variance, and based on their derivation of moments we adjust the test by the help of Wolf, in order to apply the test for mean, skewness and kurtosis.

The implementation of the hypothesis test for each moment, in R, is shown in Listings 13. The null- and alternative hypothesis are given by:

$$\begin{aligned} H_0 : \hat{\Delta}_M &= 0 \\ H_1 : \hat{\Delta}_M &\neq 0 \end{aligned}$$

where  $\hat{\Delta}_M$  denotes the difference between the sample estimate of moment M of portfolio x,  $\hat{\theta}_x$ , and portfolio y,  $\hat{\theta}_y$ ;

$$\hat{\Delta}_M = \hat{\theta}_x - \hat{\theta}_y \quad (61)$$

In our case the  $x$  represents any portfolio we construct and  $y$  represents either of the benchmark portfolios. The difference  $\hat{\Delta}_M$  between the portfolio means, - log variances, -skewness' and -kurtosis', is respectively given by:

$$\hat{\Delta}_M = \begin{cases} \hat{\mu}_x - \hat{\mu}_y & \text{if } M = 1 \\ \ln(\hat{\sigma}_x) - \ln(\hat{\sigma}_y) & \text{if } M = 2 \\ \hat{s}_x - \hat{s}_y & \text{if } M = 3 \\ \hat{k}_x - \hat{k}_y & \text{if } M = 4 \end{cases} \quad (62)$$

where  $\hat{\mu}_l$ ,  $\hat{\sigma}_l$ ,  $\hat{s}_l$  and  $\hat{k}_l$ , are the sample estimates of the first, second, third and fourth moment, respectively, for portfolio l. Note that we use the difference of the log variance and not the variance, as this is recommended by Ledoit and Wolf (2008). The standard error of the test,  $s(\hat{\Delta}_M)$ , is given by:

$$s(\hat{\Delta}_M) = \sqrt{\frac{\nabla^T f(\hat{v}) \hat{\Psi}_M \nabla f(\hat{v})}{T}} \quad (63)$$

where  $\nabla^T f(\hat{v})$  is the transposed gradient vector of function  $f$ , and  $\hat{\Psi}_M$  is estimated by the heteroscedastic and autocorrelation robust (HAC) kernel estimate. The closed-form expressions of these components are presented in appendix C. To determine whether there is a significant difference between the out-of-sample moment of portfolio x and the benchmark portfolio y, we

compute the p-value of the test given by Ledoit and Wolf (2011):

$$\hat{p} = 2P \left( Z \leq -\frac{|\hat{\Delta}|}{s(\hat{\Delta})} \right) \quad (64)$$

where  $Z$  is a standard normal variable.

## 4 Data

In this study we use an empirical data set containing the dividend-adjusted daily stock prices, in USD, of 29 stocks obtained from DataStream(c), presented in appendix D. The stocks are randomly selected from the S&P Global 1200-index in 1994, and belong to the investment universe from January 1, 1995 to December 31, 2015. The S&P Global 1200-index covers global, blue-chip stocks over the whole spectre of GICS-industries. The seventeen cyclical stocks are within the following sectors; Consumer Discretionary, Financials, Industrials, Information Technology and Materials. Most of the twelve defensive stocks are registered in Consumer Staples, but we also find stocks belonging to the Energy- and Health Care-sector. The sample period over 20 years includes both the dot-com bubble in 1996-2000, the financial crisis in 2007-2008 and the sovereign debt crisis in 2010-2012. By including such business cycle fluctuations we can observe the impact of including higher-moments during normal times but also during extremely volatile periods.

The mean, variance, skewness, kurtosis, and the Jarque-Bera statistic of each stock in the investment universe, over the period 1996 to 2015, are reported in appendix D. The overall high annual median Jarque-Bera statistics indicate non-normal returns, and the presence of skewness and leptokurtosis further confirms the need to control for higher moments.

In order to compute the Sharpe-ratio and the adjusted Sharper-ratio we use the risk-free rate from Kenneth French's Fama/French Global 3 Factors from the CRSP database (French, 2017).

## 5 Model

We construct and evaluate the performance of the higher-moment portfolios MVS and MVS<sub>K</sub>, and the two benchmark portfolios, GMVP and MV. There are several versions of the mean-

variance-, mean-variance-skewness- and mean-variance-skewness-kurtosis efficient portfolios, as the allocation of the portfolios depend on the specified preference parameters of the investor. In this thesis the MV-, MVS- and MVSK portfolio are constructed with the base case preference 1 for the relevant moments. Following the notation in section 3.4.1, the allocation of the MV portfolio is based on the preference parameter vector  $\lambda = (1, 1, 0, 0)$ , the allocation of the MVS portfolio is based on  $\lambda = (1, 1, 1, 0)$ , and the allocation of the MVSK portfolio is based on  $\lambda = (1, 1, 1, 1)$ .

The portfolios are revised quarterly with quarterly holding periods, and the revisions are based on Bayesian shrinkage estimates of the daily stock moments over a 1-year rolling window. The estimates are scaled to match a quarterly holding period. Each portfolio is revised 80 times over the 20 year period, including the initial allocation. The initial allocations of the portfolios are January 1, 1996, and based on the estimation window from January 1, 1995 to December 31, 1995. The corresponding holding period to the initial allocation is between January 1, 1996 and March 31, 1996. The last revision date for all portfolios is October 1, 2015, and the corresponding holding period ends December 31, 2015.

Portfolios are constructed with a buy-hold strategy and daily rebalancing strategy, and in addition we impose a variety of constraints, described in Table 1.

**Table 1:** Portfolio specifications

Portfolio constraints						
	Basic	Diversification		Turnover		Div.- and turn.
		Mild	Strong	Mild	Strong	Strong
Long-only	Yes	Yes	Yes	Yes	Yes	Yes
$c_1^*$	-	-	-	0.05	0.03	0.03
$c_2^{**}$	-	0.6	0.75	-	-	0.75
Strategy	BH/RB	BH/RB	BH/RB	BH	BH	BH

\* $c_1$  denotes the maximum allowed turnover per revision

\*\* $c_2$  denotes the minimum required diversification per revision.

The basic constraints are imposed on all portfolios and include the full investment- and long-only constraint. We also consider portfolios with additionally the strong or mild turnover constraint of 0.03 and 0.05, respectively. We only impose this constraint on the buy-hold portfolios as we do not want to limit the rebalancing within the holding periods for the daily rebalanced portfolios. In addition, we construct portfolios with the strong and mild diversification constraint of 0.75 and 0.6, respectively. The allocation of the portfolios with both the diversification and the turnover constraints are required to have a minimum diversification of 0.75 and a maximum turnover of 0.03, per revision.

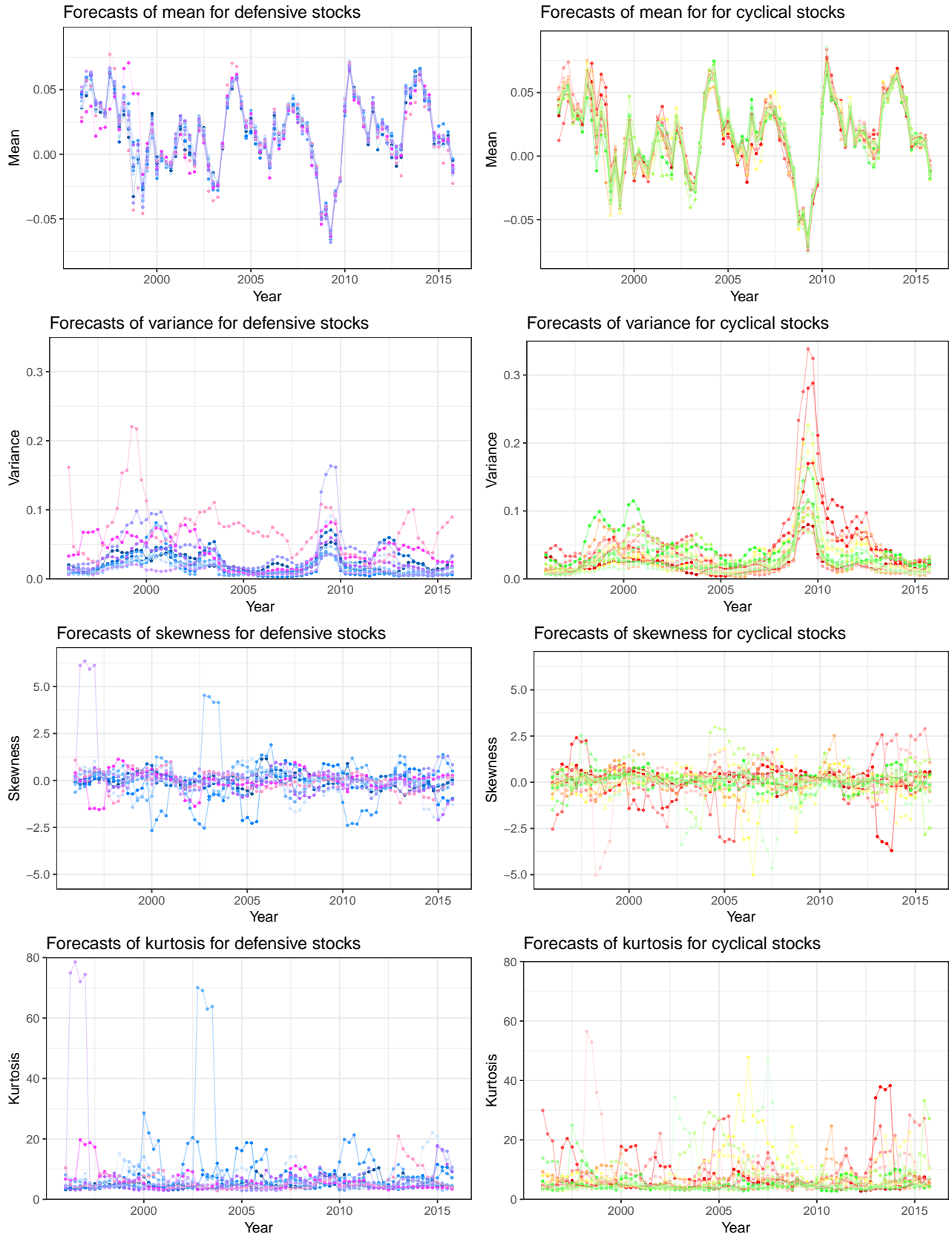
Note that we impose the constraints per revision, and that the portfolios under the buy-hold strategy with a diversification constraint may end up with a portfolio that is less diversified than the target during the holding periods.

### 5.1 Forecasts of mean, variance, skewness and kurtosis

The inputs of the asset allocation framework are the forecasts of each stock's moments. The forecasts are in turn based on Bayesian shrinkage estimates of daily returns over a rolling-window of 1 year, and the time-varying shrinkage intensities of the mean vector, covariance-, coskewness- and cokurtosis matrices, are presented in appendix E.1. For the single-index estimates we used the return series of the S&P Global 1200-index to obtain the beta of each stock and the mean, variance, skewness and kurtosis, of the market.

We only present the forecast of the mean, variance, skewness and kurtosis for each of the 29 stocks, as the forecasts of the co-moment matrices consists of 435 unique forecasted covariance elements, 4495 unique forecasted coskewness elements and 35.960 unique forecasted cokurtosis elements, for each period. Figure 6 shows the quarterly forecasts of the stocks' mean, variance, skewness and kurtosis, for each quarterly holding period. Note that due to graphical purposes the stock forecasts are separated by cyclical- and aggressive stocks, and the color code for each stock is given in appendix D.2.

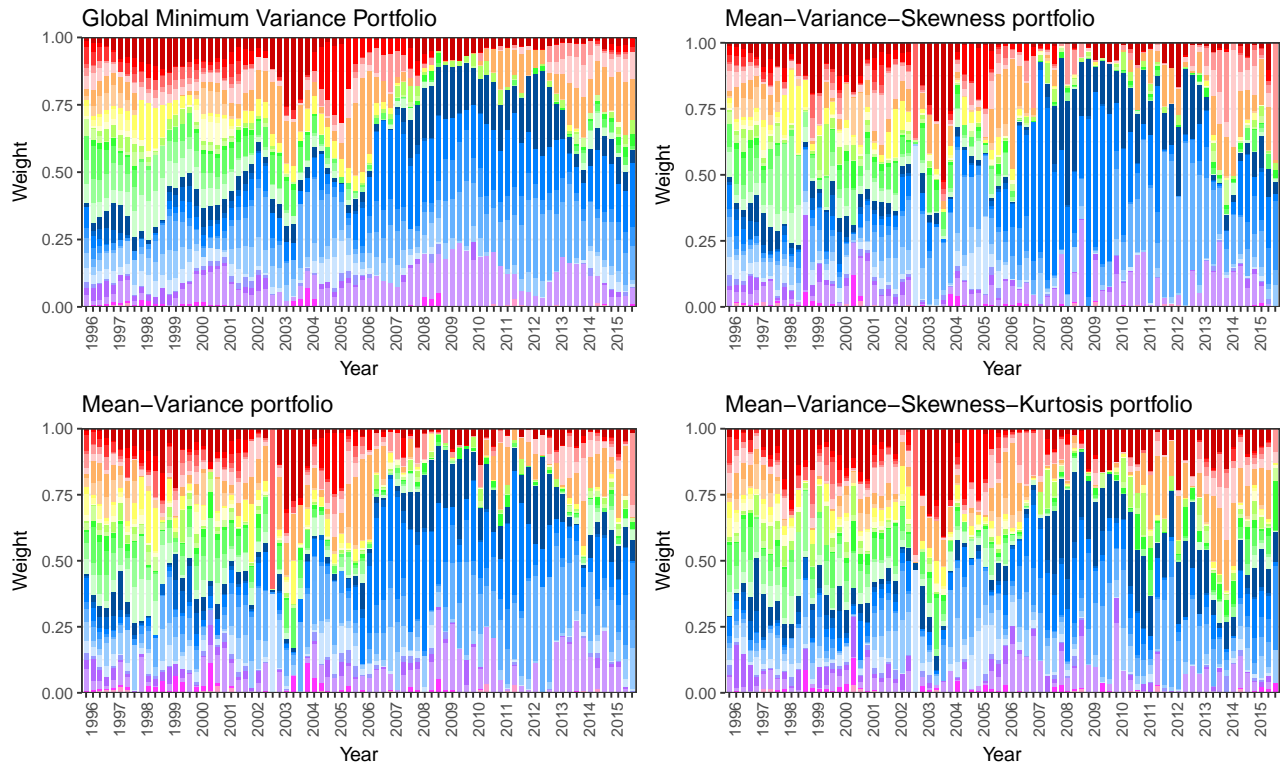
The small difference between the stocks' mean over the whole period is a result of a high shrinkage intensity over time, shown in appendix E.1., towards the common target, i.e. the GMVP. As for the forecasts of variance, we observe high values during and after the dot-com bubble and the financial crisis, and that the cyclical stocks are anticipated to have higher variance than the defensive stocks. The patterns are less clear-cut for the higher-moments; certain defensive stocks exhibit large values of positive skewness, while some defensive and cyclical stocks exhibit negative values of skewness. Especially, the skewness of some cyclical stocks are as low as -5. The forecasts of kurtosis shows that the cyclical stocks tend to have higher kurtosis than the defensive stocks, but we also observe that two of the defensive stocks have occasionally high kurtosis above 60.



**Figure 6:** Quarterly forecasts of mean, variance, skewness and kurtosis by rolling shrinkage

## 5.2 Portfolio allocation, diversification and turnover

The initial allocation and the allocations for each revision, for the GMVP and the MV-, MVS- and MVS-K with basic constraints, are shown in Figure 7. The pattern of all the portfolios' allocations is quite similar over time as they are allocated about evenly between cyclical and defensive stocks prior to the financial crisis, and heavily in defensive stocks during the crisis. The GMVP is the most diversified portfolio overall, but the MVS-K-portfolio is in fact more diversified during and after the financial crisis. One explanation for this are the increasing correlations between the assets, especially during the financial crisis, and the fact that the objective of the MVS-K portfolio also involves diversifying away idiosyncratic skewness and -kurtosis. Thus the MVS-K portfolio is less dependent on the covariances between the assets, yet we notice that this is not the case for the MVS portfolio as it is always less diversified than GMVP and the MV portfolio.



**Figure 7:** Quarterly allocations of portfolios with basic constraints

The weight plots of the portfolios with the additional diversification- and/or turnover constraint are presented in appendix E.2. As the most stable and diversified portfolio overall, the GMVP is not affected by any of the additional constraints. The MVS portfolio is on the other hand

highly affected by the strong turnover- and diversification constraints, and from Table 2 we observe that this is due to the portfolio having the highest average quarterly turnover and lowest average quarterly diversified allocation. The higher turnover of the higher-moment portfolios, compared to the GMVP and the MV portfolio, is also partially explained by the occasionally large positions in few stocks by the MVS- and MVS<sub>K</sub> portfolio. Furthermore we observe that the MVS<sub>K</sub> portfolio on average has more diversified allocations than the MV portfolio, and when the strong diversification constraint is imposed it obtains, on average, a more diversified allocation than GMVP. The MVS portfolio has, on average, the least diversified allocations, even when the strong diversification constraint is imposed.

**Table 2:** Turnover- and diversification rates

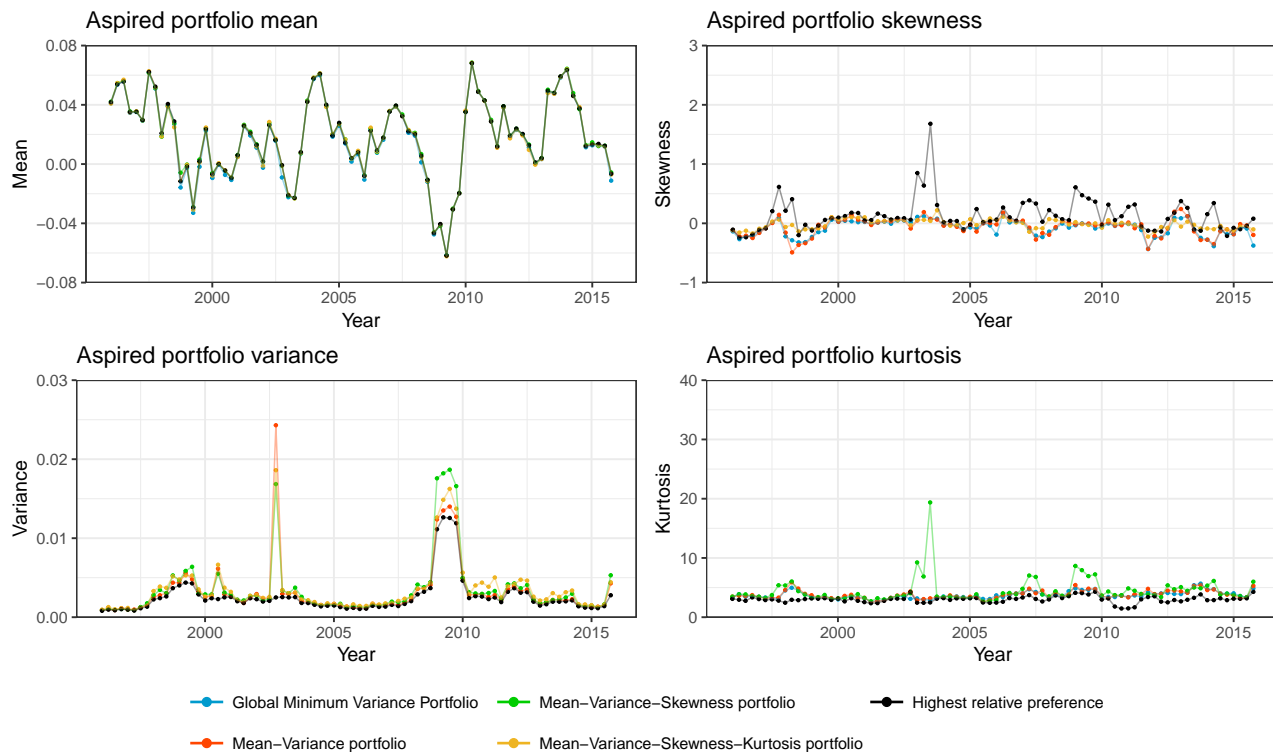
	Diversification			Turnover		
	Min	Mean	Max	Min	Mean	Max
<i>Global Minimum Variance portfolio</i>						
Basic constraints	0.800	0.887	0.949	0.004	0.011	0.034
Mild div. constraint	0.800	0.887	0.949	0.004	0.011	0.034
Strong div. constraint	0.800	0.887	0.949	0.004	0.011	0.034
Mild turn. constraint	0.800	0.887	0.949	0.004	0.011	0.034
Strong turn. constraint	0.800	0.887	0.949	0.004	0.011	0.034
Strong div.- and turn. constraint	0.800	0.887	0.949	0.004	0.011	0.034
<i>Mean-Variance portfolio</i>						
Basic constraints	0.491	0.864	0.949	0.012	0.024	0.067
Mild div. constraint	0.709	0.869	0.946	0.007	0.024	0.052
Strong div. constraint	0.771	0.870	0.944	0.014	0.026	0.055
Mild turn. constraint	0.664	0.872	0.951	0.010	0.023	0.050
Strong turn. constraint	0.611	0.859	0.945	0.010	0.024	0.034
Strong div.- and turn. constraint	0.750	0.867	0.940	0.016	0.026	0.034
<i>Mean-Variance-Skewness portfolio</i>						
Basic constraints	0.493	0.820	0.951	0.006	0.025	0.065
Mild div. constraint	0.600	0.821	0.948	0.012	0.025	0.055
Strong div. constraint	0.750	0.827	0.946	0.014	0.031	0.055
Mild turn. constraint	0.419	0.807	0.945	0.009	0.025	0.050
Strong turn. constraint	0.280	0.768	0.940	0.007	0.025	0.034
Strong div.- and turn. constraint	0.750	0.827	0.944	0.016	0.027	0.034
<i>Mean-Variance-Skewness-Kurtosis portfolio</i>						
Basic constraints	0.554	0.880	0.946	0.008	0.025	0.065
Mild div. constraint	0.716	0.881	0.950	0.014	0.024	0.050
Strong div. constraint	0.752	0.896	0.942	0.014	0.029	0.051
Mild turn. constraint	0.689	0.879	0.945	0.014	0.025	0.047
Strong turn. constraint	0.511	0.850	0.944	0.011	0.025	0.034
Strong div.- and turn. constraint	0.751	0.848	0.937	0.019	0.026	0.034



While the buy-hold strategy and the rebalancing strategy have similar allocations per revision, they differ during the holding periods. The daily rebalancing strategy leads to relatively high turnover during the holding periods, but also a constant diversification ratio. This is due to the fact that the portfolios with the rebalancing strategy only have daily fluctuations from the initial allocation of the holding period. On the other hand, the buy-hold strategy leads to portfolios with zero turnover during the holding period, and is considered to be the cheaper strategy. Yet, we sometimes observe that an initially diversified buy-hold portfolio becomes more concentrated under the holding period if a couple of the assets it holds have relatively large increases in price.

### 5.3 Portfolio aspired moments

The aspired quarterly portfolio moments, known as the in-sample moments, for the GMVP and the MV-, MVS- and MVSK portfolio with basic constraints, are shown in Figure 8. There appears to be no clear dominating portfolio in terms of in-sample quarterly mean. The higher-moment portfolios have a lower relative preference for the mean compared to the MV portfolio, yet the difference between the portfolios' means is negligible. This is a result of the stocks' forecasted means being very similar due to being shrunked towards a common target, with a high shrinkage intensity. By including higher-moments the relative preference for variance also decreases in-sample. From Figure 8 we observe that the GMVP has the lowest variance, in accordance with the definition of the portfolio, and that the MV portfolio has on average lower variance than the higher-moment portfolios. Even though the MVS portfolio has a relatively higher preference for variance than the MVSK portfolio, the trade-off between the moments leads to MVS obtaining a higher skewness but also a higher variance than the MVSK portfolio. The MVSK portfolio is the only portfolio with a preference for kurtosis, and we observe that while it obtains the lowest in-sample kurtosis, the MVS portfolio obtains the highest one.



**Figure 8:** Quarterly in-sample moments for portfolios with basic constraints

From appendix D we also show the aspired moments of the portfolios with the diversification and/or turnover constraints. Naturally, the variance of the portfolios are more similar to the variance of the GMVP when the diversification constraint is imposed, and the stronger it is. The relationships between the portfolios' in-sample moments are, in general, the same for all the different sets of constraints except when both the strict turnover- and diversification are imposed. In this case, the MVS<sub>K</sub> portfolio obtains a higher in-sample variance compared to the MVS portfolio.

## 6 Results

We evaluate the out-of-sample performance of the portfolios from the initial allocation at January 1, 1996 till the end of the last holding period at December 31, 2015. First, we evaluate the wealth of the GMVP and the MV-, MVS- and MVS<sub>K</sub> portfolio. Then we compare other criteria such as the risk-adjusted return, risk profile and overall turnover, and examine the out-of-sample performance of the higher-moment portfolios versus the two benchmark portfolios, and versus the different constraints imposed. Finally, we conduct a statistical test where we

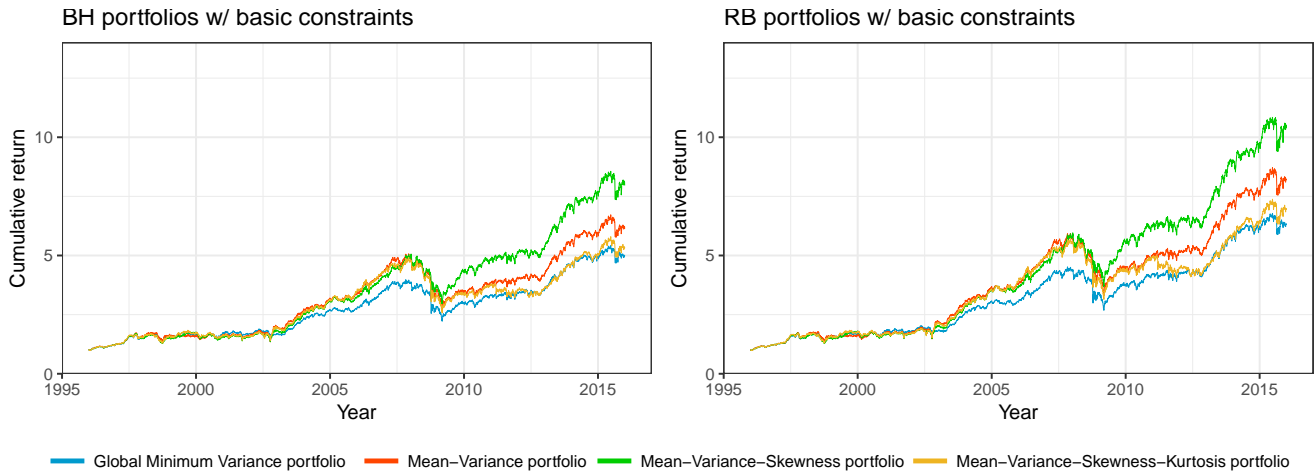
compare the out-of-sample mean, variance, skewness and kurtosis, for each portfolio against each of the benchmark portfolios.

## 6.1 Out-of-sample performance

### 6.1.1 Portfolio wealth

Figure 9 to 12 illustrate the daily cumulative return of a \$1 investment in each of the GMVP, and the MV-, MVS- and MVSK portfolio, given different sets of constraints and strategies. The portfolios are ranked based on terminal wealth for a given set of constraints, as if the investor is subject to a set of constraints and chooses the portfolio that gives him the highest wealth at the end of the period.

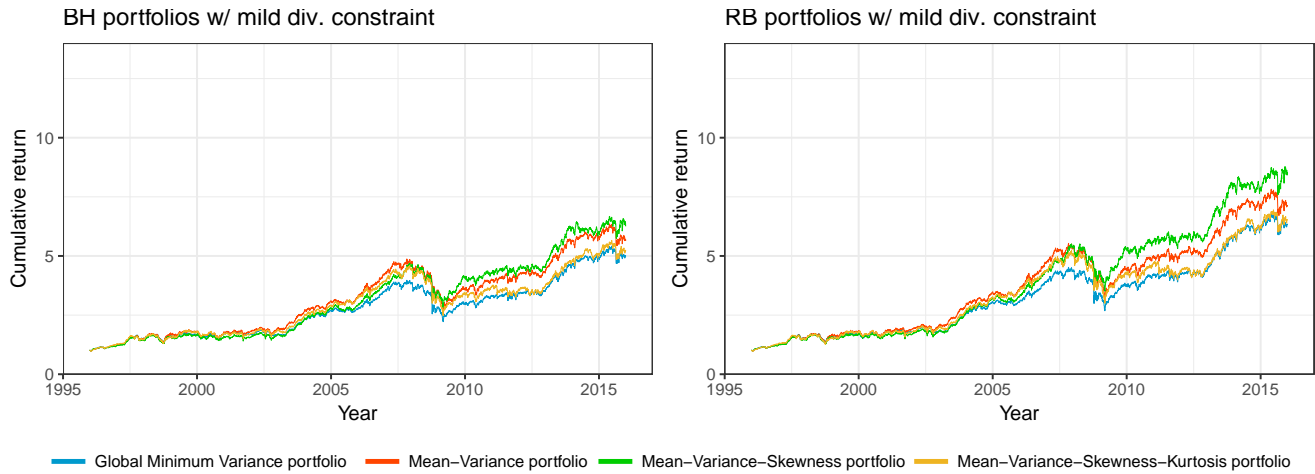
Given basic constraints the MVS portfolio obtains the highest terminal wealth, under both the buy-hold and rebalancing strategy, as shown in Figure 9. The MVSK portfolio obtains higher terminal wealth than GMVP, under both strategies, yet it loses to the MV portfolio. All of the portfolios obtain higher returns when rebalanced compared to buy-hold, and the MV- and MVS portfolio benefits the most being rebalanced.



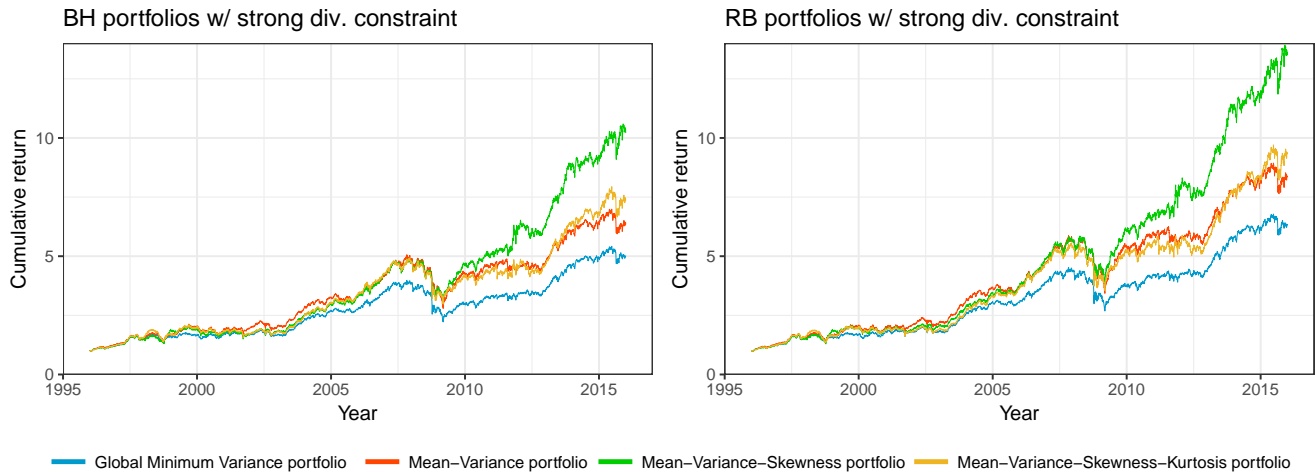
**Figure 9:** Cumulative returns for portfolios with basic constraints

The mild and strong diversification constraint have different implications for the higher-moment portfolios, as shown in Figure 10 and Figure 11. While the mild diversification constraint reduces the dominance of the MVS portfolio, and reduces the relative performance of the MVSK portfolio, the strong diversification constraint improves the relative performance of the

higher-moments portfolios. In fact, the MVS<sub>K</sub> portfolio obtains a higher terminal wealth than the MV portfolio under the strong diversification constraint, but vice versa under the mild diversification constraint and the basic constraints, regardless of strategy.

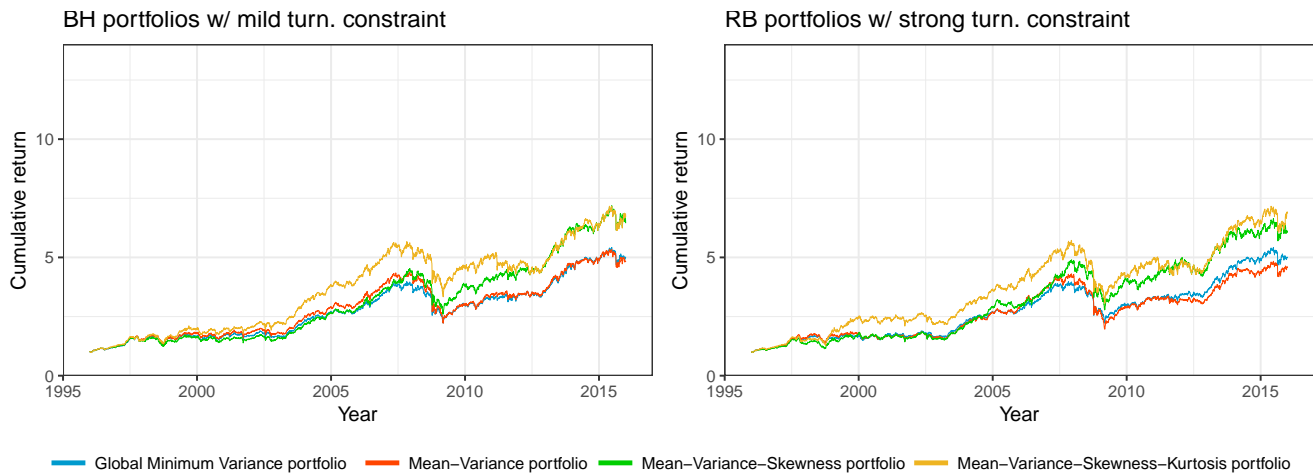


**Figure 10:** Cumulative returns for portfolios with mild div. constraint



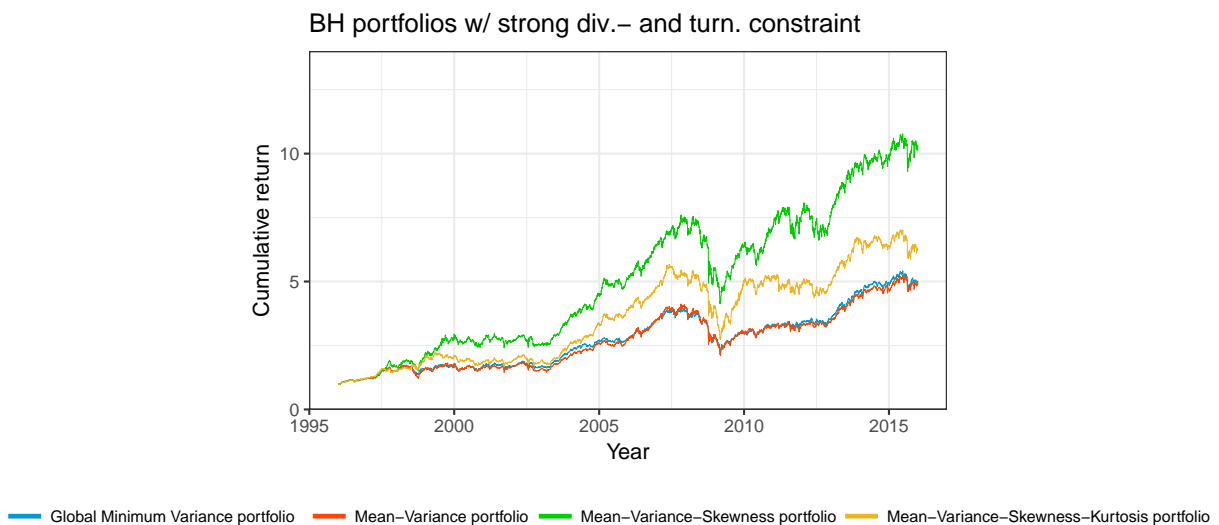
**Figure 11:** Cumulative returns for portfolios with strong div. constraint

From Figure 12 we observe that the higher-moment portfolios obtain the highest terminal wealth with the turnover constraint imposed. The MVS<sub>K</sub> portfolio obtains the highest terminal wealth, for both the mild and the strong turnover constraint, and is positively affected compared to its counterpart with basic constraints. The MVS portfolio is negatively affected by the turnover constraints, and we also observe that the MV portfolio obtains lower terminal wealth than the GMVP.



**Figure 12:** Cumulative returns for portfolios with turn. constraint

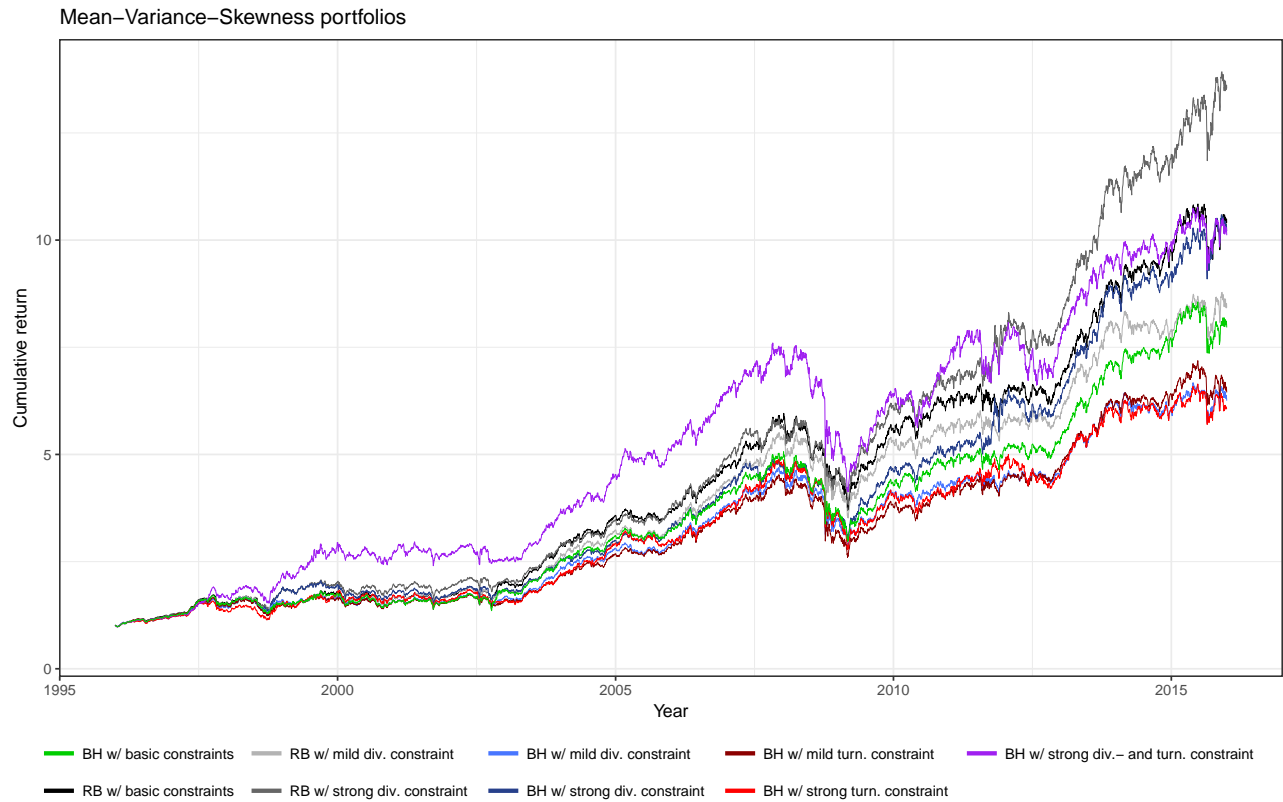
The higher-moment portfolios obtain higher terminal wealth given the strong diversification- and strong turnover constraint, as shown in Figure 13. While GMVP is unaffected by this constraint set, the MV portfolio performs relatively worse and even obtains the lowest terminal wealth. The MVS portfolio ends up with the highest wealth, followed by the MVSK portfolio, and this indicates that the strong diversification constraint is the more important driver as the MVSK portfolio dominates MVS when only the additional turnover constraint is imposed.



**Figure 13:** Cumulative returns for portfolios with strong div. and turn. constraint

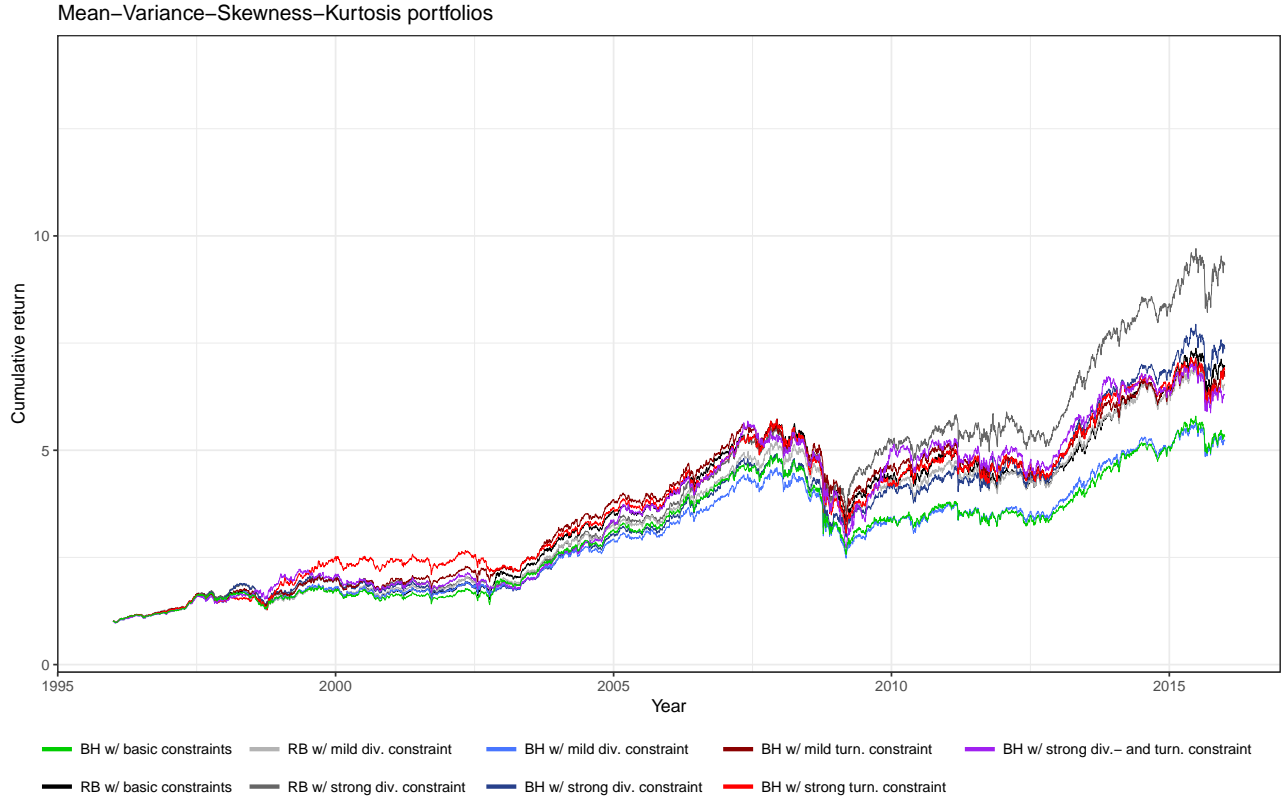
Figure 14 shows the cumulative returns of the MVS portfolio given different sets of constraints, under the buy-hold- and the rebalancing strategy. The cumulative returns of the MVS portfolio are particularly affected by the diversification constraint due to the occasional concentrated

allocations. The scenario where the MVS portfolio obtains the highest terminal wealth is with the strong diversification constraint and the rebalancing strategy. Imposing only the turnover constraint, reduces the terminal wealth of the MVS portfolio, compared to the buy-hold version with basic constraints. The MVS portfolio also performs well when both the strong diversification- and the strong turnover constraints are imposed.



**Figure 14:** Wealth of mean-variance-skewness portfolios

Figure 15 shows the cumulative returns of the MVSK portfolio given different sets of constraints under the buy-hold- and the rebalancing strategy. The strong diversified MVSK portfolios end up with the highest terminal wealth regardless of strategy, and we observe that the rebalanced MVSK portfolio with the strong diversification constraint performs the best out of all of the portfolios in Figure 15. Both the turnover and the strong diversification constraint have a positive impact on the MVSK portfolio, in fact the wealth of the MVSK portfolio with any of these constraints have been the best versions of this portfolio since 1997.



**Figure 15:** Wealth of mean-variance-skewness-kurtosis portfolios

### 6.1.2 Performance measures

Table 3 reports the annual returns, Sharpe-ratio (SR) and Adjusted Sharpe-ratio (ASR), standard deviation (SD), expected shortfall (ES) and quarterly turnover (TR), for all portfolios constructed in this thesis. The quarterly turnover is included in the table to give an indication of the relative cost of the portfolio, because higher turnover implies higher trading costs. Hence, the rebalanced portfolios might not obtain higher profit than the buy-hold counterparts, even though they obtain higher returns compared to the respective buy-hold portfolios. We compare the higher-moment portfolios to the benchmark portfolios given the basic-, diversification-, and turnover constraint, under the buy-hold and rebalancing strategy.

Given buy-hold portfolios with basic constraints, we observe that the MVS portfolio outperforms the other portfolios in terms of annualized returns, Sharpe-ratio, adjusted Sharpe-ratio and expected shortfall, while GMVP has the lowest standard deviation. The MVSK portfolio generally performs worse than the two benchmarks. The MV portfolio outperforms MVSK over any metric considered, and GMVP has a lower standard deviation and expected shortfall,

than MVSK as well. MVSK obtains higher returns than GMVP and the difference between the portfolios Sharpe-ratio when adjusted for skewness and kurtosis, is negligible. With basic constraints and rebalancing, the MVS portfolio exhibits better annualized return, standard deviation, Sharpe-ratio and adjusted Sharpe-ratio, compared to the other portfolios. The MVSK portfolio is still outperformed by the MV portfolio and only outperforms GMVP in terms of annual return and adjusted Sharpe-ratio. In general, we observe that with only basic constraints imposed the higher-moment portfolios have consistently higher turnover compared to the benchmark portfolios.

The diversification constraint has mixed implications for the higher moment portfolios compared to the basic constraints, under both the rebalanced and buy-hold strategy. The buy-hold and rebalanced higher-moment portfolios, with the additional mild diversification constraint, exhibit a better standard deviation and expected shortfall, but at the same time lower returns. The buy-hold and rebalanced higher-moment portfolios, with the additional strong diversification constraint, exhibit higher returns, higher expected shortfall, but also higher standard deviation.

Imposing the mild diversification constraint on the MVS portfolio leads to lower annualized returns, lower Sharpe-ratio, adjusted Sharpe-ratio, and higher expected shortfall. The MVSK portfolio obtains a higher adjusted Sharpe-ratio under the buy-hold strategy, compared to the counterpart with basic constraints, but a lower one given the rebalancing strategy. The strong diversification constraint has a positive impact on the portfolios affected by it, i.e. the MV-, MVS- and MVSK portfolio, both in terms of annualized returns, Sharpe-ratio, adjusted Sharpe-ratio and expected shortfall. The best version of the MS portfolio when it comes to returns and risk-adjusted return is the strong diversification constraint.

Given the mild- or strong diversification constraint under the buy-hold and the rebalancing strategy, the MVS portfolios perform better in terms of Sharpe-ratio, adjusted Sharpe-ratio and expected shortfall than the benchmark portfolios. The performance of the MVSK portfolios compared to the benchmarks are mixed. The MVSK portfolios have a higher annualized return and adjusted Sharpe-ratio than the GMVP, but lower Sharpe-ratio and higher expected shortfall. The MVSK portfolios are worse than the MV portfolios on all metrics, given the mild diversification constraint. However, given the strong diversification constraint, the MVSK portfolio obtains better annualized returns, Sharpe-ratio and adjusted Sharpe-ratio than the two benchmarks.

From a risk perspective the turnover has mixed implications on the higher moment portfolios compared to their counterparts with basic constraints. The mild turnover constraint leads



to lower standard deviations for the higher-moment portfolios, yet the expected shortfall is reduced for the MVS portfolio while it is increased for the MVSK portfolio. Imposing the turnover constraint on the MVS portfolios leads to lower annualized returns, Sharpe-ratio and adjusted Sharpe-ratio, for both the mild and strong turnover constraint. It has a positive effect on the MVSK portfolio from a return perspective as the annualized returns, Sharpe-ratio and adjusted Sharpe-ratio improves.

Given the turnover constraint the higher-moment portfolios perform better than the two benchmark portfolios, in terms of returns and risk-adjusted returns. The MVSK portfolio has the highest Sharpe- and adjusted Sharpe-ratio, followed by the MVS portfolio. The higher-moment portfolios also have the highest fluctuations overall as their standard deviations are consistently higher than the standard deviations of the benchmark portfolios, in addition to having the highest turnover.

By imposing both the diversification and turnover constraint, the higher-moment portfolios obtain relatively high standard deviations compared to the portfolio with basic constraints. Yet, additional constraint leads the higher- moment portfolio to outperform their counterpart with basic constraints, in terms of the annualized returns, Sharpe-ratio and adjusted Sharpe-ratio. As opposed to the situation with basic constraints, the MVSK portfolio with both the diversification constraint and the turnover constraint imposed, obtains a lower expected shortfall than GMVP. Yet the standard deviation of the higher-moment portfolio are still relatively high, and the turnover of the MVS portfolio is the highest among the portfolios, given the additional turnover and diversification constraint.

**Table 3:** Annualized performance measures

	Return	SD	SR*	ASR*	ES 5%	Turnover**
<i>Buy-Hold portfolios with basic constraints</i>						
Global Minimum Variance portfolio	0.081	0.116	0.483	0.462	0.216	0.011
Mean-Variance portfolio	0.092	0.128	0.523	0.485	0.231	0.024
Mean-Variance-Skewness portfolio	0.105	0.129	0.620	0.572	0.210	0.025
Mean-Variance-Skewness-Kurtosis portfolio	0.084	0.132	0.448	0.462	0.258	0.025
<i>Rebalanced portfolios with basic constraints</i>						
Global Minimum Variance portfolio	0.093	0.118	0.580	0.517	0.172	0.132
Mean-Variance portfolio	0.106	0.129	0.632	0.538	0.194	0.144
Mean-Variance-Skewness portfolio	0.119	0.130	0.720	0.628	0.176	0.139
Mean-Variance-Skewness-Kurtosis portfolio	0.098	0.134	0.545	0.533	0.231	0.152
<i>Buy-hold portfolios with mild div. constraint</i>						
Global Minimum Variance portfolio	0.081	0.116	0.483	0.462	0.216	0.011
Mean-Variance portfolio	0.087	0.122	0.510	0.499	0.204	0.024
Mean-Variance-Skewness portfolio	0.093	0.128	0.529	0.514	0.193	0.025
Mean-Variance-Skewness-Kurtosis portfolio	0.083	0.128	0.455	0.464	0.230	0.024
<i>Rebalanced portfolios with mild div. constraint</i>						
Global Minimum Variance portfolio	0.093	0.118	0.580	0.517	0.172	0.132
Mean-Variance portfolio	0.099	0.124	0.598	0.554	0.168	0.146
Mean-Variance-Skewness portfolio	0.108	0.129	0.640	0.587	0.158	0.139
Mean-Variance-Skewness-Kurtosis portfolio	0.095	0.130	0.538	0.525	0.188	0.150
<i>Buy-hold portfolios with strong div. constraint</i>						
Global Minimum Variance portfolio	0.081	0.116	0.483	0.462	0.216	0.011
Mean-Variance portfolio	0.093	0.125	0.547	0.506	0.205	0.026
Mean-Variance-Skewness portfolio	0.119	0.136	0.686	0.612	0.201	0.031
Mean-Variance-Skewness-Kurtosis portfolio	0.101	0.134	0.569	0.545	0.232	0.029
<i>Rebalanced portfolios with strong div. constraint</i>						
Global Minimum Variance portfolio	0.093	0.118	0.580	0.517	0.172	0.132
Mean-Variance portfolio	0.108	0.126	0.653	0.563	0.171	0.149
Mean-Variance-Skewness portfolio	0.134	0.137	0.786	0.665	0.157	0.149
Mean-Variance-Skewness-Kurtosis portfolio	0.113	0.136	0.650	0.596	0.195	0.153
<i>Buy-hold portfolios with mild turn. constraint</i>						
Global Minimum Variance portfolio	0.081	0.116	0.483	0.462	0.216	0.011
Mean-Variance portfolio	0.079	0.122	0.445	0.445	0.261	0.023
Mean-Variance-Skewness portfolio	0.094	0.128	0.541	0.527	0.224	0.025
Mean-Variance-Skewness-Kurtosis portfolio	0.096	0.131	0.546	0.531	0.188	0.025
<i>Buy-hold portfolios with strong turn. constraint</i>						
Global Minimum Variance portfolio	0.081	0.116	0.483	0.462	0.216	0.011
Mean-Variance portfolio	0.076	0.124	0.415	0.426	0.286	0.024
Mean-Variance-Skewness portfolio	0.091	0.136	0.487	0.499	0.207	0.025
Mean-Variance-Skewness-Kurtosis portfolio	0.097	0.137	0.531	0.525	0.253	0.025
<i>Buy-hold portfolios with strong div. and turn. constraint</i>						
Global Minimum Variance portfolio	0.081	0.116	0.483	0.462	0.216	0.011
Mean-Variance portfolio	0.080	0.128	0.431	0.440	0.224	0.026
Mean-Variance-Skewness portfolio	0.118	0.144	0.644	0.574	0.218	0.027
Mean-Variance-Skewness-Kurtosis portfolio	0.093	0.143	0.477	0.488	0.213	0.026

\* Using an annual risk-free rate of return at 2.3% \*\* Average quarterly turnover

## 6.2 Out-of-sample portfolio moments

In this section we evaluate the differences between the portfolios' monthly out-of-sample mean, variance, skewness and kurtosis, using 240 monthly observations. In addition, we make inferences based on the statistical difference test, presented in section 3.5.2, between the respective moments of each portfolio constructed and each of the benchmark portfolios.

### 6.2.1 Test of out-of-sample portfolio mean

Table 4 reports the differences, and the corresponding p-values, between the monthly out-of-sample mean of GMVP and the MV-, MVS- and MVSK portfolio, given the constraints and strategies considered in this thesis. We observe that all of the portfolios obtain a higher out-of-sample mean than GMVP, given any constraints with the buy-hold or rebalancing strategy, and that the portfolios given certain constraints also obtain significantly higher mean. The relative higher means of the MV-, MVS- and MVSK portfolios are consistent with the in-sample means compared to GMVP, as the latter does not include mean in the optimization problem and the other portfolios do. According to the relative preferences, the MV portfolio has the highest preference for mean, yet we observe that the MVS portfolio has the highest mean given any set of constraints and under both of the strategies.

**Table 4:** Difference test of monthly out-of-sample mean with GMVP as benchmark

Strategy	Constraint	MV		MVS		MVSK	
		$\Delta_1$	p-value	$\Delta_1$	p-value	$\Delta_1$	p-value
BH	Basic	0.001	0.432	0.002	0.0591*	0.000	0.790
BH	Mild div.	0.001	0.475	0.001	0.253	0.000	0.807
BH	Strong div.	0.001	0.159	0.003	0.015**	0.002	0.098*
BH	Mild turn.	0.000	0.699	0.001	0.435	0.001	0.269
BH	Strong turn.	0.000	0.841	0.001	0.177	0.001	0.176
BH	Strong div.- and turn.	0.000	0.934	0.003	0.038**	0.001	0.484
RB	Basic	0.001	0.343	0.002	0.041**	0.001	0.709
RB	Mild div.	0.001	0.485	0.001	0.172	0.000	0.844
RB	Strong div.	0.001	0.085*	0.003	0.010**	0.002	0.098*

\*  $p < 0.1$  \*\*  $p < 0.05$  \*\*\*  $p < 0.01$

Table 5 reports the differences, and the corresponding p-values, between the monthly out-of-sample mean of MV and the global minimum variance-, MVS- and MVSK portfolio. The MVS portfolio has a significantly higher out-of-sample mean than the MV portfolio, given the strong diversification constraint or the strong diversification- and strong turnover constraint. The out-

of-sample mean between the MV- and MVSK portfolio is different over the constrains; under basic constraints the MVSK portfolio obtains a lower mean than the MV portfolio and also when the mild diversification constraint is imposed (negative on the fourth decimal). Given the strong diversification constraint, both with buy-hold or rebalancing strategy, the MVSK portfolio obtains a higher out-of-sample mean than the MV, although neither of their differences are significant.

**Table 5:** Difference test of monthly out-of-sample mean with MV portfolio as benchmark

Strategy	Constraint	GMVP		MVS		MVSK	
		$\Delta_1$	p-value	$\Delta_1$	p-value	$\Delta_1$	p-value
BH	Basic	-0.001	0.432	0.001	0.197	-0.001	0.522
BH	Mild div.	-0.001	0.475	0.001	0.559	0.000	0.716
BH	Strong div.	-0.001	0.159	0.002	0.097*	0.001	0.523
BH	Mild turn.	0.000	0.699	0.001	0.328	0.002	0.188
BH	Strong turn.	0.000	0.841	0.001	0.138	0.001	0.162
BH	Strong div.- and turn.	0.000	0.934	0.003	0.037**	0.001	0.515
RB	Basic	-0.001	0.343	0.001	0.248	-0.001	0.493
RB	Mild div.	-0.001	0.485	0.001	0.380	0.000	0.705
RB	Strong div.	-0.001	0.085*	0.002	0.092*	0.001	0.628

\*  $p < 0.1$  \*\*  $p < 0.05$  \*\*\*  $p < 0.01$

### 6.2.2 Test of out-of-sample portfolio variance

Table 6 reports the differences between the monthly out-of-sample log variance of GMVP and the MV-, MVS- and MVSK portfolio, and the corresponding p-values. Given any constraint and strategy, GMVP obtains a lower out-of-sample variance than the other portfolios. This is consistent with the in-sample prediction as GMVP has the highest relative preference of this moment and also has the lowest in-sample variance. Most of the differences are significant, except the difference for the rebalanced MV portfolio with basic constraints.

Table 7 reports the differences between the monthly out-of-sample log variance of MV and the global minimum variance-, MVS- and MVSK portfolio, and the corresponding p-values. The higher-moment portfolios obtain generally higher out-of-sample variance than the MV portfolio, but given the mild diversification constraint the MVS portfolio obtains a lower variance, and given the strong turnover constraint the MVSK obtains a lower variance. According to the relative preferences, the MV portfolio should have the lowest in-sample variance compared to the MVS- and MVSK portfolio, yet there might be an allocation where the MV portfolio obtains

**Table 6:** Difference test of monthly out-of-sample log variance with GMVP as benchmark

Strategy	Constraint	MV		MVS		MVSK	
		$\Delta_2$	p-value	$\Delta_2$	p-value	$\Delta_2$	p-value
BH	Basic	0.165	0.081*	0.187	0.008***	0.191	0.014**
BH	Mild div.	0.104	0.016**	0.100	0.091*	0.153	0.000***
BH	Strong div.	0.082	0.048**	0.245	0.006***	0.269	0.000***
BH	Mild turn.	0.178	0.006***	0.312	0.000***	0.288	0.000***
BH	Strong turn.	0.137	0.000***	0.155	0.001***	0.123	0.030**
BH	Strong div.- and turn.	0.308	0.000***	0.466	0.000***	0.438	0.000***
RB	Basic	0.179	0.170	0.203	0.011**	0.236	0.012**
RB	Mild div.	0.111	0.017**	0.093	0.094*	0.173	0.001***
RB	Strong div.	0.137	0.002**	0.207	0.017**	0.301	0.000***

\*  $p < 0.1$  \*\*  $p < 0.05$  \*\*\*  $p < 0.01$

a higher mean, but also has to accept a higher variance than the higher-moment portfolios. We observe that given the mild diversification constraint or the strong turnover constraint, the MVS- or the MVSK portfolio obtain higher out-of-sample mean and lower variance, thus each of the higher-moment portfolios are better than the MV portfolio, given these constraints, in the out-of-sample mean-variance space.

**Table 7:** Difference test of monthly out-of-sample log variance with MV portfolio as benchmark

Strategy	Constraint	GMVP		MVS		MVSK	
		$\Delta_2$	p-value	$\Delta_2$	p-value	$\Delta_2$	p-value
BH	Basic	-0.165	0.081*	0.021	0.760	0.025	0.687
BH	Mild div.	-0.104	0.016**	-0.004	0.942	0.049	0.311
BH	Strong div.	-0.082	0.048**	0.163	0.112	0.187	0.065*
BH	Mild turn.	-0.178	0.006***	0.134	0.079*	0.110	0.242
BH	Strong turn.	-0.137	0.000***	0.018	0.722	-0.014	0.839
BH	Strong div.- and turn.	-0.308	0.000***	0.158	0.064*	0.130	0.162
RB	Basic	-0.179	0.170	0.025	0.754	0.058	0.370
RB	Mild div.	-0.111	0.017**	-0.018	0.793	0.062	0.167
RB	Strong div.	-0.137	0.002***	0.070	0.502	0.165	0.043**

\*  $p < 0.1$  \*\*  $p < 0.05$  \*\*\*  $p < 0.01$

### 6.2.3 Test of out-of-sample portfolio skewness

Table 8 reports the differences, and the corresponding p-values, between the monthly out-of-sample skewness of GMVP and the MV-, MVS-, and MVSK portfolio, given the constraints and strategies considered in this thesis. The higher-moment portfolios, who also incorporate

skewness into the portfolio selection process, obtain higher skewness compared to GMVP, except the rebalanced MVSK portfolio given the mild diversification constraint. The MVS portfolio has the highest skewness differences versus GMVP, and they are significantly higher, except given the basic constraint or the strong turnover constraint. According to the in-sample moments we expect MVS, and then MVSK, to have the highest out-of-sample skewness, and based on the out-of-sample differences this relationship seems to persist to some extent.

**Table 8:** Difference test of monthly out-of-sample skewness with GMVP as benchmark

Strategy	Constraint	MV		MVS		MVSK	
		$\Delta_3$	p-value	$\Delta_3$	p-value	$\Delta_3$	p-value
BH	Basic	0.267	0.485	0.245	0.227	0.345	0.177
BH	Mild div.	-0.004	0.976	0.321	0.033**	0.002	0.986
BH	Strong div.	-0.030	0.773	0.602	0.020**	0.301	0.147
BH	Mild turn.	-0.107	0.479	0.375	0.018**	0.247	0.081*
BH	Strong turn.	-0.172	0.024**	0.171	0.180	0.088	0.556
BH	Strong div.- and turn.	0.045	0.759	0.422	0.024**	0.044	0.873
RB	Basic	0.599	0.351	0.401	0.124	0.487	0.167
RB	Mild div.	-0.146	0.318	0.309	0.018**	-0.063	0.656
RB	Strong div.	-0.071	0.536	0.478	0.013**	0.249	0.092*

\*  $p < 0.1$  \*\*  $p < 0.05$  \*\*\*  $p < 0.01$

Table 9 reports the differences, and the corresponding p-values, between the monthly out-of-sample skewness of MV and the global minimum variance-,MVS- and MVSK portfolio, given the constraints and strategies considered in this thesis. The MVS portfolio obtains significantly higher out-of-sample skewness than the MV portfolio, except when the basic constraint is imposed. The MVSK portfolio obtains higher out-of-sample skewness than the MV portfolio, except given strong diversification- and strong turnover constraint, or when the portfolios are rebalanced given mild diversification constraint.

#### 6.2.4 Test of out-of-sample portfolio kurtosis

Table 10 reports the differences, and the corresponding p-values, between the monthly out-of-sample kurtosis of GMVP and the MV-,MVS- and MVSK portfolio, given the constraints and strategies considered in this thesis. The only portfolio to incorporate kurtosis in the allocation framework is the MVSK portfolio, yet it only obtains a significantly lower out-of-sample kurtosis under buy-hold with the mild diversification constraint. We also observe that the MVS portfolio obtains significantly lower out-of-sample kurtosis compared to GMVP, given certain constraints. The MVS portfolio also has a lower kurtosis compared to the MVSK portfolio given most of the

**Table 9:** Difference test of monthly out-of-sample skewness with MV portfolio as benchmark

Strategy	Constraint	GVMP		MVS		MVSK	
		$\Delta_3$	p-value	$\Delta_3$	p-value	$\Delta_3$	p-value
BH	Basic	-0.267	0.485	-0.023	0.936	0.078	0.695
BH	Mild div.	0.004	0.976	0.324	0.007***	0.006	0.964
BH	Strong div.	0.030	0.773	0.632	0.022**	0.331	0.078*
BH	Mild turn.	0.107	0.479	0.482	0.006***	0.354	0.077*
BH	Strong turn.	0.172	0.024**	0.343	0.005***	0.259	0.083*
BH	Strong div.- and turn.	-0.045	0.759	0.376	0.044**	-0.001	0.995
RB	Basic	-0.599	0.351	-0.199	0.633	-0.112	0.716
RB	Mild div.	0.146	0.318	0.456	0.034**	0.083	0.448
RB	Strong div.	0.071	0.536	0.549	0.017**	0.321	0.097*

\*  $p < 0.1$  \*\*  $p < 0.05$  \*\*\*  $p < 0.01$

constraints, as shown by the negative coefficients and the magnitudes of the differences versus GMVP.

**Table 10:** Difference test of monthly out-of-sample kurtosis with GMVP as benchmark

Strategy	Constraint	MV		MVS		MVSK	
		$\Delta_4$	p-value	$\Delta_4$	p-value	$\Delta_4$	p-value
BH	Basic	1.194	0.212	-0.605	0.256	-0.276	0.660
BH	Mild div.	-0.175	0.681	-0.939	0.061**	-0.504	0.094*
BH	Strong div.	-0.331	0.259	0.582	0.424	-0.646	0.278
BH	Mild turn.	0.182	0.707	-1.210	0.023**	-0.630	0.171
BH	Strong turn.	0.355	0.098	-0.741	0.054*	-0.747	0.115
BH	Strong div.- and turn.	-0.092	0.855	-0.378	0.491	0.101	0.905
RB	Basic	3.232	0.134	0.160	0.792	0.935	0.330
RB	Mild div.	0.616	0.336	-0.767	0.031**	0.229	0.697
RB	Strong div.	0.116	0.797	-0.028	0.950	-0.484	0.172

\*  $p < 0.1$  \*\*  $p < 0.05$  \*\*\*  $p < 0.01$

Table 11 reports the difference between the monthly out-of-sample kurtosis of the MV portfolio and the global minimum variance-, MVS- and MVSK portfolio, given the constraints and strategies considered in this thesis. We observe that the higher-moment portfolios obtain significantly lower out-of-sample kurtosis than the MV portfolio, given most constraints. The magnitude of the out-of-sample kurtosis difference between the MVS and MVSK portfolio reveals that the value of kurtosis is highly dependent on constraint and strategy.

While kurtosis is not considered in the asset allocation framework of the MVS portfolio, the portfolio obtains a more optimal value for this moment out-of-sample than the MVSK portfolio given the basic constraints, under both the buy-hold and rebalancing strategy. This

**Table 11:** Difference test of monthly out-of-sample kurtosis with MV portfolio as benchmark

Strategy	Constraint	GMVP		MVS		MVSK	
		$\Delta_4$	p-value	$\Delta_4$	p-value	$\Delta_4$	p-value
BH	Basic	-1.194	0.212	-1.799	0.020**	-1.469	0.007***
BH	Mild div.	0.175	0.681	-0.764	0.004***	-0.329	0.263
BH	Strong div.	0.331	0.241	0.913	0.138	-0.315	0.420
BH	Mild turn.	-0.182	0.707	-1.391	0.003***	-0.811	0.101
BH	Strong turn.	-0.355	0.098*	-1.095	0.003***	-1.101	0.018**
BH	Strong div.- and turn.	0.092	0.855	-0.286	0.475	0.193	0.729
RB	Basic	-3.232	0.134	-3.072	0.059**	-2.298	0.071*
RB	Mild div.	-0.616	0.336	-1.383	0.124	-0.387	0.095*
RB	Strong div.	-0.116	0.797	-0.144	0.856	-0.600	0.392

\*  $p < 0.1$  \*\*  $p < 0.05$  \*\*\*  $p < 0.01$

might be an indication of the kurtosis estimate containing too much error; there are large discrepancies between the relative in-sample and out-of-sample kurtosis between the portfolios and constraints, complicating the inference.



## 7 Conclusion

In this thesis we have provided dynamic moment-based portfolios, with and without higher-moments, over a 20-year horizon with quarterly revisions, given different sets of constraints imposed. We evaluate the performance of the higher-moment portfolios, the MVS- and MVSK portfolio, against two benchmark portfolios in terms of annualized returns, risk-adjusted returns, risk and allocation. The benchmark portfolios are two popular portfolios based on the traditional Markowitz framework; the mean-variance portfolio and the global minimum variance portfolio. In addition, we test if the relationship between the in-sample moments of the higher-moment portfolios persist out-of-sample; i.e. whether the higher-moment portfolios also obtain better out-of-sample higher-moments, and worse out-of-sample mean and/or variance, compared to the benchmark portfolios, as anticipated by the in-sample moments.

The MVS- and the MVSK portfolio require relatively high turnover for the quarterly revisions, and are in general not as diversified as GMVP and the MV portfolio. All of the moment-based portfolios follow a similar trend of allocating about evenly between cyclical and defensive stocks before the financial crisis occurs mid 2007, while during the crisis either of them allocate 70%, or more, of the portfolio wealth in defensive stocks. Since variance is the only moment connecting all of the portfolios, this indicates that this is the most important determinant of the allocation for all of the portfolios.

Our findings suggest that by adding skewness to the traditional Markowitz portfolio one can improve performance in terms of return and risk-adjusted return. We find that the MVS portfolio, for any given constraints considered, obtains a higher annualized return, Sharpe-ratio and adjusted Sharpe-ratio, compared to GMVP and the MV portfolio. The MVS portfolio also outperforms the MVSK portfolio for these performance measures, except when the high or the low turnover constraint is imposed. The MVSK portfolio yields mixed results as the portfolio only outperforms the MV portfolio when the high diversification constraint, or any combination with turnover constraint, is imposed. This may indicate that kurtosis to some extent steers the portfolio allocation away from stocks with high returns at the cost of focusing on minimizing kurtosis, or that the estimation error related to this moments leads to a misleading objective function.

The higher-moment portfolios are riskier in terms of annualized standard deviation, than GMVP and the MV portfolio. Yet, the tail risk is lower for MVS as it obtains a lower expected shortfall for any given constraints compared to the MV portfolio, and a lower expected shortfall in general for GMVP. The MVSK only obtains a better expected shortfall than the benchmark portfolios

given a turnover constraint, this indicates that incorporating kurtosis into portfolio selection does not improve the out-of-sample tail risk. With the high turnover constraint or the high diversification- and low turnover constraint, the MVS portfolio also outperforms MVS in terms of expected shortfall. Yet, for any other constraint the MVS portfolio has a better expected shortfall and also a lower annualized standard deviation, compared to the MVS portfolio. Thus we conclude that MVS clearly dominates the traditional Markowitz portfolio, as well as the MVS portfolio, when it comes to risk management.

From the aspired moments plots, shown in section 5.3 and appendix XX, we observe the higher-moment portfolios having a slightly higher variance, but in return have more optimal skewness and/or kurtosis. The out-of-sample test for the mean, variance, skewness and kurtosis reveals that the MVS portfolio over most constraints obtains the most optimal out-of-sample mean, skewness and kurtosis, while GMVP consistently obtains the lowest variance. There is a trend of the out-of-sample mean increasing when the portfolios are rebalanced, and all the portfolios obtain the highest respective out-of-sample mean under daily rebalancing with a high diversification constraint imposed. The MVS- and MVS portfolio have the highest variances for any given constraints, in fact both of the portfolios even have significantly higher variance than the MV portfolio for certain strong constraints.

There is a clear tendency of both the MVS- and MVS portfolio obtaining higher out-of-sample skewness compared to the other portfolios. This relationship is consistent with the differences in in-sample skewness, as the higher-moment portfolios incorporate skewness in the asset allocation framework. As for kurtosis, the higher-moment portfolios obtain, in general, significantly lower kurtosis compared to the MV portfolio with basic constraints. This finding is somewhat diffuse as the allocation framework of the MVS portfolio does not explicitly take into account portfolio kurtosis.

On the basis of this thesis' findings we recommend the incorporation of skewness into the traditional Markowitz framework, even if the investor is limited to a certain degree of turnover or a required diversification target. Ideally the MVS portfolio is implemented based on frequent rebalancing, but this strategy also incur higher transaction costs and the trade-off has to be considered. Finally, the higher-moment portfolios do what they promise to a certain extent; our findings show that the MVS- and MVS portfolio obtain significantly higher out-of-sample skewness and lower out-of-sample kurtosis, given most constraints, at the cost of higher variance.

## 8 Limitations

Although it is encouraging to discover that by incorporating skewness into the portfolio selection process the investor can improve the out-of-sample return and risk-adjusted return, we caution readers that the improved performance can be driven by estimation error, selection bias, and/or the length of the estimation window, frequency of revisions, and holding period.

The moments are prone to estimation error, even though we reduce estimation error to some extent by Bayesian shrinkage. The improved estimates still contains error so that the allocation of the higher-moment portfolios could be driven by rare outliers, some of which can easily be observed in the forecasts of skewness and kurtosis in Figure 6. For instance, a positive skewed stock that exhibit extra-ordinary right tail performance over a one year estimation period might be overvalued.

In addition, our results are prone to selection bias as the selected S&P Global 1200-index consists of multi-national, well-established and stable companies with a long history of consistent growth and dividend payments. These stocks often have a relatively low volatility and usually require lower risk tolerance, while growth stocks usually require an higher risk tolerance. Consequently, the constructed portfolios can have different characteristics for different sets of stocks and assets.

Constructed portfolios with different length of the estimation- and holding periods are likely to exhibit different characteristics. For instance, many investors bought into the market, prior to the financial crisis in 2007, when the market was booming and showing positive skewness, and experienced a massive decline after the financial crisis 2008-2009. Thus, a too narrow estimation window may encourage risk-seeking behavior and a too wide estimation window may encourage a more risk-averse behavior.

We highlight two additional suggestions for further research. First, we believe it is possible to make progress on the last two limitations to generalize the results. This involves an analysis of more portfolios with various length of the estimation- and holding periods on different sets of stocks and asset classes. We further suggest an analysis of the length of the holding period to reflect the trading costs; because a shorter holding period should, in theory, better reflect the investor's preference for statistical moments, but at a higher cost. Secondly, PGP is an adequate framework, but it relies on good forecasts of moments. Hence, we suggest research on more robust estimators for higher-moments.

## References

- David Ardia, Kris Boudt, Peter Carl, Katharine M Mullen, and Brian G Peterson. Differential Evolution with DEoptim. *The R Journal*, 3(1):27–34, 2011. ISSN 20734859. URL [http://journal.r-project.org/archive/2011-1/RJournal\\_{ }2011-1\\_{ }Ardia{~}et{~}al.pdf](http://journal.r-project.org/archive/2011-1/RJournal_{ }2011-1_{ }Ardia{~}et{~}al.pdf).
- David Ardia, Kris Boudt, Katharine Mullen, and Brian Peterson. Large-Scale Portfolio Optimization with DEoptim. pages 69–77, 2014. URL <http://www.universal-publishers.com/book.php?method=ISBN{&}book=1627345035>.
- Fred D. Arditti. Risk and the Required Return on Equity. *The Journal of Finance*, 22(1):19–36, 1967.
- Fred D. Arditti. Risk and the required return on equity. *Journal of Finance*, 22(1):19–36, 1971.
- Fred D. Arditti and Haim Levy. Portfolio Efficiency Analysis in Three Moments: The Multiperiod Case. *The Journal of Finance*, 30(3):797–809, 1975.
- Gustavo M. Athayde and Renato G. Flores. Finding a maximum skewness portfolio - A general solution to three-moments portfolio choice. *Journal of Economic Dynamics and Control*, 28(7):1335–1352, 2004. ISSN 01651889. doi: 10.1016/S0165-1889(02)00084-2.
- Xiaoxin W Beardsley, Brian Field, and Mingqing Xiao. Mean-Variance-Skewness-Kurtosis Portfolio Optimization with Return and Liquidity. *Communications in Mathematical Finance*, 1(1):13–49, 2012.
- Gary P. Brinson, L. Randolph Hood, and Gilbert L. Beebower. Determinants of Portfolio Performance. *Financial Analysts Journal*, 51(1):133–138, 1995. ISSN 0015-198X. doi: 10.2469/faj.v51.n1.1869.
- Mark Britten-Jones. The sampling error in estimates of mean-variance efficient portfolio weights. *Journal of Finance*, 54(2):655–671, 1999. ISSN 00221082. doi: 10.1111/0022-1082.00120.
- Patrick L. Brockett and James R. Garven. A Reexamination of the Relationship Between Preferences and Moment Orderings by Rational Risk-Averse Investors. *The GENEVA Papers on Risk and Insurance - Theory*, 23(2):127–137, 1998. ISSN 0926-4957. doi: 10.1023/A:1008674127340. URL <http://dx.doi.org/10.1023/A:1008674127340>.
- Chris Brooks, Simon P Burke, Saeed Heravi, and Gita Persaud. Autoregressive Conditional Kurtosis. *Journal of Financial Econometrics*, 3, 2005.

- Chris Brooks, Aleš Černý, and Joëlle Miffre. Optimal hedging with higher moments. *Journal of Futures Markets*, 32(10):909–944, 2011.
- Honghui Chen, Hemang Desai, and Srinivasan Krishnamurthy. A First Look at Mutual Funds That Use Short Sales. *Journal of Financial and Quantitative Analysis*, 48(03):761–787, 2013. ISSN 0022-1090. doi: 10.1017/S0022109013000264. URL <http://www.journals.cambridge.org/abstract%7B%7D%7DS0022109013000264>.
- Vijay Kumar Chopra and William T Ziemba. The effect of errors in means , variances , and covariances on optimal portfolio choices. *Journal of Portfolio Management*, 19(2):6, 1993.
- Pomchai Chunchachinda, Krishnan Dandapani, Shahid Hamid, and Arun J Prakash. Portfolio selection and skewness: Evidence from international stock markets. *journal of Banking & Finance*, 21:143–167, 1997. ISSN 03784266. doi: 10.1016/S0378-4266(96)00032-5.
- Roger Clarke, Harindra de Silva, and Steven Thorley. Minimum-Variance Portfolio Composition. *The Journal of Portfolio Management*, 37(2):31–45, 2011. ISSN 0095-4918. doi: 10.3905/jpm.2011.37.2.031.
- Roger G Clarke, Harindra de Silva, and Steven Thorley. Minimum-Variance Portfolios in the U.S. Equity Market. *The Journal of Portfolio Management*, 33(1):10–24, 2006. ISSN 0095-4918. doi: 10.3905/jpm.2006.661366.
- Paul H. Cootner. Review: The Random Character of Stock Market Prices. *The Journal of Business*, 39(4):532–535, 1966.
- Ryan J Davies, Harry M Kat, and Sa Lu. Fund of hedge funds portfolio selection: A multiple-objective approach. *Journal of Derivatives & Hedge Funds*, 15(2):91–115, 2009. ISSN 1753-9641. doi: 10.1057/jdhf.2009.1. URL <http://www.palgrave-journals.com/doi/finder/10.1057/jdhf.2009.1>.
- Ajit Dayanandan and Minh Lam. Portfolio Rebalancing – Hype or Hope? *The Journal of Business Inquiry*, 14(2):79–92, 2015.
- Victor DeMiguel, Lorenzo Garlappi, and Raman Uppal. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies*, 22(5):1915–1953, 2009. ISSN 14657368. doi: 10.1093/rfs/hhm075.
- David J Disatnik and Simon Benninga. Shrinking the Covariance Matrix - Simpler is Better. *Journal of Portfolio Management*, 33(4):56–63, 2007. ISSN 0095-4918. doi: 10.3905/jpm.2007.690606.

- Edwin J. Elton and Martin J. Gruber. Estimating the Dependence Structure of Sharpe Prices: Implications for Portfolio Selection, 1973.
- Frank J. Fabozzi, Sergio M. Focardi, and Petter N. Kolm. *Quantitative Equity Investing*. Hoboken, New Jersey, 2010. URL <https://books.google.no/books?id=eAmGGITptUMC&printsec=frontcover&hl=no#v=onepage&q&f=false>.
- Eugene F. Fama. The Behavior of Stock-Market Prices. *The Journal of Business*, 38(1):34–105, 1965.
- Kenneth R French. Fama French Research Data Factors, 2017. URL [http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/ftp/F-F\\_Research\\_Data\\_Factors\\_CSV.zip](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/ftp/F-F_Research_Data_Factors_CSV.zip).
- Kenneth R French, G William Schwert, and Robert F Stambaugh. Expected Stock Returns and Volatility. *Journal of Financial Economics*, 19:3–29, 1987.
- Henrik Alsing Friberg. Rmosek : The R-to-MOSEK Optimization Interface, 2014. URL <https://mran.microsoft.com/web/packages/Rmosek/Rmosek.pdf>.
- Peter A. Frost and James E. Savarino. For better performance constrain portfolio weights. *The Journal of Portfolio Management*, 15(1):29–34, 1988. URL 10.3905/jpm.1988.409181.
- Campbell R Harvey and Akhtar Siddique. Conditional Skewness in Asset Pricing Tests. *The Journal of Finance*, 55(3):1263–1295, 2000. ISSN 0022-1082. doi: 10.1111/0022-1082.00247.
- Campbell R. Harvey, John C. Liechty, Merrill W. Liechty, and Peter Müller. Portfolio Selection with Higher Moments. *Quantitative Finance*, 10(5):469–485, 2010. ISSN 1469-7688. doi: 10.1080/14697681003756877.
- John Haslem. Mutual Fund Service Advantages. In *Mutual Funds: Risk and Performance Analysis for Decision Making*. Wiley-Blackwell, 2003. ISBN 978-0-631-21561-5. URL <http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0631215611.html>.
- Thomas Heinze. Markowitz 3 . 0 : Including Diversification Targets in Portfolio Optimization via Diversification Functions. *SSRN Electronic Journal*, pages 1–20, 2016.
- Soosung Hwang and Stephen Satchell, E. Modelling emerging market risk premia using higher moments. In *Return Distributions in Finance*, pages 75–117. Butterworth-Heinemann, 2001. URL <http://www.sciencedirect.com/science/book/9780750647519>.

- J. D. Jobson and B Korkie. Estimation for Markowitz Efficient Portfolios. *Journal of the American Statistical Association*, 75(371):544–554, 1980. ISSN 0162-1459. doi: 10.1080/01621459.1980.10477507.
- Eric Jondeau and Michael Rockinger. Optimal portfolio allocation under higher moments. *European Financial Management*, 12(1):29–55, 2006. ISSN 13547798. doi: 10.1111/j.1354-7798.2006.00309.x.
- Eric Jondeau, Ser-Huang Poon, and Michael Rockinger. Statistical Properties of Financial Market Data. *Financial Modeling Under Non-Gaussian Distributions*, 2007.
- Philippe Jorion. Bayes-Stein Estimation for Portfolio Analysis. *The Journal of Financial and Quantitative Analysis*, 21(3):279, 1986. ISSN 00221090. doi: 10.2307/2331042. URL <http://www.jstor.org/stable/2331042?origin=crossref>.
- Emmanuel Jurczenko and Bertrand Maillet. Between asset pricing and asset allocation. In *Multi-moment asset allocation and pricing models*, pages 113–163. 2006.
- Emmanuel Jurczenko, Bertrand Maillet, and Paul Merlin. Hedge Fund Portfolio Selection with Higher-order Moments: A Nonparametric Mean-Variance-Skewness-Kurtosis Efficient Frontier. *Multi-moment Asset Allocation and Pricing Models*, 1:51–66, 2015. ISSN 1556-5068. doi: 10.1002/9781119201830.ch3.
- Alex Kane. Skewness Preference and Portfolio Choice. *Journal of Financial & Quantitative Analysis*, 17(1):15–25, 1982. ISSN 00221090. doi: 10.2307/2330926. URL <http://www.lib.lsu.edu/apps/onoffcampus.php?url=http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=5721966&site=ehost-live&scope=site>.
- Giri Kumar, Tayi, and Paul A. Leonard. Bank Balance-Sheet Management: An Alternative Multi-Objective Model. *The Journal of the Operational Research Society*, 39(4):401–410, 1988. doi: 10.1057/palgravejors.2601357. URL <http://www.jstor.org/stable/2582120>.
- Kin Keung Lai, Lean Yu, and Shouyang Wang. Mean-variance-skewness-kurtosis-based portfolio optimization. In *First International Multi-Symposiums on Computer and Computational Sciences, IMSCCS'06*, volume 2, pages 292–297, 2006. ISBN 0769525814. doi: 10.1109/IMSCCS.2006.252.
- Tsong Yue Lai. Portfolio selection with skewness: A multiple-objective approach. *Review of Quantitative Finance and Accounting*, 1(3):293–305, 1991. ISSN 0924865X. doi: 10.1007/BF02408382.

- Olivier Ledoit and Michael Wolf. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance*, 10(5):603–621, 2003. ISSN 09275398. doi: 10.1016/S0927-5398(03)00007-0.
- Olivier Ledoit and Michael Wolf. Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, 15:850–859, 2008.
- Olivier Ledoit and Michael Wolf. Robust Performances Hypothesis Testing With the Variance. *Wilmott*, 2011(55):86–89, 2011. ISSN 15406962. doi: 10.1002/wilm.10036. URL <http://doi.wiley.com/10.1002/wilm.10036>.
- S. Lummer and M. Reipe. The Role of Asset Allocation in Portfolio Management. *Global Asset Allocation: Techniques for Optimizing Portfolio Management*, 1994.
- Tongshu Ma and Ravi Jagannathan. Risk Reduction in Large Portfolios: A Role for Portfolio Weight Constraints. 2001.
- Hans Malmsten and Timo Teräsvirta. And, Stylized Facts of Financial Time Series Volatility, Three Popular Models of. *SSE/EFI Working Paper Series in Economics and Finance*, (563), 2004. URL <http://www-stat.wharton.upenn.edu/~steele/Resources/FTSResources/StylizedFacts/MalmstenTerasvirta04.pdf>.
- Benoit Mandelbrot. The Variation of Certain Speculative Prices. *The Journal of Business*, 36(4):394–419, 1963.
- Harry Markowitz. Portfolio Selection. *The Journal of Finance*, 7(1):77–91, 1952.
- Lionel Martellini and Volker Ziemann. Improved estimates of higher-order comonents and implications for portfolio selection. *Review of Financial Studies*, 23(4):1467–1502, 2010. ISSN 08939454. doi: 10.1093/rfs/hhp099.
- Edward L. Melnick and Brian S. Everitt. Value at Risk (VaR) and Risk Measures. In *Encyclopedia of Quantitative Risk Analysis and Assessment*, page 1825. 2008.
- Robert C. Merton. An Analytic Derivation of the Efficient Portfolio Frontier. *The Journal of Financial and Quantitative Analysis*, 7(4):1851–1872, 1972.
- Richard Michaud. The Markowitz Optimization Enigma: Is ‘Optimized’ Optimal? *Financial Analysts Journal*, 45(1):31–42, 1989. ISSN 0015-198X. doi: 10.2469/faj.v45.n1.31.



- Michael B. Miller. Chapter 3. Basic Statistics. In *Mathematics and Statistics for Financial Risk Management*, pages 53–56. John Wiley & Sons, Inc., Hoboken, New Jersey, 2 edition, 2014. ISBN 978-1-118-75029-2.
- R. R. Officer. The Distribution of Stock Returns. *Journal of the American Statistical Association*, 67(340):807–812, 1972. ISSN 01621459. doi: 10.2307/2284641.
- Brian Peterson, Peter Carl, Kris Boudt, Ross Bennett, Guy Yollin, and R Douglas Martin. Package ‘PortfolioAnalytics’ (R). Technical report, 2015. URL <https://cran.r-project.org/web/packages/PortfolioAnalytics/PortfolioAnalytics.pdf>.
- Arun J. Prakash, Chun Hao Chang, and Therese E. Pactwa. Selecting a portfolio with skewness: Recent evidence from US, European, and Latin American equity markets. *Journal of Banking and Finance*, 27(7):1375–1390, 2003. ISSN 03784266. doi: 10.1016/S0378-4266(02)00261-3.
- Kenneth V Price, Rainer M Storn, and Jouni A Lampinen. *Differential Evolution: A Practical Approach to Global Optimization*, volume 28. 2005. ISBN 978-3-540-31306-9. doi: 10.1007/3-540-31306-0. URL <http://www.springer.com/mathematics/book/978-3-540-20950-8?detailsPage=reviews>.
- Juliane Proelss and Denis Schweizer. Polynomial goal programming and the implicit higher moment preferences of US institutional investors in hedge funds. *Financial Markets and Portfolio Management*, 28(1):1–28, 2014. ISSN 15554961. doi: 10.1007/s11408-013-0221-x.
- Jean Charles Richard and Thierry Roncalli. Smart Beta: Managing Diversification of Minimum Variance Portfolios. *Risk-Based and Factor Investing*, (2):31–63, 2015. doi: 10.1016/B978-1-78548-008-9.50002-2.
- Paul a. Samuelson. The Fundamental Approximation Theorem of Portfolio Analysis in terms of Means, Variances and Higher Moments. *The Review of Economic Studies*, 37(4):537–542, 1970. ISSN 0034-6527. doi: 10.2307/2296483. URL <http://www.jstor.org/stable/2296483>. URL <http://www.jstor.org/stable/pdfplus/2296483.pdf?acceptTC=true>.
- Arthur C. Sanderson and Jingqiao Zhang. JADE: Adaptive differential evolution with optional external archive. *IEEE Transactions on Evolutionary Computation*, 13(5):945–958, 2009. ISSN 1089778X. doi: 10.1109/TEVC.2009.2014613.
- K. Saranya and P. Krishna Prasanna. Portfolio Selection and Optimization with Higher Moments: Evidence from the Indian Stock Market. *Asia-Pacific Financial Markets*, 21(2): 133–149, 2014. ISSN 13872834. doi: 10.1007/s10690-014-9180-0.

- Bernd Scherer. A New Look at Minimum Variance Investing. *SSRN Electronic Journal*, (August):1–23, 2010. ISSN 1556-5068. doi: 10.2139/ssrn.1681306. URL <http://www.ssrn.com/abstract=1681306>.
- John Schreiner. Portfolio Revision: A Turnover-Constrained Approach. *Financial Management*, 9(1):67–75, 1980.
- William F. Sharpe. A Simplified Model for Portfolio Analysis. *Management Science*, 9(2):277–293, 1963. ISSN 0025-1909. doi: 10.1287/mnsc.9.2.277. URL <http://pubsonline.informs.org/doi/abs/10.1287/mnsc.9.2.277>.
- A Signer and Laurent Favre. The difficulties of measuring the benefits of hedge funds. *The Journal of Alternative Investments*, 33(0), 2002. URL <http://www.iijournals.com/doi/abs/10.3905/jai.2002.319041>.
- Tihana Škrinjarić. Portfolio Selection with Higher Moments and Application on Zagreb Stock Exchange. *Zagreb International Review of Economics & Business*, 16(1):65–78, 2013.
- Stoyan Stoyanov. Advantages and shortcomings of minimum variance portfolios, 2011. URL [http://www.edhec-risk.com/latest{\\_}news/featured{\\_}analysis/RISKArticle.2011-11-15.4803?newsletter=yes](http://www.edhec-risk.com/latest{_}news/featured{_}analysis/RISKArticle.2011-11-15.4803?newsletter=yes).
- Qian Sun and Yuxing Yan. Skewness persistence with optimal portfolio selection. *Journal of Banking and Finance*, 27(6):1111–1121, 2003. ISSN 03784266. doi: 10.1016/S0378-4266(02)00247-9.
- Ruey S Tsay. Value at Risk, Expected Shortfall & Risk Management, 2013. URL <http://faculty.chicagobooth.edu/ruey.tsay/teaching/bs41202/sp2013/lec9-13.pdf>.
- Ursula Walther. Diversification of higher moments in stock portfolios - an empirical investigation. *Journal of Applied Operational Research*, 6(4):255–263, 2014.
- Qi-fa Xu, Cui-xia Jiang, and Pu Kang. Dynamic Portfolio Selection with Higher Moments Risk Based on Polynomial Goal Programming. *2007 International Conference on Management Science and Engineering*, pages 2152–2157, 2008.

# Appendix

## A Components for shrinkage intensities

Ledoit and Wolf (2003) and Martellini and Ziemann (2010) derive the formulas for  $\pi_{n,t}$ ,  $\rho_{n,t}$  and  $\gamma_{n,t}$ , for  $n = 2, 3, 4$ , by the delta-method. We present the closed-form expressions of the shrinkage intensity for the covariance matrix, derived by Ledoit and Wolf (2003), and for the coskewness- and cokurtosis matrix, derived by Martellini and Ziemann (2010), and briefly comment on the components. The reader is referred to Martellini and Ziemann (2010) for a thorough derivation. Note that we use the same notation as defined in methodology, i.e. length of estimation window is  $L$ , the sample estimates are denoted  $S$  and the single-index estimates are denoted by  $SI$ .

The sum of asymptotic variances of the sample estimates of moment  $n$  at time  $t$  is given by:

$$\hat{\pi}_{2,t} = \sum_{i=1}^N \sum_{j=1}^N \pi_{ij,t} \quad \hat{\pi}_{3,t} = \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \hat{\pi}_{ijk,t} \quad \hat{\pi}_{4,t} = \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \sum_{l=1}^N \hat{\pi}_{ijkl,t} \quad (65)$$

where

$$\begin{aligned} \hat{\pi}_{ij,t} &= \text{AsyVar}[\sqrt{L}\sigma_{ij,t}^S] = \frac{1}{L} \sum_{x=t-L}^{t-1} \{(r_{i,x} - \bar{r}_{i,t})(r_{j,x} - \bar{r}_{j,t}) - \sigma_{ij,t}^S\}^2 \\ \hat{\pi}_{ijk,t} &= \text{AsyVar}[\sqrt{L}s_{ijk,t}^S] = \frac{1}{L} \sum_{x=t-L}^{t-1} \{(r_{i,x} - \bar{r}_{i,t})(r_{j,x} - \bar{r}_{j,t})(r_{k,x} - \bar{r}_{k,t}) - s_{ijk,t}^S\}^2 \\ \hat{\pi}_{ijkl,t} &= \text{AsyVar}[\sqrt{L}k_{ijkl,t}^S] = \frac{1}{L} \sum_{x=t-L}^{t-1} \{(r_{i,x} - \bar{r}_{i,t})(r_{j,x} - \bar{r}_{j,t})(r_{k,x} - \bar{r}_{k,t})(r_{l,x} - \bar{r}_{l,t}) - k_{ijkl,t}^S\}^2 \end{aligned} \quad (66)$$

where  $\sigma_{ij,t}^S$ ,  $s_{ijk,t}^S$  and  $k_{ijkl,t}^S$  are the sample estimates of the covariance, coskewness and cokurtosis, respectively, for time  $t$ . Connecting (26) and (27) we observe that  $\hat{\pi}_n$ , for  $n = 2, 3, 4$ , increases as a result of an increase in the squared difference between the product of the excess returns of one or several assets and the the sample estimate. This illustrates that as  $\hat{\pi}_n$  increases, ceteris paribus, we weight more in the single-index estimate, and less in the sample estimate as its bias increases.

The misspecification of the single-index estimate of moment  $n$  at time  $t$  is given by:

$$\widehat{\gamma}_{2,t} = \sum_{i=1}^N \sum_{j=1}^N \widehat{\gamma}_{ij,t} \quad \widehat{\gamma}_{3,t} = \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \widehat{\gamma}_{ijk,t} \quad \widehat{\gamma}_{4,t} = \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \sum_{l=1}^N \widehat{\gamma}_{ijkl,t} \quad (67)$$

where

$$\begin{aligned} \widehat{\gamma}_{ij,t} &= (\sigma_{ij,t}^{SI} - \sigma_{ij,t}^S)^2 & \forall i, j \\ \widehat{\gamma}_{ijk,t} &= (s_{ijk,t}^{SI} - s_{ijk,t}^S)^2 & \forall i, j, k \\ \widehat{\gamma}_{ijkl,t} &= (k_{ijkl,t}^{SI} - k_{ijkl,t}^S)^2 & \forall i, j, k, l \end{aligned} \quad (68)$$

The greater the difference is between the sample and single-index estimate, the larger the denominator of the shrinkage intensity is, and the less weight put in the single-index estimate of the respective moment.

The sum of asymptotic covariances of the single-index estimates and sample estimates for moment  $n$  is given by:

$$\widehat{\rho}_{2,t} = \sum_{i=1}^N \sum_{j=1}^N \widehat{\rho}_{ij,t} \quad \widehat{\rho}_{3,t} = \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \widehat{\rho}_{ijk,t} \quad \widehat{\rho}_{4,t} = \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \sum_{l=1}^N \widehat{\rho}_{ijkl,t} \quad (69)$$

where

$$\begin{aligned} \widehat{\rho}_{ij,t} &= \text{AsyCov}[\sqrt{L}\sigma_{ij,t}^{SI}, \sqrt{L}\sigma_{ij,t}^S] & \forall i, j \\ \widehat{\rho}_{ijk,t} &= \text{AsyCov}[\sqrt{L}s_{ijk,t}^{SI}, \sqrt{L}s_{ijk,t}^S] & \forall i, j, k \\ \widehat{\rho}_{ijkl,t} &= \text{AsyCov}[\sqrt{L}k_{ijkl,t}^{SI}, \sqrt{L}k_{ijkl,t}^S] & \forall i, j, k, l \end{aligned} \quad (70)$$

where  $\sigma_{ij,t}^y$ ,  $s_{ijk,t}^y$  and  $k_{ijkl,t}^y$  are the time  $t$  estimates of the covariance, coskewness and cokurtosis, respectively, where  $y = S, SI$ . We observe that as the difference between the estimates increases, the shrinkage intensity decreases and we place more weight in the sample estimate and less in the single-index estimate.

The sample estimate of the variance, skewness and kurtosis, of each asset, is the same as the single-index estimate, i.e.  $\sigma_{ii}^S = \sigma_{ii}^{SI}$ ,  $s_{iii}^S = s_{iii}^{SI}$  and  $k_{iiii}^S = k_{iiii}^{SI}$ . Hence the asymptotic covariance of the single-index and sample estimate of each these moments equals the asymptotic variance of the corresponding moment, given for asset  $i$  by:

$$\begin{aligned} \widehat{\rho}_{ii,t} &= \text{AsyCov}[\sqrt{L}\sigma_{ii,t}^{SI}, \sqrt{L}\sigma_{ii,t}^S] = \text{AsyVar}[\sqrt{L}\sigma_{ii,t}^S] = \widehat{\pi}_{ii,t} & \forall i \\ \widehat{\rho}_{iii,t} &= \text{AsyCov}[\sqrt{L}s_{iii,t}^{SI}, \sqrt{L}s_{iii,t}^S] = \text{AsyVar}[\sqrt{L}s_{iii,t}^S] = \widehat{\pi}_{iii,t} & \forall i \\ \widehat{\rho}_{iiii,t} &= \text{AsyCov}[\sqrt{L}k_{iiii,t}^{SI}, \sqrt{L}k_{iiii,t}^S] = \text{AsyVar}[\sqrt{L}k_{iiii,t}^S] = \widehat{\pi}_{iiii,t} & \forall i \end{aligned} \quad (71)$$

The closed-form solution of the asymptotic covariance of the single-index and sample estimates of the covariance, coskewness and cokurtosis, are more complex as the single-index estimates consists of betas and the market moments, e.g.  $\sigma_{ij,t}^{SI} = \beta_{i,t}\beta_{j,t}\sigma_{M,t}$ .

$$\begin{aligned}
\widehat{\rho}_{ij,t} &= \frac{\sigma_{jM,t}}{\sigma_{M,t}^2} \text{AsyCov}[\sqrt{L}\sigma_{iM,t}, \sqrt{L}\sigma_{ij,t}] \\
&+ \frac{\sigma_{iM,t}}{\sigma_{M,t}^2} \text{AsyCov}[\sqrt{L}\sigma_{jM,t}, \sqrt{L}\sigma_{ij,t}] \\
&- \frac{\sigma_{iM,t}\sigma_{jM,t}}{(\sigma_{M,t}^2)^2} \text{AsyCov}[\sqrt{L}\sigma_{M,t}^2, \sqrt{L}\sigma_{ij,t}] \\
\widehat{\rho}_{ijk,t} &= \frac{\sigma_{jM,t}\sigma_{kM,t}s_{M,t}^3}{(\sigma_{M,t}^2)^3} \text{AsyCov}[\sqrt{L}\sigma_{iM,t}, \sqrt{L}s_{ijk,t}] \\
&+ \frac{\sigma_{iM,t}\sigma_{kM,t}s_{M,t}^3}{(\sigma_{M,t}^2)^3} \text{AsyCov}[\sqrt{L}\sigma_{jM,t}, \sqrt{L}s_{ijk,t}] \\
&+ \frac{\sigma_{iM,t}\sigma_{jM,t}s_{M,t}^3}{(\sigma_{M,t}^2)^3} \text{AsyCov}[\sqrt{L}\sigma_{kM,t}, \sqrt{L}s_{ijk,t}] \\
&- 3 * \frac{\sigma_{iM,t}\sigma_{jM,t}\sigma_{kM,t}s_{M,t}^3}{(\sigma_{M,t}^2)^4} \text{AsyCov}[\sqrt{L}\sigma_{M,t}^2, \sqrt{L}s_{ijk,t}] \\
&+ \frac{\sigma_{iM,t}\sigma_{jM,t}\sigma_{kM,t}}{(\sigma_{M,t}^2)^3} \text{AsyCov}[\sqrt{L}s_{M,t}^3, \sqrt{L}s_{ijk,t}] \tag{72} \\
\widehat{\rho}_{ijkl,t} &= \frac{\sigma_{jM,t}\sigma_{kM,t}\sigma_{lM,t}k_{m,t}^4}{(\sigma_{M,t}^2)^4} \text{AsyCov}[\sqrt{L}\sigma_{iM,t}, \sqrt{L}k_{ijkl,t}] \\
&+ \frac{\sigma_{iM,t}\sigma_{kM,t}\sigma_{lM,t}k_{m,t}^4}{(\sigma_{M,t}^2)^4} \text{AsyCov}[\sqrt{L}\sigma_{jM,t}, \sqrt{L}k_{ijkl,t}] \\
&+ \frac{\sigma_{iM,t}\sigma_{jM,t}\sigma_{lM,t}k_{m,t}^4}{(\sigma_{M,t}^2)^4} \text{AsyCov}[\sqrt{L}\sigma_{kM,t}, \sqrt{L}k_{ijkl,t}] \\
&+ \frac{\sigma_{iM,t}\sigma_{jM,t}\sigma_{kM,t}k_{m,t}^4}{(\sigma_{M,t}^2)^4} \text{AsyCov}[\sqrt{L}\sigma_{lM,t}, \sqrt{L}k_{ijkl,t}] \\
&+ 4 * \frac{\sigma_{iM,t}\sigma_{jM,t}\sigma_{kM,t}\sigma_{lM,t}k_{m,t}^4}{(\sigma_{M,t}^2)^5} \text{AsyCov}[\sqrt{L}\sigma_{M,t}^2, \sqrt{L}k_{ijkl,t}] \\
&+ \frac{\sigma_{iM,t}\sigma_{jM,t}\sigma_{kM,t}\sigma_{lM,t}}{(\sigma_{M,t}^2)^3} \text{AsyCov}[\sqrt{L}k_{M,t}^4, \sqrt{L}k_{ijkl,t}] \\
&+ r_{ijkl,t}
\end{aligned}$$

for  $z = i, j, M$ .

for  $\forall i, j, k, l$  and the idiosyncratic terms of the cokurtosis elements are:

$$\begin{aligned}
r_{iii,t} &= 3 * \left( \frac{\sigma_{jM,t} \bar{\varepsilon}_{ii,t}}{\sigma_{M,t}^2} \text{AsyCov}[\sqrt{L}\sigma_{iM,t}, \sqrt{L}k_{iii,t}] + \frac{\sigma_{jM,t} \bar{\varepsilon}_{jj,t}}{\sigma_{M,t}^2} \text{AsyCov}[\sqrt{L}\sigma_{jM,t}, \sqrt{L}k_{iii,t}] \right. \\
&\quad \left. - \frac{\sigma_{iM,t} \sigma_{jM,t} \bar{\varepsilon}_{ii,t}}{(\sigma_{M,t}^2)^2} \text{AsyCov}[\sqrt{L}\sigma_{M,t}^2, \sqrt{L}k_{iii,t}] + \frac{\sigma_{iM,t} \sigma_{jM,t}}{\sigma_{M,t}^2} \text{AsyCov}[\sqrt{L}\bar{\varepsilon}_{ii,t}, \sqrt{L}k_{iii,t}] \right) \\
r_{iij,t} &= \frac{2\sigma_{iM,t} \bar{\varepsilon}_{jj,t}}{\sigma_{M,t}^2} \text{AsyCov}[\sqrt{L}\sigma_{iM,t}, \sqrt{L}k_{iij,t}] + \frac{2\sigma_{jM,t} \bar{\varepsilon}_{ii,t}}{\sigma_{M,t}^2} \text{AsyCov}[\sqrt{L}\sigma_{jM,t}, \sqrt{L}k_{iij,t}] \\
&\quad - \frac{\sigma_{iM,t}^2 \bar{\varepsilon}_{jj,t}}{(\sigma_{M,t}^2)^2} \text{AsyCov}[\sqrt{L}\sigma_{M,t}^2, \sqrt{L}k_{iij,t}] - \frac{\sigma_{jM,t}^2 \bar{\varepsilon}_{ii,t}}{(\sigma_{M,t}^2)^2} \text{AsyCov}[\sqrt{L}\sigma_{M,t}^2, \sqrt{L}k_{iij,t}] \\
&\quad + \frac{\sigma_{iM,t}^2}{\sigma_{M,t}^2} \text{AsyCov}[\sqrt{L}\bar{\varepsilon}_{jj,t}, \sqrt{L}k_{iij,t}] + \frac{\sigma_{jM,t}^2}{\sigma_{M,t}^2} \text{AsyCov}[\sqrt{L}\bar{\varepsilon}_{ii,t}, \sqrt{L}k_{iij,t}] \\
&\quad + \bar{\varepsilon}_{jj,t} \text{AsyCov}[\sqrt{L}\bar{\varepsilon}_{ii,t}, \sqrt{L}k_{iij,t}] + \bar{\varepsilon}_{ii,t} \text{AsyCov}[\sqrt{L}\bar{\varepsilon}_{jj,t}, \sqrt{L}k_{iij,t}] \\
r_{iijk,t} &= \frac{\sigma_{kM,t} \bar{\varepsilon}_{ii,t}}{\sigma_{M,t}^2} \text{AsyCov}[\sqrt{L}\sigma_{jM,t}, \sqrt{L}k_{iijk,t}] + \frac{\sigma_{jM,t} \bar{\varepsilon}_{ii,t}}{\sigma_{M,t}^2} \text{AsyCov}[\sqrt{L}\sigma_{kM,t}, \sqrt{L}k_{iijk,t}] \\
&\quad - \frac{\sigma_{jM,t} \sigma_{kM,t} \bar{\varepsilon}_{ii,t}}{(\sigma_{M,t}^2)^2} \text{AsyCov}[\sqrt{L}\sigma_{M,t}^2, \sqrt{L}k_{iijk,t}] + \frac{\sigma_{jM,t} \sigma_{kM,t}}{\sigma_{M,t}^2} \text{AsyCov}[\sqrt{L}\bar{\varepsilon}_{ii,t}, \sqrt{L}k_{iijk,t}] \\
r_{ijkl,t} &= 0
\end{aligned} \tag{73}$$

Note that  $\sigma_{M,t}^2 = \sigma_{MM,t}$ , so that the asymptotic covariance of the market variance and the sample estimates of covariance, coskewness and cokurtosis for all assets  $i, j, k, l$  are also given by (81), (82) and (83), respectively.

The AsyCov-terms are given by:

$$\begin{aligned}
& \text{AsyCov}[\sqrt{L}\sigma_{zM,t}, \sqrt{L}\sigma_{ij,t}] = \\
& \frac{1}{L} \sum_{x=t-L}^{t-1} \{(r_{z,x} - \bar{r}_{z,t})(r_{M,x} - \bar{r}_{M,t}) - \sigma_{zM,t}\} \{(r_{i,x} - \bar{r}_{i,t})(r_{j,x} - \bar{r}_{j,t}) - \sigma_{ij,t}\} \\
& \text{AsyCov}[\sqrt{L}\sigma_{zM,t}, \sqrt{L}s_{ijk,t}] = \\
& \frac{1}{L} \sum_{x=t-L}^{t-1} \{(r_{z,x} - \bar{r}_{z,t})(r_{M,x} - \bar{r}_{M,t}) - \sigma_{zM,t}\} \{(r_{i,x} - \bar{r}_{i,t})(r_{j,x} - \bar{r}_{j,t})(r_{k,x} - \bar{r}_{k,t}) - s_{ijk,t}\} \\
& \text{AsyCov}[\sqrt{L}\sigma_{zM,t}, \sqrt{L}k_{ijkl,t}] = \\
& \frac{1}{L} \sum_{x=t-L}^{t-1} \{(r_{z,x} - \bar{r}_{z,t})(r_{M,x} - \bar{r}_{M,t}) - \sigma_{zM,t}\} \{(r_{i,x} - \bar{r}_{i,t})(r_{j,x} - \bar{r}_{j,t})(r_{k,x} - \bar{r}_{k,t})(r_{l,x} - \bar{r}_{l,t}) - k_{ijkl,t}\} \\
& \text{AsyCov}[\sqrt{T}s_{M,t}^3, \sqrt{L}s_{ijk,t}] = \\
& \frac{1}{L} \sum_{x=t-L}^{t-1} \{(r_{M,x} - \bar{r}_{M,t})^3 - s_{M,t}^3\} \{(r_{i,x} - \bar{r}_{i,t})(r_{j,x} - \bar{r}_{j,t})(r_{k,x} - \bar{r}_{k,t}) - s_{ijk,t}\} \\
& \text{AsyCov}[\sqrt{T}k_{M,t}^4, \sqrt{L}k_{ijkl,t}] = \\
& \frac{1}{L} \sum_{x=t-L}^{t-1} \{(r_{M,x} - \bar{r}_{M,t})^3 - k_{M,t}^4\} \{(r_{i,x} - \bar{r}_{i,t})(r_{j,x} - \bar{r}_{j,t})(r_{k,x} - \bar{r}_{k,t})(r_{l,x} - \bar{r}_{l,t}) - k_{ijkl,t}\} \\
& \text{AsyCov}[\sqrt{T}\sigma_{zm}, \sqrt{L}k_{ijkl,t}] = \\
& \frac{1}{L} \sum_{x=t-L}^{t-1} \{(r_{z,x} - \bar{r}_{z,t})(r_{M,x} - \bar{r}_{M,t}) - \sigma_{zM,t}\} \{(r_{i,x} - \bar{r}_{i,t})(r_{j,x} - \bar{r}_{j,t})(r_{k,x} - \bar{r}_{k,t})(r_{l,x} - \bar{r}_{l,t}) - k_{ijkl,t}\} \\
& \text{AsyCov}[\sqrt{T}\bar{\varepsilon}_{ii,t}, \sqrt{L}k_{ijkl,t}] = \\
& \frac{1}{L} \sum_{x=t-L}^{t-1} \{(\bar{\varepsilon}_{ii,x}^2 - \sum_{y=t-L}^{t-1} \bar{\varepsilon}_{ii,y}^2)\} \{(r_{i,x} - \bar{r}_{i,t})(r_{j,x} - \bar{r}_{j,t})(r_{k,x} - \bar{r}_{k,t})(r_{l,x} - \bar{r}_{l,t}) - k_{ijkl,t}\}
\end{aligned} \tag{74}$$

for  $z = i, j, M$ .

## B Modifying turnover constraint for convex set

By imposing the turnover constraint (X) in the mean- or variance optimization, the problems become non-convex and we might lose accuracy and, definitely, efficiency. We regain convexity by modifying the turnover constraint, where the first step is to introduce auxiliary variables. We let  $b_i$  and  $s_i$  denote the percent-wise amount bought and sold of asset  $i$ , respectively, so that the difference of wealth invested from time  $t-1$  to  $t$  is described by:

$$x_{t,i} - b_{t,i} + s_{t,i} = x_{t-1,i} \quad \forall i \quad (75)$$

where the initial position  $x_{0,i}$  is found solving the optimization problem for the first portfolio revision, i.e. disregarding the turnover constraint. We impose the turnover constraint for the  $N$ -asset portfolio by precluding an average turnover above a specified turnover target of  $c_2$ :

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N (b_{t,i} + s_{t,i}) &\leq c_1 \\ b_{t,i}, s_{t,i} &\geq 0 \quad \forall i \end{aligned} \quad (76)$$

- Write about implication of last constraint

## C Standard error of hypothesis test

In this section we present the standard error of the HAC-hypothesis test. We assume the period of evaluation is  $T$ , where  $T$  is the time between the initial allocation and the end of the holding period of the last revision, for simplicity assumed to be  $t = 1, \dots, T$ . The asymptotic standard error is given by:

$$s(\hat{\Delta}_M) = \sqrt{\frac{\nabla^T f(\hat{v}) \hat{\Psi}_M \nabla f(\hat{v})}{T}} \quad (77)$$

where  $\nabla^T f(\hat{v})$  is the transposed gradient vector for function  $f$ , and  $\hat{\Psi}$  is estimated by the heteroscedastic and autocorrelation robust (HAC) kernel estimate.

The difference between the moments can be written as a function of the differences

$$f(\hat{v}_M) = \hat{\Delta}_M = h(\hat{v}_{x,M}) - h(\hat{v}_{y,M}) \quad (78)$$



where we assume that the moment  $\theta_l$  for portfolio  $l$  can be written as a function of up till  $M$  uncentered moments:

$$h(\hat{v}_{l,M}) = h(\hat{v}_l^1, \hat{v}_l^2, \dots, \hat{v}_l^M) \quad l = x, y \quad (79)$$

where  $M \geq 1$ ,  $h$  is a continuously differentiable function, and  $v_l^{(m)}$  is the uncentered  $m$ th-moment given by:

$$\hat{v}_l^{(m)} = \frac{1}{T} \sum_{t=1}^T R_{l,t}^m \quad l = x, y \quad (80)$$

We simplify the notation by setting the uncentered first moment as  $a_l$ , i.e.  $a_l = \hat{v}_l^{(1)}$ , the uncentered second moment as  $b_l$ , i.e.  $b_l = \hat{v}_l^{(2)}$ , the uncentered third moment as  $c_l$ , i.e.  $c_l = \hat{v}_l^{(3)}$ , and the uncentered fourth moment as  $d_l$ , i.e.  $d_l = \hat{v}_l^{(4)}$ , for portfolio  $l$ . Then we can write the mean, log variance, skewness and kurtosis, of portfolio  $l$ , as:

$$h(\hat{v}_{l,M}) = \begin{cases} a_l & M = 1 \\ \ln(b_l - a_l^2) & M = 2 \\ \frac{2a_l^3 + c_l - 3a_l b_l}{(b_l - a_l^2)^{1.5}} & M = 3 \\ \left[ \frac{-3a_l^4 + d_l - 4a_l c_l + 6a_l^2 b_l}{(b_l - a_l^2)^2} - 3 \right] & M = 4 \end{cases} \quad \text{for } l = x, y \quad (81)$$

Note that we use the logarithm of variance instead of just variance, due to the recommendation of Ledoit and Wolf (2011), portfolio  $l = x, y$

The difference between portfolio  $x$  and  $y$  as a function of the uncentered moments is thus given by:

$$f(\hat{v}_M) = \begin{cases} a_x - a_y & M = 1 \\ \ln(b_x - a_x^2) - \ln(b_y - a_y^2) & M = 2 \\ \frac{2a_x^3 + c_x - 3a_x b_x}{(b_x - a_x^2)^{1.5}} - \frac{2a_y^3 + c_y - 3a_y b_y}{(b_y - a_y^2)^{1.5}} & M = 3 \\ \left[ \frac{-3a_x^4 + d_x - 4a_x c_x + 6a_x^2 b_x}{(b_x - a_x^2)^2} - \frac{-3a_y^4 + d_y - 4a_y c_y + 6a_y^2 b_y}{(b_y - a_y^2)^2} \right] & M = 4 \end{cases} \quad (82)$$

The transposed gradient vector of the difference between the mean of portfolio  $x$  and  $y$  is given by:

$$\nabla^T f(\hat{v}_1) = \left( \frac{\partial f(\hat{v}_1)}{\partial a_x}, \frac{\partial f(\hat{v}_1)}{\partial a_y} \right)^T = (1, -1)^T \quad (83)$$

The transposed gradient vector of the difference between the log variance of portfolio  $x$  and  $y$ ,

is given by:

$$\nabla^T f(\hat{v}_2) = \left( \frac{\partial f(\hat{v}_2)}{\partial a_x}, \frac{\partial f(\hat{v}_2)}{\partial a_y}, \frac{\partial f(\hat{v}_2)}{\partial b_x}, \frac{\partial f(\hat{v}_2)}{\partial b_y} \right)^T = \left( \frac{-2a_x}{b_x - a_x^2}, \frac{2a_y}{b_y - a_y^2}, \frac{1}{b_y - a_y^2}, \frac{-1}{b_y - a_y^2} \right)^T \quad (84)$$

The transposed gradient vector of the difference between the skewness of portfolio x and y, is given by:

$$\begin{aligned} \nabla^T f(\hat{v}_3) &= \left( \frac{\partial f(\hat{v}_3)}{\partial a_x}, \frac{\partial f(\hat{v}_3)}{\partial a_y}, \frac{\partial f(\hat{v}_3)}{\partial b_x}, \frac{\partial f(\hat{v}_3)}{\partial b_y}, \frac{\partial f(\hat{v}_3)}{\partial c_x}, \frac{\partial f(\hat{v}_3)}{\partial c_y} \right)^T \\ &= \left( \frac{-3b_x^2 + 3a_x c_x}{(b_x - a_x^2)^{2.5}}, \frac{3b_y^2 - 3a_y c_y}{(b_y - a_y^2)^{2.5}}, \frac{-3c_x + 3a_x b_x}{2(b_x - a_x^2)^{2.5}}, \frac{3c_y - 3a_y b_y}{2(b_y - a_y^2)^{2.5}}, \frac{1}{(b_l - a_l^2)^{1.5}}, \frac{-1}{(b_l - a_l^2)^{1.5}} \right)^T \end{aligned} \quad (85)$$

The transposed gradient vector of the difference between the skewness of portfolio x and y, is given by:

$$\begin{aligned} \nabla^T f(\hat{v}_4) &= \left( \frac{\partial f(\hat{v}_4)}{\partial a_x}, \frac{\partial f(\hat{v}_4)}{\partial a_y}, \frac{\partial f(\hat{v}_4)}{\partial b_x}, \frac{\partial f(\hat{v}_4)}{\partial b_y}, \frac{\partial f(\hat{v}_4)}{\partial c_x}, \frac{\partial f(\hat{v}_4)}{\partial c_y}, \frac{\partial f(\hat{v}_4)}{\partial d_x}, \frac{\partial f(\hat{v}_4)}{\partial d_y} \right)^T \\ &= \left( \frac{12a_x b_x^2 - 12a_x^2 c_x + 4a_x d_x - 4b_x c_x}{(b_x - a_x^2)^3}, \frac{-12a_y b_y^2 + 12a_y^2 c_y - 4a_y d_y + 4b_y c_y}{(b_y - a_y^2)^3}, \right. \\ &\quad \left. \frac{-6a_x^2 b_x + 8a_x c_x - 2d_x}{(b_x - a_x^2)^3}, \frac{6a_y^2 b_y - 8a_y c_y + 2d_y}{(b_y - a_y^2)^3}, \frac{-4a_x}{(b_x - a_x^2)^2}, \frac{4a_y}{(b_y - a_y^2)^2}, \right. \\ &\quad \left. \frac{1}{(b_x - a_x^2)^2}, \frac{-1}{(b_y - a_y^2)^2} \right)^T \end{aligned} \quad (86)$$

The kernel estimate of moment  $M$  is given by:

$$\hat{\Psi}_M = \frac{T}{T - 2M} \sum_{j=-T+1}^{T-1} \kappa \left( \frac{j}{\hat{S}_T} \right) \hat{\Gamma}_T(j) \quad (87)$$

where  $k(\cdot)$  is a kernel function and  $S_T$  the bandwidth of this function, and  $\hat{\Gamma}_T(j)$  is given by:

$$\hat{\Gamma}_T(j) = \begin{cases} \frac{1}{T} \sum_{t=j+1}^T \hat{y}_t \hat{y}_{t-j}^T & \text{for } j \geq 0 \\ \frac{1}{T} \sum_{t=-j+1}^T \hat{y}_{t+j} \hat{y}_t^T & \text{for } j < 0 \end{cases} \quad (88)$$

with the  $Tx2M$ -dimensional vector  $\hat{y}_t^T$  given by

$$\hat{y}_t^T = (R_{x,t} - \hat{v}_x^{(1)}, \dots, R_{x,t}^M - \hat{v}_x^{(M)}, R_{y,t} - \hat{v}_y^{(1)}, \dots, R_{y,t}^M - \hat{v}_y^{(M)}) \quad (89)$$

where  $R_{l,t}$  is the return of portfolio  $l$  at time  $t$ , for  $l = x, y$ . There are different ways of estimating the kernel density  $\kappa(\cdot)$ , and as we follow the procedure of Ledoit-Wolf (var-paper), we use the formula of the Parzen kernel estimate from Andrew (1991), given by:

$$\kappa(x) = \begin{cases} 1 - 6x^2 + 6|x|^3 & \text{for } 0 \leq |x| \leq 0.5 \\ 2(1 - |x|)^3 & \text{for } 0.5 < |x| \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (90)$$

The bandwidth of the Parzen kernel estimate  $S_T$  is given by:

$$\hat{S} = 2.6614 (\hat{\alpha}(2)T)^{1/5} \quad (91)$$

where  $\hat{\alpha}(2)$  is given by:

$$\hat{\alpha}(2) = \left( \sum_{i=1}^{2M} \frac{4\hat{\rho}_i^2 \hat{\sigma}_i^4}{(1 - \hat{\rho}_i)^8} \right) / \left( \sum_{i=1}^{2M} \frac{\hat{\sigma}_i^4}{(1 - \hat{\rho}_i)^4} \right) \quad (92)$$

where  $\hat{\rho}_i$  and  $\hat{\sigma}_i^2$  denote the autoregressive parameter and innovation variance parameter, respectively, for  $i = 1, 2, \dots, 2M$ , as a result of running OLS with autoregressive errors, using an AR(1)-model, on each vector  $\hat{y}_t^T$ . We obtain these parameters using the R-function.

## D Description of stocks in empirical data set

### D.1 Descriptive statistics of the empirical data set

The composition of stocks for S&P Global 1200 (1994) was found from the Bloomberg Terminal, whereas we collected the sample data from Datastream. The descriptive statistics of the stocks are given below.

**Table 12:** Descriptive statistics for cyclical stocks

Bloomberg Ticker	Datastream Ticker	Stock Name	Annual return	SD	Skewness	Kurtosis	JB*
6201 JT	932034	Toyota Industries	0.04	0.30	0.35	6.9	33
NYT US	912377	The New York Times Company	0.01	0.36	0.53	13.5	55
NC UN	741924	NACCO Industries Inc.	0.08	0.44	0.44	10.5	76
CAL UN	921334	Caleres Inc	0.04	0.48	0.49	13.7	160
NKE UN	993249	NIKE Inc	0.14	0.32	0.17	11.5	234
NXT LN	901203	NEXT PLC	0.13	0.32	-0.21	14.1	80
23 HK	951410	The Bank of East Asia, Ltd.	0.03	0.32	0.46	13.6	61
BBVA SM	779090	Banco Bilbao Vizcaya Argentaria, S.A.	0.05	0.35	0.38	10.6	22
DLX UN	916704	Deluxe Corp.	0.06	0.35	0.38	18.8	129
TKR UN	903720	The Timken Company	0.05	0.37	-0.03	14.5	134
VOLVB SS	944486	Volvo AB	0.06	0.37	0.18	7.8	31
6702 JT	912718	Fujitsu Limited	-0.03	0.38	0.31	6.9	27
HRS UN	905409	Harris Corporation	0.12	0.33	0.41	11.5	216
5711 JT	930960	Mitsubishi Materials Corporation	-0.02	0.42	0.49	7.6	31
CSR AT	904735	CSR Limited	0.03	0.34	-0.15	8.4	13
JMAT LN	901152	Johnson Matthey PLC	0.07	0.32	0.24	7.9	26
ASH UN	905342	Ashland Global Holdings Inc.	0.06	0.33	-0.43	16.9	52

\* Median of annual Jarque-Bera statistic

**Table 13:** Descriptive statistics for defensive stocks


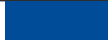





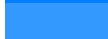

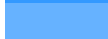

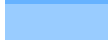

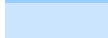






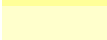



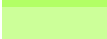
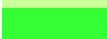

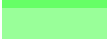
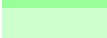
Bloomberg Ticker	Datastream Ticker	Stock Name	Annual Return	SD	Skewness	Kurtosis	JB*
4911 JT	905323	Shiseido Company	0.03	0.30	0.23	7.2	33
CA FP	922029	Carrefour SA	0.01	0.32	0.13	6.1	21
KO UN	904282	Coca-Cola Company	0.05	0.22	0.20	10.2	55
KR UN	912134	The Kroger Co.	0.11	0.29	-0.50	13.2	113
HSY UN	905077	The Hershey Company	0.10	0.23	1.25	27.1	105
RB/ LN	900484	Reckitt Benckiser Group PLC	0.10	0.27	0.21	12.9	52
AD NA	779426	Koninklijke Ahold Delhaize N.V.,	0.07	0.29	-0.08	9.1	82
BHI UN	921431	Baker Hughes Inc	0.06	0.41	0.17	9.5	30
SAN FP	992594	Sanofi	0.09	0.30	0.16	6.9	31
NOVN VX	992594	Novartis AG	0.08	0.21	0.75	15.6	40
TIT IM	923374	Telecom Italia S.p.A., Centrais Eletricas	-0.05	0.40	0.11	8.6	37
ELET6 BS	501893	Brasileiras S.A. (Eletrobras)	-0.03	0.54	0.57	10.1	40

\* Median of annual Jarque-Bera statistic

## D.2 Encoded colors for defensive and cyclical stocks

The table presents the encoded colors for the forecasts of statistical moments (5.1) and weight plots for the basic portfolio (5.2) and for the constrained portfolios in appendix (A.X)

**Table 14:** Encoded colors for stocks in empirical data set

<b>Cyclical</b>			<b>Defensive</b>		
Bloomberg	GICS	Color	Bloomberg	GICS	Color
6201 JT	Consumer Discretionary		4911 JT	Consumer Staples	
NYT US	Consumer Discretionary		CA FP	Consumer Staples	
NC UN	Consumer Discretionary		KO UN	Consumer Staples	
CAL UN	Consumer Discretionary		KR UN	Consumer Staples	
NKE UN	Consumer Discretionary		HSY UN	Consumer Staples	
NXT LN	Consumer Discretionary		RB/ LN	Consumer Staples	
23 HK	Financials		AD NA	Consumer Staples	
BBVA SM	Financials		BHI UN	Energy	
DLX UN	Industrials		SAN FP	Health Care	
TKR UN	Industrials		NOVN VX	Health Care	
VOLVB SS	Industrials		TIT IM	Telecommunication	
6702 JT	Information Technology		ELET6 BS	Utilities	
HRS UN	Information Technology				
5711 JT	Materials				
CSR AT	Materials				
JMAT LN	Materials				
ASH UN	Materials				

# E Model

## E.1 Shrinkage intensities

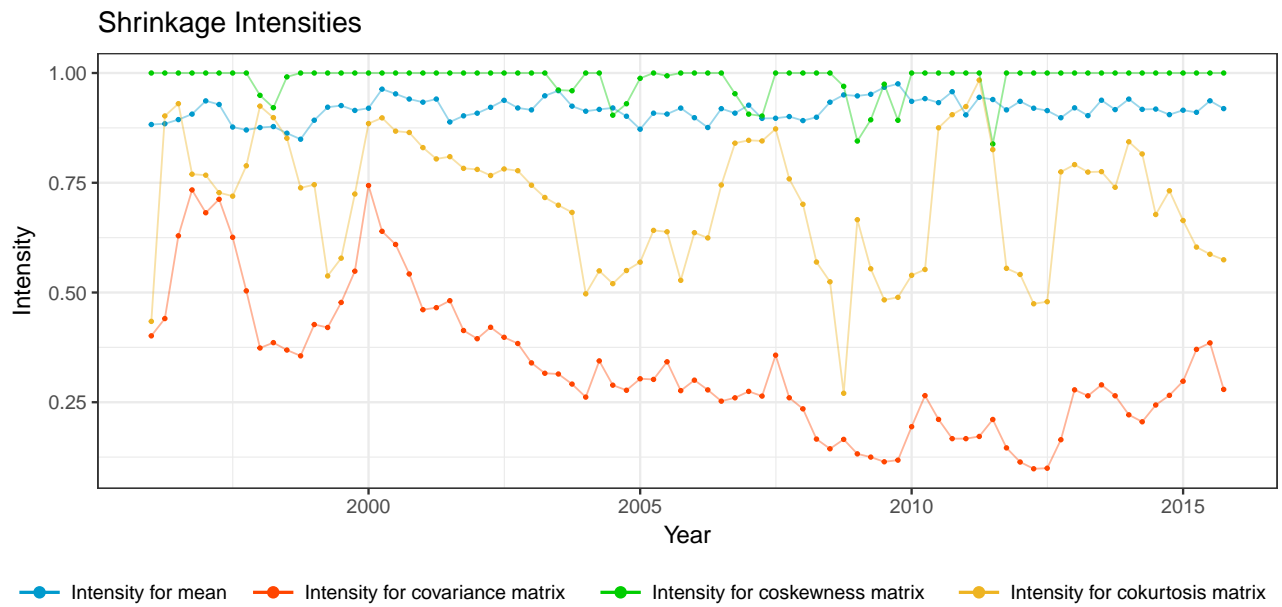


Figure 16: Shrinkage intensities

## E.2 Portfolio weight plots

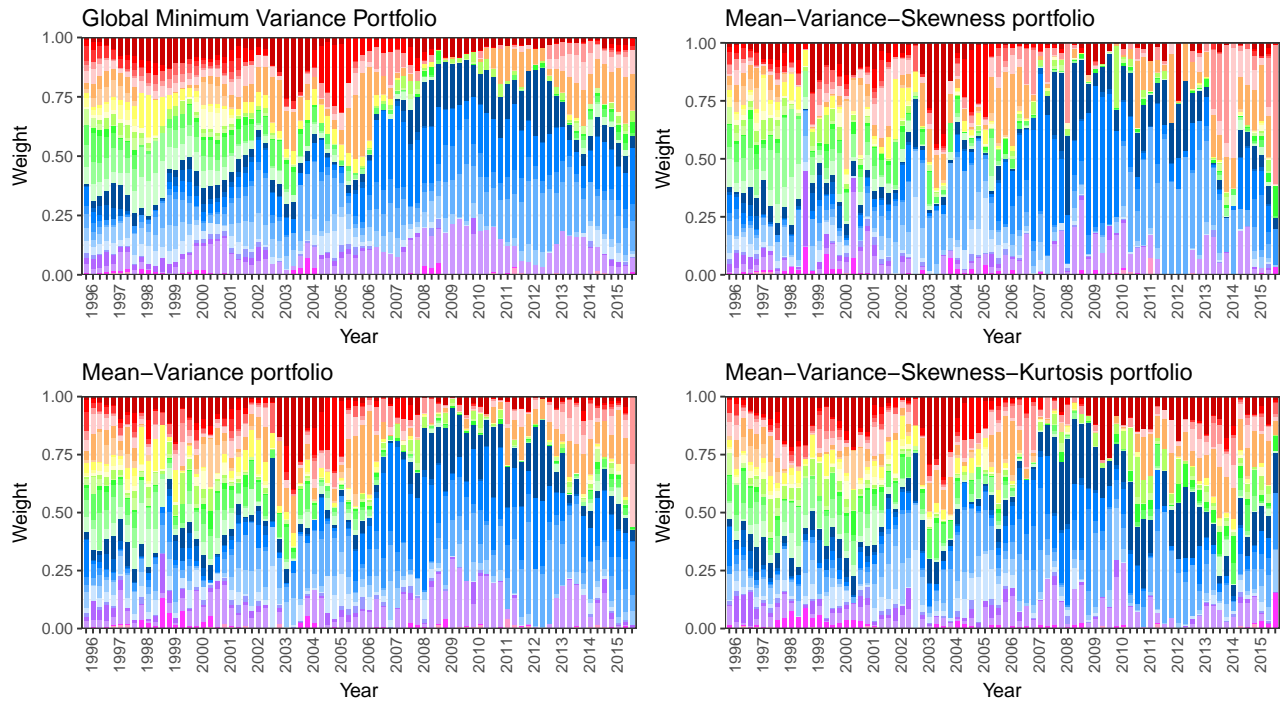


Figure 17: Quarterly allocation of portfolios with mild div. constraint

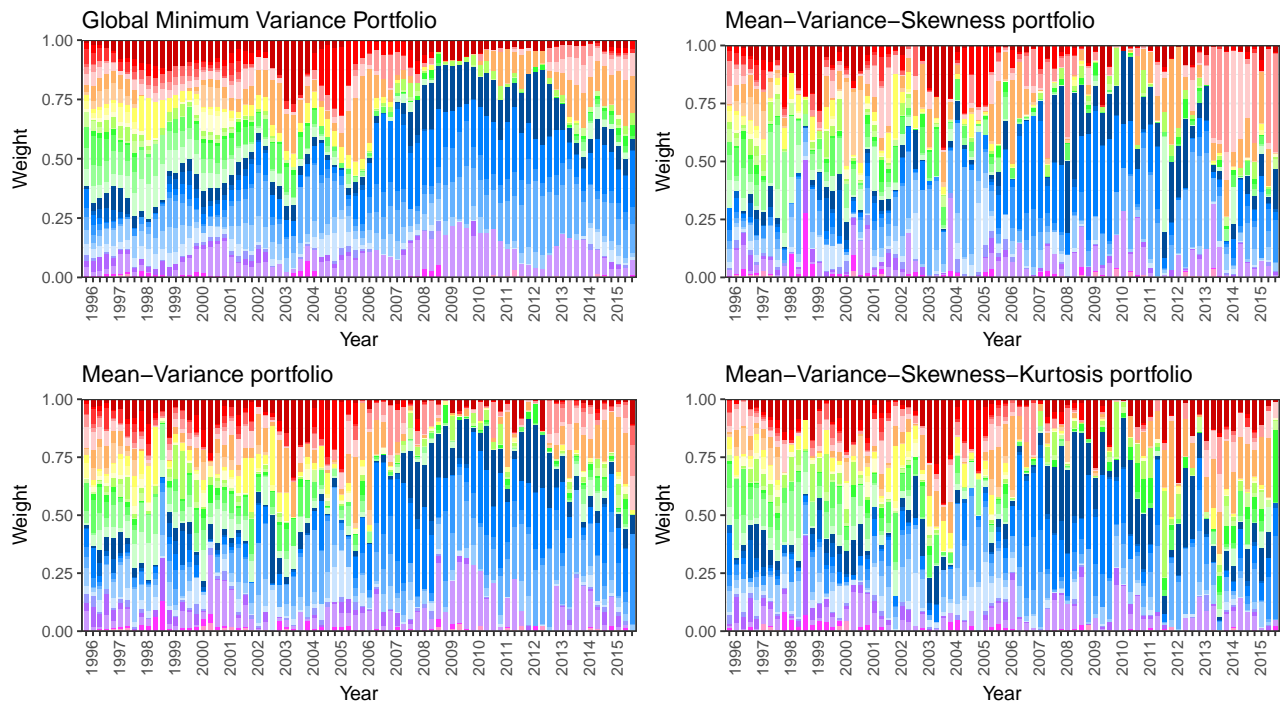


Figure 18: Quarterly allocation of portfolios with strong div. constraint



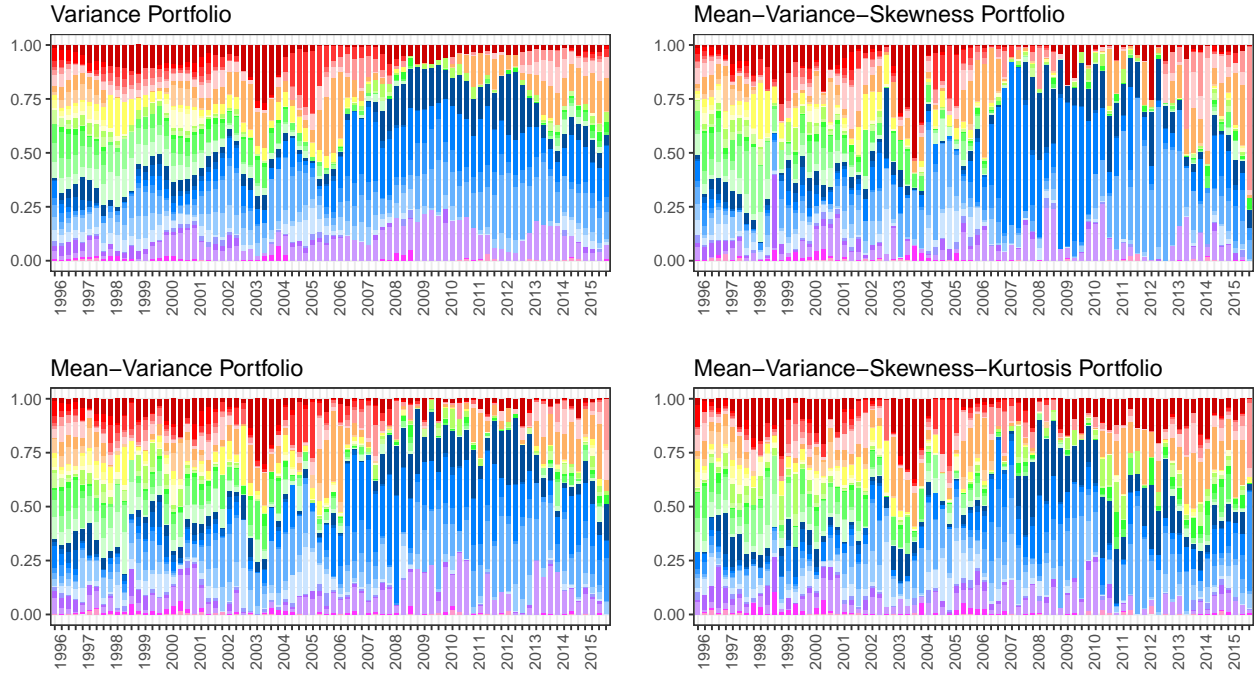


Figure 19: Quarterly allocation of portfolios with mild turn. constraint

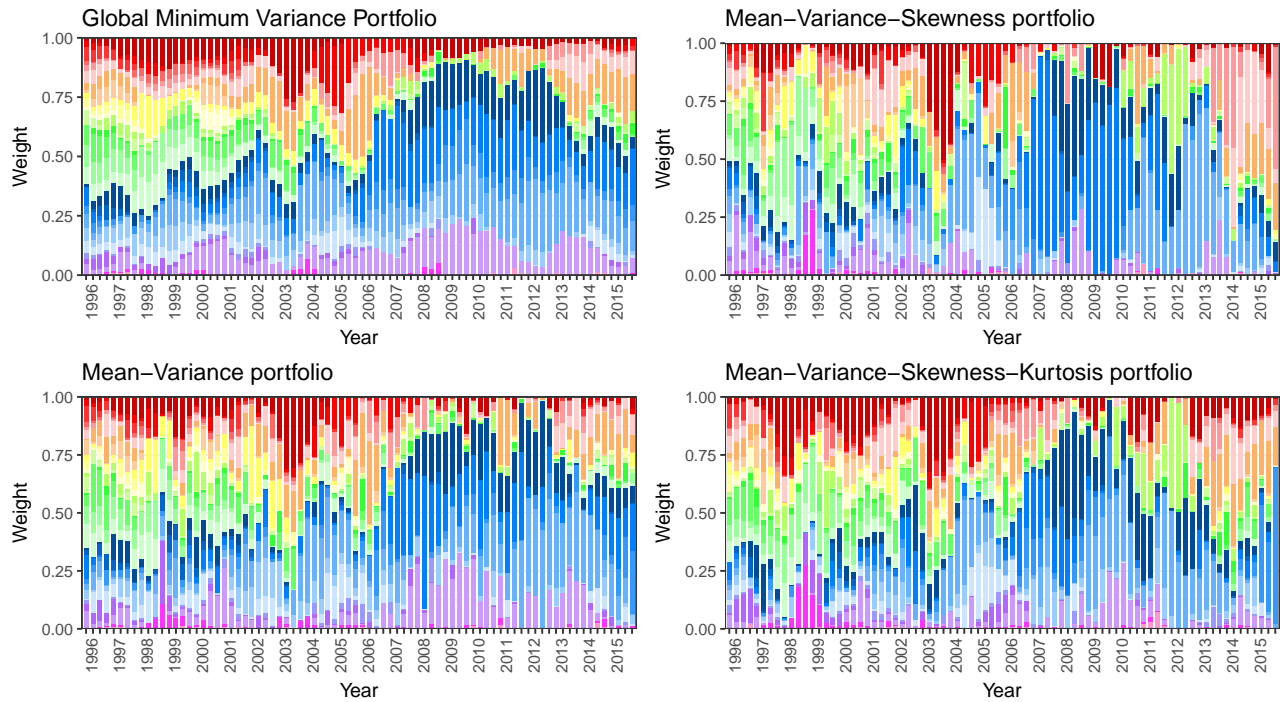


Figure 20: Quarterly allocation of portfolios with strong turn. constraint

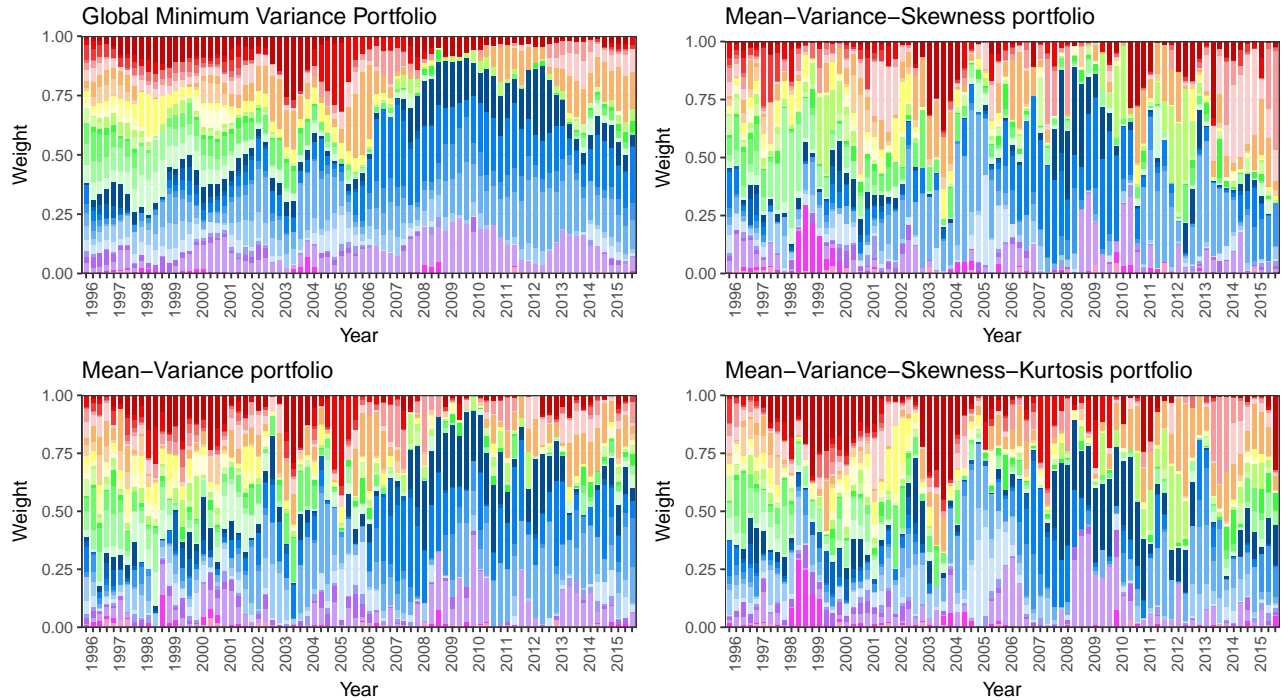


Figure 21: Quarterly allocation of portfolios with strong div.- and turn. constraint

### E.3 Quarterly in-sample portfolio mean, variance skewness and kurtosis

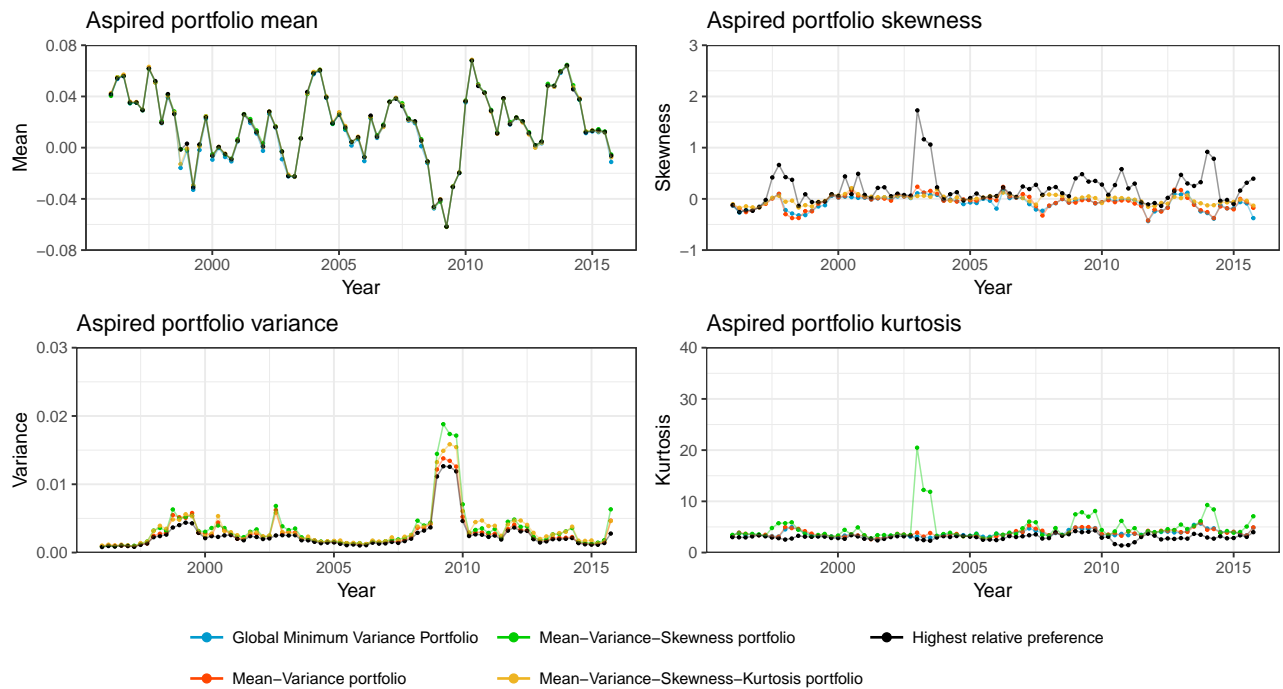


Figure 22: Quarterly aspired portfolio moments given mild div. constraint

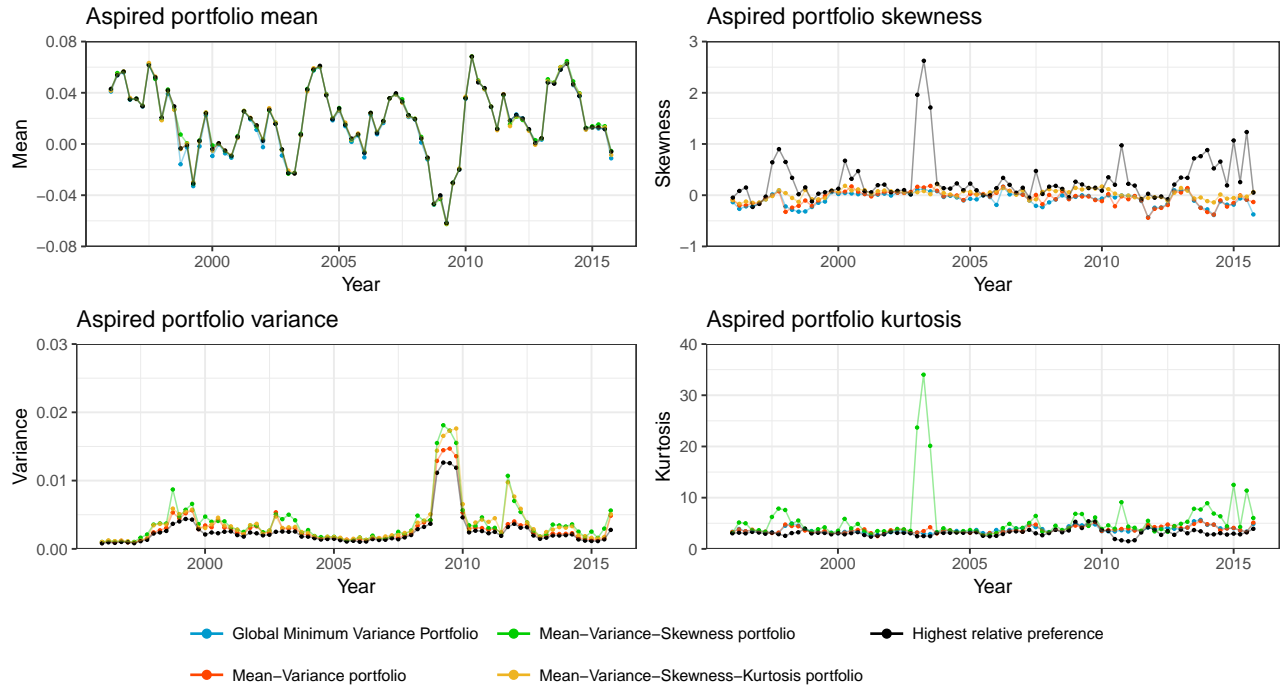


Figure 23: Quarterly aspired portfolio moments given strong div. constraint

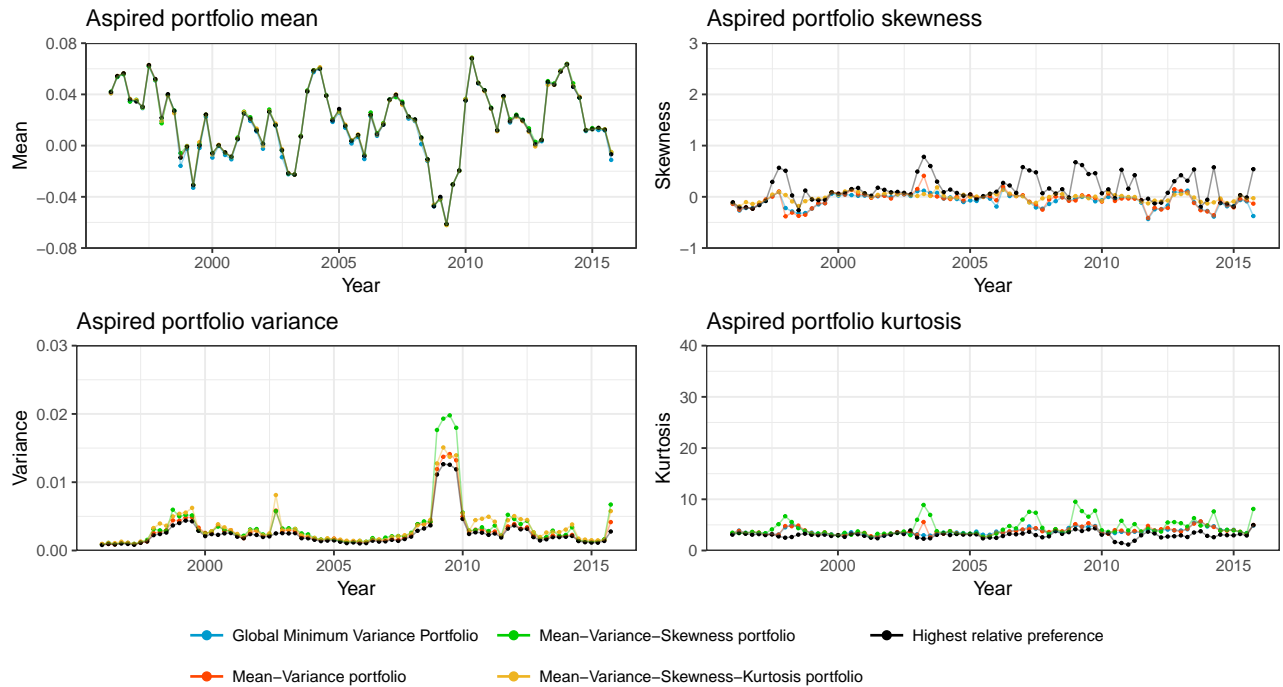


Figure 24: Quarterly aspired portfolio moments given mild turn. constraint

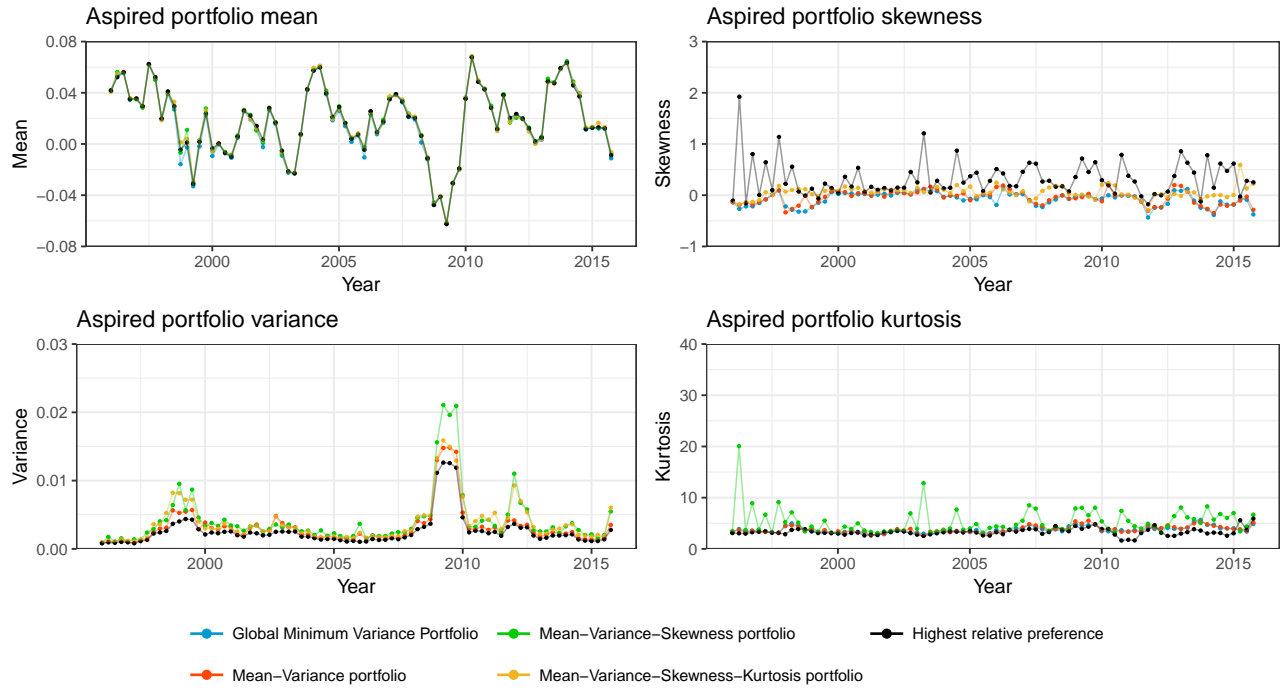


Figure 25: Quarterly aspired portfolio moments given strong turn. constraint

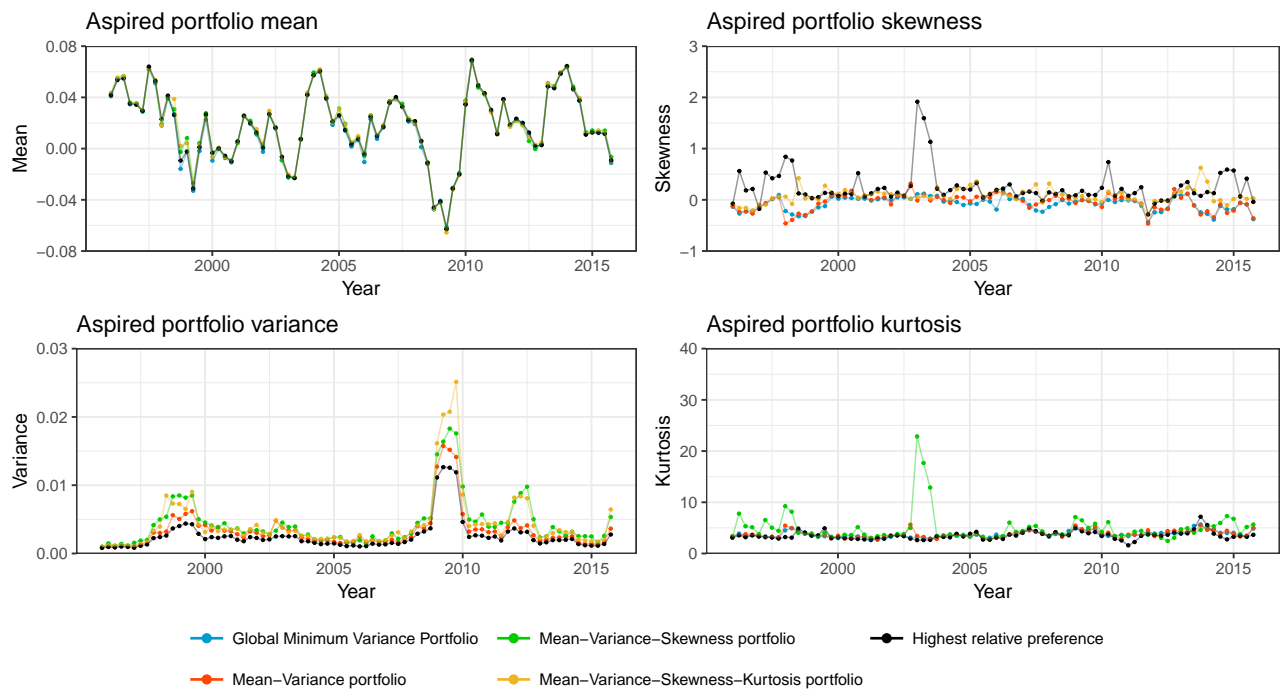


Figure 26: Quarterly aspired portfolio moments given strong div.- and turn. constraint

## F Cumulative returns for benchmark portfolios

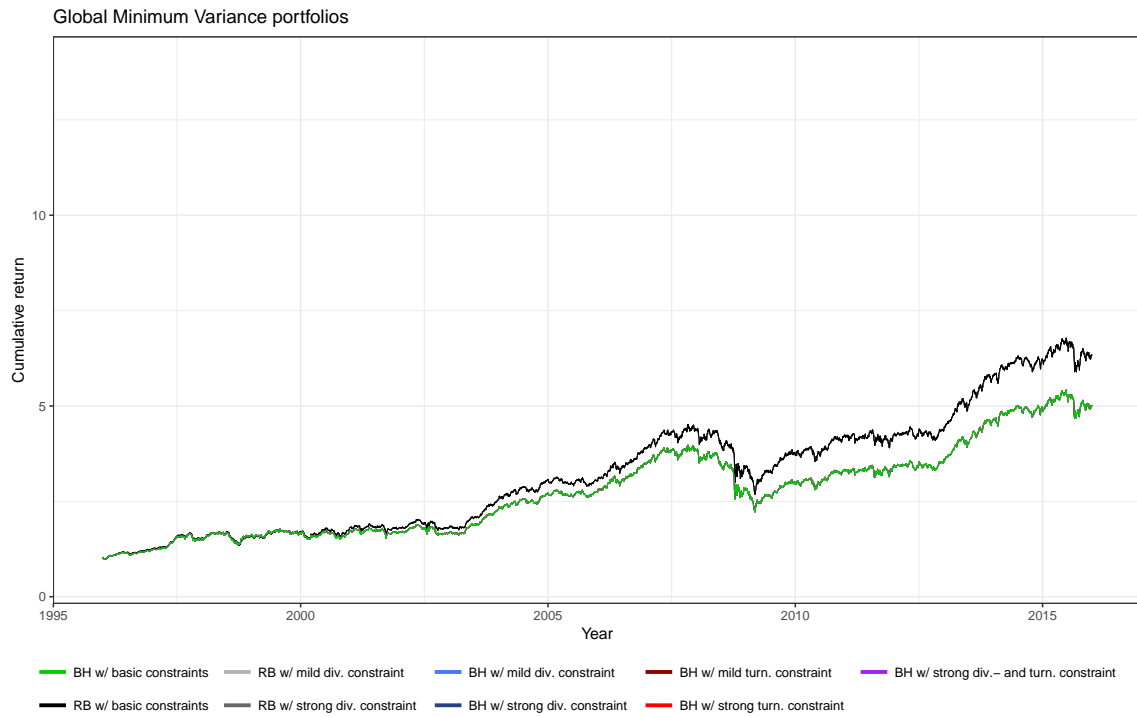


Figure 27: Wealth of global minimum variance portfolios

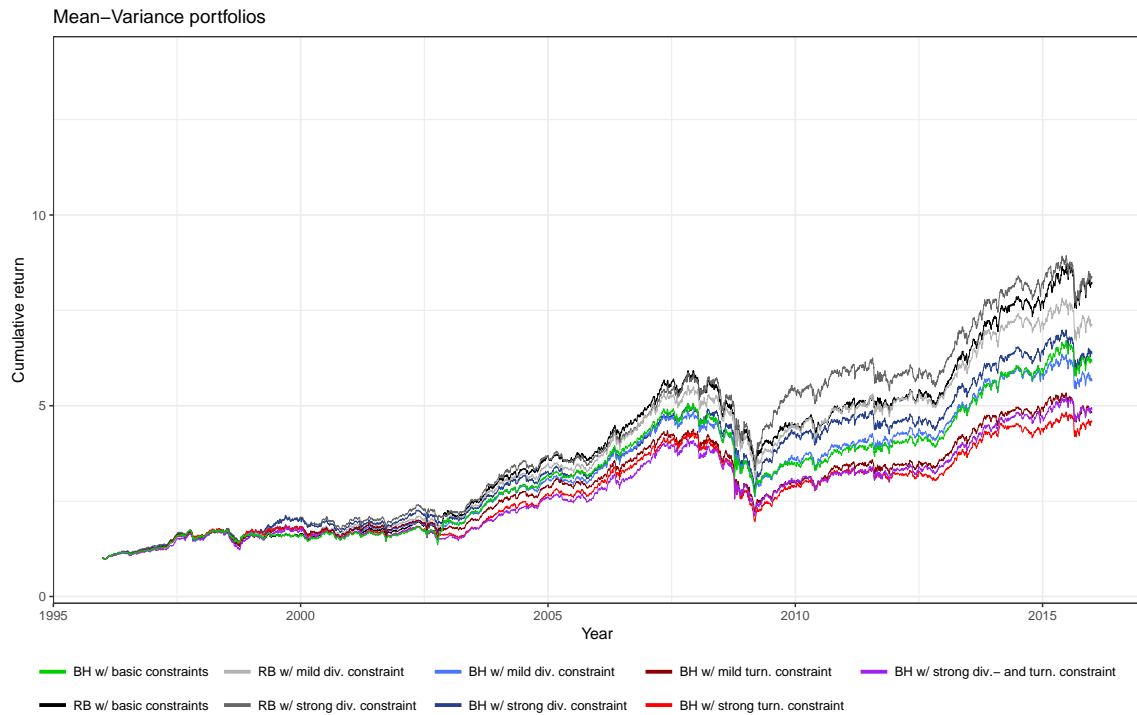


Figure 28: Wealth of mean-variance portfolios

## G R-code Appendix

Listing 1: Load packages

---

```
1 # Load packages
2 library(sqldf); library(zoo); library(xts); library(ROI); library(Rmosek)
3 library(RiskPortfolios); library(optimbase); library(ggplot2); library(PortfolioAnalytics)
4 library(DEoptim); library(boot); library(tseries); library(PerformanceAnalytics)
```

---

Listing 2: Forecast of stock moments by shrinkage

---

```
6 rolling_shrinkage ←function(R, R2) {
7   # Storing returns as matrix
8   r ← as.matrix(R)
9
10  # Sample estimate of covariance, coskewness and cokurtosis
11  M2 ← covp(R2); M3 ←M3.MM(R2); M4 ←M4.MM(R2)
12
13  # Regression of market on each return series to obtain betas and residuals
14  b ← sapply(1:N, function(i) lm(R[,i] ~ M)$coefficients[2])
15  et ← sapply(1:N, function(i) summary(lm(R[,i]~M))$residuals)
16
17  # Market mean & centered market moments of n-degree
18  mrktmean ←mean(M); mu_02 ←mrktmom(2); mu_03 ←mrktmom(3); mu_04 ←mrktmom(4)
19
20  # Idiosyncratic var(e2), skewness(e3) and kurtosis(e4); stored in a matrix
21  e2 ← e(2)*diag(x = 1, nrow = N, ncol = N)
22  e3 ← matrix(0, nrow = N, ncol = N^2)
23  for(i in 1:N) {
24    e3[i, (1 + (N+1)*(i - 1))] ←e(3)[i]
25  }
26
27  # Identifying elements of cokurtosis matrix
28  cok_p ←function(N){
29    x ← matrix(nrow = N, ncol = 1)
30    for(i in 1:N) { #When Q != K
31      for(j in 1:N) {
32        m ← matrix(5, nrow = N, ncol = N)
33        m[,j] ←4; m[j,] ←4 ; diag(m) ←4; m[,i] ←4; m[i,] ←4
34        m[j,i] ←3; m[i,j] ←3
35        m[j,j] ←2; m[i,i] ←2
36        x ← cbind(x,m)
37      }
38    }
39    x ← x[,2:(N^3 + 1)]
40
41    for(i in 1:N) { #When Q != K
42      m ← matrix(4, nrow = N, ncol = N)
43      diag(m) ←3
44      m[,i] ←2; m[i,] ←2
45      m[i,i] ←1
46      x[, (1+N*((N+1)*(i-1))):(N+N*((N+1)*(i-1)))] ←m
47    }

```

```

48   return(x)
49 }
50
51 e4 ← cok_p(N)
52
53 for(k in 1:N^3) {
54   for(l in 1:N) {
55     if(e4[l,k] == 1) { #viiii
56       e4[l,k] ← e(4)[l]
57     } else if(e4[l,k] == 2) { #viiij
58       e4[l,k] ← 3*b[I[k]]*b[J[k]]*mu_02*e(2)[I[k]]
59     } else if(e4[l,k] == 3) { #viiijj
60       e4[l,k] ← ((b[I[k]]^2)*mu_02*e(2)[J[k]] + (b[J[k]]^2)*mu_02*e(2)[I[k]] +
61                 (e(2)[J[k]])*(e(2)[I[k]]))
62     } else if(e4[l,k] == 4) { #viiijk
63       e4[l,k] ← b[J[k]]*b[K[k]]*mu_02*e(2)[I[k]]
64     } else if(e4[l,k] == 5) { #vijjkl
65       e4[l,k] ← 0
66     }
67   }
68 }
69
70 # Single-Index covariance-, coskewness- and cokurtosis matrix
71 cov ← (b%*%t(b))*mu_02 + e2
72 coskew ← kronecker(b%*%t(b),t(b))*mu_03 + e3
73 cokurt ← kronecker(kronecker(b%*%t(b),t(b)),t(b))*mu_04 + e4
74
75 ## Bayesian Shrinkage
76
77 # Vector containing the covariance of assets and market
78 cm ← sapply(1:N, mrktcov, f = R)
79
80 # Shrinkage of mean (Bayes-Stein, Formula: 16)
81 sM1 ← meanEstimation(r, control = list(type = "bs"))
82
83 # Shrinkage of the covariance matrix(M2)
84
85 # Asymptotic covariance between single-index and sample estimate
86 covpie ← function(f, R, covar) {
87   pie ← matrix(ncol = N, nrow = N)
88   for(i in 1:N) {
89     for(j in 1:N) {
90       pie[j,i] ← sum((f(R,i)*f(R,j) - covar[j,i])^2)/T
91     }
92   }
93   return(pie)
94 }
95
96 p2 ← covpie(f = ex_ri, R = R, covar = M2)
97 pie2 ← sum(p2)
98
99 # Asymptotic variance of single-index covariance matrix (rho)
100 r_diag ← sum(diag(p2))
101

```

```

102 #Asymptotic covariance for covariance by single-index
103 AsyCov_V ←function(f, R, n = 0, x, y) {
104   o ← (f(R,x)*f(R,y) - M2[x,y])
105   if(n == 0) {
106     return(sum((f(R,x)*(M-mean(M)) - mrktcov(R,x))*o)/T)
107   } else {
108     return(sum(((M - mean(M))^n - mrktmom(n))*o)/T)
109   }
110 }
111
112 # Asymptotic covariance of single-index covariance matrix (rho)
113 l1 ← 0; l2 ← 0
114 for(i in 1:N) {
115   for(j in 1:N) {
116     if(i != j) {
117       ij ← list(f=ex_ri, R=R2, x=i, y=j)
118       l1 ← (l1 + do.call(AsyCov_V, ij)*cm[j]/mu_02)
119       l2 ← (l2 + do.call(AsyCov_V, c(n=2, ij))*cm[i]*cm[j]/(mu_02^2))
120     }
121   }
122 }
123
124 rho2 ← r_diag + (2*l1 - l2)
125 # Asymptotic covariance of single-index- and sample covariance matrix (gamma)
126 gamma2 ← sum((cov - M2)^2)
127
128 # Shrinkage Intensity
129 k2 ← (pie2-rho2)/gamma2
130 a2 ← max(0, min(1, k2/T))
131
132 # Shrinkage estimate of covariance matrix
133 sM2 ← a2*cov + (1 - a2)*M2
134
135 # Shrinkage of the coskewness matrix
136 # Asymptotic variance of sample coskewness
137
138 #Asymptotic covariance for coskewness matrix by single-index estimate
139 cospie ←function(f, N, R, coskew) {
140   p ← 0; pie ←matrix(nrow = N, ncol = N)
141   for(i in 1:N) {
142     for(j in 1:N) {
143       for(k in 1:N) {
144         pie[j, k] ←sum((f(R,i)*f(R,j)*f(R,k) - coskew[i,(k + (j - 1)*N)])^2)/T
145       }
146     }
147   }
148   p ← cbind(p, pie)
149   return(p[,2:(N^2 + 1)])
150 }
151
152 p3 ← cospie(f = ex_ri, N = N, R = R, coskew = M3)
153 pie3 ←sum(p3)
154
155 # Asymptotic variance of single-index coskewness

```



```

156 r_diag ← sum(sapply(1:N, function(i) p3[i,(1 + (N + 1)*(i - 1))]))
157
158 #Asymptotic covariance for skewness by single-index
159 AsyCov_S ← function(f, R, n = 0, x, y, z) {
160   o ← (f(R,x)*f(R,y)*f(R,z) - M3[x,(z + (y - 1)*N)])
161   if(n == 0) {
162     return(sum((f(R,x)*(M-mean(M)) - mrktcov(R,x))*o)/T)
163   } else {
164     return(sum(((M - mean(M))^n - mrktmom(n))*o)/T)
165   }
166 }
167
168 # Asymptotic covariance of single-index coskewness
169 l1 ← 0; l2 ← 0; l3 ← 0
170 for(i in 1:N) {
171   for(j in 1:N) {
172     for(k in 1:N) {
173       if(i != j | i != k) { #for ijk
174         ijk ← list(f=ex_ri,R=R,x=i,y=j,z=k)
175         l1 ← l1 + do.call(AsyCov_S, ijk)*cm[j]*cm[k]*mu_03/(mu_02^3)
176         l2 ← l2 + do.call(AsyCov_S, c(ijk, n=2))*cm[i]*cm[j]*cm[k]*mu_03/(mu_02^4)
177         l3 ← l3 + do.call(AsyCov_S, c(ijk, n=3))*cm[i]*cm[j]*cm[k]/(mu_02^3)
178       }
179     }
180   }
181 }
182
183 rho3 ← r_diag + (3*l1 - 3*l2 + l3)
184
185 # Asymptotic covariance of sample- and single-index coskewness (gamma)
186 gamma3 ← sum((coskew - M3)^2)
187
188 # Shrinkage Intensity
189 k3 ← (pie3 - rho3)/gamma3
190 a3 ← max(0, min(1, k3/T))
191
192 # Shrinkage estimate of coskewness matrix
193 sM3 ← a3*coskew + (1 - a3)*M3
194
195 # Shrinkage of the cokurtosis matrix (M4)
196 # Asymptotic variance of sample cokurtosis
197 p4 ← cokpie(f = ex_ri, N = N, R = R, cokurt = M4)
198 pie4 ← sum(p4)
199
200 # Asymptotic variance of single-index cokurtosis
201 r_diag ← sum(sapply(1:N, function(i) p4[i,(1 + (N^2 + N + 1)*(i - 1))]))
202
203 cokpie ← function(f, N, R, cokurt) {
204   p ← 0; pie ← matrix(nrow = N, ncol = N)
205   for(i in 1:N) {
206     for(j in 1:N) {
207       for(k in 1:N) {
208         for(l in 1:N) {
209           pie[k, l] ← sum((f(R, i)*f(R, j)*f(R, k)*f(R, l) -

```

```

210         cokurt[i,(1 + (k - 1)*N + (j - 1)*(N^2)))]^2)/T
211     }
212 }
213 p ← cbind(p, pie)
214 }
215 }
216 return(p[, 2:(N^3 + 1)])
217 }
218
219 #Asymptotic covariance for kurtosis by single-index
220 AsyCov_K ←function(f, R=R, n = 0, x, y, z,v) {
221   o ← (f(R,x)*f(R,y)*f(R,z)*f(R,v) - M4[x,(v + (z - 1)*N + (y - 1)*(N^2))])
222   if(n == 0) {
223     return(sum((f(R,x)*(M-mean(M)) - mrktcov(R,x))*o)/T)
224   } else if(n == "e") {
225     return(sum((et[,x]^2 - e(2)[x])*o)/T)
226   } else {
227     return(sum(((M - mean(M))^n - mrktmom(n))*o)/T)
228   }
229 }
230
231 # Asymptotic covariance of single-index cokurtosis
232 l1 ← 0; l2 ← 0; l3 ← 0; l4 ← 0; l5 ← 0; l6 ← 0
233 for(i in 1:N) {
234   for(j in 1:N) {
235     if(i != j) { #for riiij & riiij
236       iijj ← list(f=ex_ri,R=R,y=i,z=i); iijj ← list(f=ex_ri,R=R,y=i,v=j)
237       l4 ← (l4 + 3*((cm[j]*e(2)[i]/mu_02)*do.call(AsyCov_K, c(iijj,x=i,v=j)) +
238             (cm[i]*e(2)[i]/mu_02)*do.call(AsyCov_K, c(iijj,x=j,v=i)) -
239             (cm[i]*cm[j]*e(2)[i]/(mu_02^2))*do.call(AsyCov_K, c(iijj,n=2,x=i,v=j)) +
240             (cm[i]*cm[j]/mu_02)*do.call(AsyCov_K, c(iijj,n="e",x=i,v=j))))
241       l5 ← (l5 + (2*cm[i]*e(2)[j]/mu_02)*do.call(AsyCov_K, c(iijj,x=i,z=j)) -
242             ((cm[i]^2)*e(2)[j]/(mu_02^2))*do.call(AsyCov_K, c(iijj,n=2,x=i,z=j)) +
243             ((cm[i]^2)/mu_02)*do.call(AsyCov_K, c(iijj,n="e",x=j,z=i)) +
244             2*((cm[j]^2)*e(2)[i]/mu_02)*do.call(AsyCov_K, c(iijj,x=j,z=i)) -
245             ((cm[j]^2)*e(2)[i]/(mu_02^2))*do.call(AsyCov_K, c(iijj,n=2,x=i,z=j)) +
246             ((cm[j]^2)/mu_02)*do.call(AsyCov_K, c(iijj,n="e",x=i,z=j)) +
247             e(2)[j]*do.call(AsyCov_K, c(iijj,n="e",x=i,z=j)) +
248             e(2)[i]*do.call(AsyCov_K, c(iijj,n="e",x=j,z=i)))
249     }
250   for(k in 1:N) {
251     if(i != j & i != k & j != k) { #for riijk
252       iijk ← list(f=ex_ri,R=R,y=i)
253       l6 ← (l6 + (cm[k]*e(2)[i]/mu_02)*do.call(AsyCov_K, c(iijk,x=j,z=i,v=k)) +
254             (cm[j]*e(2)[i]/mu_02)*do.call(AsyCov_K, c(iijk,x=k,z=i,v=j)) -
255             (cm[j]*cm[k]*e(2)[i]/(mu_02^2))*do.call(AsyCov_K, c(iijk,n=2,x=i,z=j,v=k)) +
256             (cm[j]*cm[k]/mu_02)*do.call(AsyCov_K, c(iijk,x=i,z=j,v=k)))
257     }
258   for(l in 1:N) {
259     if(i != j | i != k | i != l) { #for rijkl
260       ijkl ← list(f=ex_ri,R=R,x=i,y=j,z=k,v=l)
261       l1 ← (l1 + do.call(AsyCov_K, ijkl)*cm[j]*cm[k]*cm[l]*
262             mu_04/(mu_02^4))
263       l2 ← (l2 + do.call(AsyCov_K, c(ijkl,n=2))*cm[i]*cm[j]*

```

```

264         cm[k]*cm[l]*mu_04/(mu_02^5))
265     l3 ← (l3 + do.call(AsyCov_K, c(ijkl,n=4))*cm[i]*cm[j]*
266         cm[k]*cm[l]/(mu_02^4))
267     }
268   }
269 }
270 }
271 }
272
273 r_spes ← 4*14 + 3*15 + 6*16
274 rho4 ← r_diag + (4*11 - 4*12 + l3) + r_spes
275
276 # Asymptotic covariance of single-index- and sample cokurtosis matrix (gamma)
277 gamma4 ← sum((cokurt - M4)^2)
278
279 # Shrinkage Intensity
280 k4 ← (pie4 - rho4)/gamma4
281 a4 ← max(0, min(1, k4/T))
282
283 # Shrinkage estimate of cokurtosis matrix
284 sM4 ← a4*cokurt + (1 - a4)*M4
285
286 shrink_list ← list(sM1, sM2, sM3, sM4)
287 return(shrink_list)
288 }
289 x ← rep(0, length(m))

```

---

### Listing 3: Apply portfolio constraints

---

```

290 # Apply (1) / Exclude (0) turnover constraint
291 constraint_t ← 1
292 # Specify the turnover constraint, reset constraint if excluded
293 turnoverrate ← 0.05; if(constraint_t==0){turnoverrate ← 0}
294
295 # Apply (1) / Exclude (0) diversification constraint
296 constraint_d ← 1
297 # Specify the diversification constraint, reset constraint if excluded
298 divrate = 0.6; if(constraint_d==0){divrate ← 0}

```

---

### Listing 4: Maximization of portfolio mean (Formula 50)

---

```

300 optmean ← function(m, N, div = 0, tr = Inf, oldw = NULL) {
301   cqo1 ← list(sense = "max"); cqo1$c ← m
302   if(div == 0) {
303     # If no diversification- and turnover constraint
304     if(tr == Inf) {
305       # Return maximum mean
306       x ← rep(0, length(m))
307       for(i in 1:length(m)) {
308         if(m[i]==max(m)) {
309           x[i] ← 1
310           break
311         } else {
312           x[i] ← 0

```

```

313   }
314 }
315 return(x)
316 } else {
317   m1 ← c(as.numeric(m), rep(0,2*N)); cqo1$c ←c(m,rep(0,(2*N)))
318   #Full-investment and turnover constraint
319   cqo1$A ←Matrix(rbind(rbind(c(rep(1,N),rep(0,2*N)),
320                         c(rep(0,N),rep(1,2*N))), cbind(diag(1,N),
321                         diag(-1,N),diag(1,N))), sparse = TRUE)
322   cqo1$bc ←rbind(blc = c(1,-Inf, oldw), buc = c(1,tr, oldw))
323   #Inequality constraint: long-only
324   cqo1$bx ←rbind(blx = rep(0,3*N), bux = c(rep(Inf,(3*N))))
325   #Solve optimization procedure by interior-point method
326   return(mosek(cqo1, opts = list(verbose = 1))$sol$itr$xx[1:N])
327 }
328 #
329 } else {
330   # If diversification constraint, but no turnover constraint
331   if(tr == Inf) {
332     #Equality constraint: Full-investment
333     cqo1$A ←Matrix(c(rep(1,N), rep(0,N)), nrow=2, byrow=TRUE, sparse=TRUE)
334     oprc ←sapply(1:N, function(j) list("POW", 2, j, 1, 2, 0))
335     rownames(oprc) ←c("type","i","j","f","g","h")
336     cqo1$scopt ←list(oprc=oprc)
337     cqo1$bc ←rbind(blc = c(1,-Inf), buc = c(1,(1-div)))
338     #Box constraints: long-only
339     cqo1$bx ←rbind(blx = rep(0,N), bux = rep(1,N))
340     #Solve optimization procedure by interior-point method
341     return(mosek(cqo1, opts = list(verbose = 1))$sol$itr$xx[1:N])
342     # If diversification- and turnover constraint
343   } else {
344     m1 ← c(as.numeric(m), rep(0,2*N)); cqo1$c ←c(m,rep(0,(2*N)))
345     #Full-investment, diversification-, and turnover constraint
346     cqo1$A ←Matrix(rbind(rbind(c(rep(1,N),rep(0,2*N)),
347                               c(rep(0,N),rep(1,2*N))), cbind(diag(1,N),
348                               diag(-1,N),diag(1,N)), rep(0,3*N)), sparse = TRUE)
349     oprc ←sapply(1:N, function(j) list("POW", (N+3), j, 1, 2, 0))
350     rownames(oprc) ←c("type","i","j","f","g","h")
351     cqo1$scopt ←list(oprc=oprc)
352     cqo1$bc ←rbind(blc = c(1,-Inf, oldw,-Inf), buc = c(1,tr, oldw,(1-div)))
353     #Inequality constraint:
354     cqo1$bx ←rbind(blx = rep(0,3*N), bux = c(rep(Inf,(3*N))))
355     #Solve optimization procedure by interior-point method
356     return(mosek(cqo1, opts = list(verbose = 1))$sol$itr$xx[1:N])
357   }
358 }
359 }

```

---

**Listing 5:** Minimization of portfolio variance (Formula 50)

---

```

360 optvar ←function(Q, N, div = 0, tr = Inf, oldw = NULL) {
361   qo1 ← list(sense = "min")
362   # If no diversification- and turnover constraint
363   if(div== 0) {

```

```

364 if(tr == Inf) {
365   n ← (1+nrow(Q))*(nrow(Q)/2); qo1$c ← rep(0,N)
366   i ← do.call(c,sapply(1:N, function(i) seq(i,N)))
367   j ← do.call(c,sapply(1:N, function(j) rep(j, (N+1-j))))
368   v ← as.numeric((2*sapply(1:n, function(x,Q) Q[j[x],i[x]], Q=Q)))
369   qo1$qobj ← list(i = c(i),j = c(j), v= c(v))
370   # Constraints: Full investments
371   qo1$A ← Matrix(rep(1,N), nrow=1, byrow=TRUE, sparse=TRUE)
372   qo1$bc ← rbind(blc = 1, buc = 1)
373   # Box constraints: Long-only
374   qo1$bx ← rbind(blx = rep(0.0,N), bux = rep(1.0,N))
375   # Solve optimization procedure by interior-point method
376   return(mosek(qo1, opts = list(verbose = 1))$sol$itr$xx[1:N])
377   # If turnover constraint, but no diversification constraint
378 } else {
379   Q ← cbind(rbind(Q,matrix(0,2*N,N)), matrix(0,3*N,2*N)); qo1$c ← c(rep(0,(3*N)))
380   n ← (1+nrow(Q))*(nrow(Q)/2)
381   i ← do.call(c,sapply(1:(3*N), function(i) seq(i,3*N)))
382   j ← do.call(c,sapply(1:(3*N), function(j) rep(j, (3*N+1-j))))
383   v ← as.numeric((2*sapply(1:n, function(x,Q) Q[j[x],i[x]], Q=Q)))
384   qo1$qobj ← list(i = c(i),j = c(j), v= c(v))
385   # Full-investment and turnover constraint
386   qo1$A ← Matrix(rbind(rbind(c(rep(1,N),rep(0,2*N)), c(rep(0,N),rep(1,2*N))),
387                     cbind(diag(1,N), diag(-1,N),diag(1,N))), sparse = TRUE)
388   qo1$bc ← rbind(blc = c(1,-Inf, oldw), buc = c(1,tr, oldw))
389   # Inequality constraint: Portfolio with basic constraints
390   qo1$bx ← rbind(blx = rep(0.0,3*N), bux = c(rep(Inf,(3*N))))
391   # Solve optimization procedure by interior-point method
392   return(mosek(qo1, opts = list(verbose = 1))$sol$itr$xx[1:N])
393 }
394 # If diversification constraint, but no turnover constraint
395 } else {
396   if(tr == Inf) {
397     n ← (1+nrow(Q))*(nrow(Q)/2); qo1$c ← rep(0,N)
398     i ← do.call(c,sapply(1:N, function(i) seq(i,N)))
399     j ← do.call(c,sapply(1:N, function(j) rep(j, (N+1-j))))
400     v ← as.numeric((2*sapply(1:n, function(x,Q) Q[j[x],i[x]], Q=Q)))
401     qo1$qobj ← list(i = c(i),j = c(j), v= c(v))
402     # Constraints: Full investments & diversification
403     qo1$A ← Matrix(c(rep(1,N),rep(0,N)), nrow=2, byrow=TRUE, sparse=TRUE)
404     oprc ← sapply(1:N, function(j) list("POW", 2, j, 1, 2, 0))
405     rownames(oprc) ← c("type","i","j","f","g","h")
406     qo1$scopt ← list(oprc=oprc)
407     qo1$bc ← rbind(blc = c(1,-Inf), buc = c(1,(1-div)))
408     # Box constraints: Long-only
409     qo1$bx ← rbind(blx = rep(0,N), bux = rep(1,N))
410     # Solve optimization procedure by interior-point method
411     return(mosek(qo1, opts = list(verbose = 1))$sol$itr$xx[1:N])
412     # If diversification- and turnover constraint
413   } else {
414     Q ← cbind(rbind(Q,matrix(0,2*N,N)), matrix(0,3*N,2*N)); qo1$c ← c(rep(0,(3*N)))
415     n ← (1+nrow(Q))*(nrow(Q)/2)
416     i ← do.call(c,sapply(1:(3*N), function(i) seq(i,3*N)))
417     j ← do.call(c,sapply(1:(3*N), function(j) rep(j, (3*N+1-j))))

```

```

418 v ← as.numeric((2*sapply(1:n, function(x,Q) Q[j[x],i[x]], Q=Q)))
419 qo1$qobj ←list(i = c(i),j = c(j), v= c(v))
420 # Full-investment, turnover and div constraint
421 qo1$A ←Matrix(rbind(rbind(c(rep(1,N),rep(0,2*N))), c(rep(0,N),rep(1,2*N))),
422               cbind(diag(1,N), diag(-1,N),diag(1,N)), rep(0,3*N)), sparse = TRUE)
423 oprc ←sapply(1:N, function(j) list("POW", (N+3), j, 1, 2, 0))
424 rownames(oprc) ←c("type","i","j","f","g","h")
425 qo1$scopt ←list(oprc=oprc)
426 qo1$bc ←rbind(blc = c(1,-Inf, oldw,-Inf), buc = c(1,tr, oldw,(1-div)))
427 # Inequality constraint:
428 qo1$bx ←rbind(blx = rep(0,3*N), bux = c(rep(Inf,(3*N))))
429 return(mosek(qo1, opts = list(verbose = 1))$sol$itr$xx[1:N])
430 }
431 }
432 }

```

---

**Listing 6:** Load modified DEoptim scripts from code "constrained\_objective.R"

---

```

434 # Line numbers refer to the lines in "constrained_objective.R" found from
435 # https://github.com/R-Finance/PortfolioAnalytics/tree/master/R
436
437 # First step of PGP
438
439 # Formula 45 (G1c)
440 # Code 1: Function for maximization of skewness for MVS portfolio
441 source("../TurnandDist/Step1/MVS_maxskew.R")
442 # Code 2: Function for maximization of skewness for MVSK portfolio
443 source("../TurnandDist/Step1/MVSK_maxskew.R")
444
445 # Formula 46 (G1d)
446 # Code 3: Function for minimization of kurtosis for MVSK portfolio
447 source("../TurnandDist/Step1/MVSK_minkurt.R")
448
449 # Second step of PGP
450
451 # Formula 47 (G2)
452 # Code 4: Function for minimizing Z (MV portfolio)
453 source("../Moments/Step2/MV.R")
454 # Code 5: Function for minimizing Z (MVS portfolio)
455 source("../Moments/Step2/MVS.R")
456 # Code 6: Function for minimizing Z (MVSK portfolio)
457 source("../Moments/Step2/MVSK.R")
458
459 # Load dependent scripts to our modified DEoptim script
460 source("../PortfolioAnalytics-master/R/utils.R")
461 source("../PortfolioAnalytics-master/R/constraints.R")
462 source("../PortfolioAnalytics-master/R/generics.R")

```

---

**Listing 7:** objective function for non-convex optimization problems

---

```

464 # Modified from "constrained_objective.R"
465 # First step of PGP
466 # Formula 39 (G1c) and Formula 40 (G1d)
467 # For Code 1 and 2: Maximize Skew (G1c) for MVS- and MVSK portfolio

```

```

468 #(Formula 39 excl. constraints)
469 # "-" used to convert the equation to a maximization problem from a minimization problem.
470 tmp_measure ← -((w%%sM3%%kronecker(transpose(w),transpose(w)))/
471 ((w%%sM2d%%transpose(w))^1.5))
472
473 # For Code 3: Compute the minimum kurtosis (G1d) for MVSK portfolio
474 #(Formula 40 excl. constraints)
475 tmp_measure ← ((w%%sM4%%kronecker(kronecker(transpose(w),transpose(w)),
476 transpose(w)))/((w%%sM2d%%transpose(w))^2))w%%sM3%%kronecker(transpose(w),
477 transpose(w)))/((w%%sM2d%%transpose(w))^1.5))
478
479 # Second step of PGP
480 # Formula 41 (G2)
481
482 # For Code 4: Compute PGP (Step 2) for the MV portfolio (Formula 7)
483 # The skewness and kurtosis expressions in PGP equal 1, the skewness and kurtosis
484 # expression are adapted from Code 6
485 tmp_measure ← (abs((mvRet - m%%transpose(w))/mvRet)^1 + abs((w%%sM2%%transpose(w) - mvV)/mvV)^1)
486 abs((mvskSk - ((w%%sM3%%kronecker(transpose(w),transpose(w)))/((w%%sM2d%%transpose(w))^1.5)))/
487 mvskSk)^0 + abs((((w%%sM4%%kronecker(kronecker(transpose(w),transpose(w)),transpose(w)))/
488 (w%%sM2d%%transpose(w))^2) - mvskk)/mvskk)^0)
489
490 # For Code 5: Compute PGP (Step 2) for the MV portfolio (Formula 7)
491 # The kurtosis expression in PGP equal 1, the kurtosis expression is adapted from Code 6
492 tmp_measure ← (abs((mvsRet - m%%transpose(w))/mvsRet)^1 + abs((w%%sM2%%transpose(w) - mvsV)/
493 mvsV)^1 + abs((mvsSk - ((w%%sM3%%kronecker(transpose(w),transpose(w)))/
494 ((w%%sM2d%%transpose(w))^1.5)))/mvsSk)^1) + abs((((w%%sM4%%kronecker(kronecker(transpose(w),
495 transpose(w)),transpose(w)))/((w%%sM2d%%transpose(w))^2) - mvskk)/mvskk)^0)
496
497 # For Code 6: Compute PGP (Step 2) for the MVSK portfolio (Formula 7)
498 tmp_measure ← (abs((mvskRet - m%%transpose(w))/mvskRet)^1 + abs((w%%sM2%%transpose(w) - mvskV)/
499 mvskV)^1 + abs((mvskSk - ((w%%sM3%%kronecker(transpose(w),transpose(w)))/
500 ((w%%sM2d%%transpose(w))^1.5)))/mvskSk)^1) + abs((((w%%sM4%%kronecker(kronecker(transpose(w),
501 transpose(w)),transpose(w)))/((w%%sM2d%%transpose(w))^2) - mvskk)/mvskk)^1)

```

---

### Listing 8: Diversification penalty

---

```

503 # Apply diversification constraint if added, skip otherwise
504 if(!is.null(constraints$div_target)){
505   # Define diversification target (Formula: 7)
506   div_target ← constraints$div_target
507   # Calculate the diversification rate for the estimated weights
508   div ← diversification(w)
509   # Penalize the portfolio if the div.rate is lower than div. target (Formula: 7)
510   if((div < div_target) | (div > 1)){
511     out ← out + penalty * abs(div - div_target)
512   }
513 }

```

---

### Listing 9: Turnover penalty

---

```

515 # Apply diversification constraint if added, skip otherwise
516 if(!is.null(constraints$turnover_target)){
517   # Define turnover target (Formula: 8)

```

```

518 turnover_target ← constraints$turnover_target
519 # Changes in Code 4 for the MV weights
520 # Calculate the turnover rate from the last to next quarter
521 to ← turnover(w, mvlw)
522 # Penalize the portfolio if the turnover rate is above the turnover target (Formula 8)
523 if((to < turnover_target * 0) | (to > turnover_target)){
524   out = out + penalty * mult * abs(to - turnover_target)
525 }
526 }
527
528 # Code 1 and 5: Customized penalties for the MVS portfolio
529 to ← turnover(w, mvslw)
530 # Code 2,3 and 6: Customized penalties for the MVS portfolio
531 to ← turnover(w, mvsklw)

```

---

#### Listing 10: Adjust objective function to include penalties

---

```

533 out ← out + abs(objective$multiplier)*tmp_measure
534 # out is calculated from the penalties of the div. constraint
535 # and turn. constraint
536 # tmp_measure is the objective function excluding div. and turn. constraint

```

---

#### Listing 11: Dynamic portfolio optimization

---

```

537 # Length of estimation window (trading days)
538 L ← 261
539 # Number of portfolio revisions
540 K ← 80
541
542 for(i in 1:K) {
543   for(o in 1:nrow(ln_returns)) {
544     # search column for the "i+4" for the first date of the estimation period
545     # Replace aggreb4 med aggreb
546     if (ln_returns$aggreb4[o] == (i+4)) {
547       # acquire the row number for the first date of the estimation period
548       number ← ln_returns$obs[o+1]
549     }
550   }
551   # extract the first day of the estimation period
552   a ← c(index$Date[1 + number-L])
553   # extract the last day of the estimation period
554   b ← c(index$Date[number-1])
555   # store the date for the start of the start of the forecast period
556   date[i,] ← c(index$Date[number])
557
558   for(o in 1:nrow(ln_returns)) {
559     # search column for the "i+5" for the last date of the forecast period
560     if (ln_returns$aggreb4[o] == (i+5)) {
561       # acquire the row number for the last date of the forecast period
562       number2 ← ln_returns$obs[o]
563     }
564   }
565
566   # Generate the length of the forecast period

```



```

567 # H is the length of the holding period
568 H ← (number2 - number)
569 # Extract the rolling period using the eXtensible Time Series (xts) package
570 R2 ← (window(index, start = a, end = b))
571 # Convert the extracted rolling period to a data frame
572 R ← as.data.frame(window(index, start = a, end = b))
573
574 # Repeat the procedure for the market vector
575 M ← num(window(market, start = a, end = b))
576
577 # Acquire the time period for the estimated period
578 T ← nrow(R)
579 # Acquire the number of stocks for the estimated period
580 N ← ncol(R)
581
582 # Extract shrinkage estimates
583 m ← shrink_list[[1]]
584 sM2 ← shrink_list[[2]]
585 sM3 ← shrink_list[[3]]
586 sM4 ← shrink_list[[4]]
587
588 # Specify turnover constraint and assign weight for previous date
589 if(i == 1) {
590   # set last year's weights to null
591   mvplw ← NULL; mvslw ← NULL; mvklw ← NULL; mvsklw ← NULL; mvlw ← NULL
592   # store turnover rate
593   secondturnover ← turnover rate
594   # set turnover rate to inf
595   turnover rate ← Inf
596 } else {
597   # set stored turnover rate to the second period
598   if((constraint_t == 1) & (i == 2)) {
599     turnover rate ← secondturnover
600   }
601   if(constraint_t == 1) {
602     # read last year's weights (assign to variable)
603     mvplw ← as.double(W1[(i - 1), -(1:1)])
604     mvlw ← as.double(W2[(i - 1), -(1:1)])
605     mvslw ← as.double(W3[(i - 1), -(1:1)])
606     mvsklw ← as.double(W4[(i - 1), -(1:1)])
607   }
608 }
609
610 # Diversification constraint follows the same formula for all portfolios
611 # divrate ← 0 is equivalent to no diversification constraint
612
613 # Interior-point method to solve maximization of mean
614 res$mvuw ← optmean(m=m, N=N, div=divrate, tr=turnover rate*N, oldw=mvlw); mvRet ← m%*%res$mvuw
615 res$mvsw ← optmean(m=m, N=N, div=divrate, tr=turnover rate*N, oldw = mvslw); mvsRet ← m%*%res$mvsw
616 res$mvskw ← optmean(m=m, N=N, div=divrate, tr=turnover rate*N, oldw=mvsklw); mvskRet ← m%*%res$mvskw
617
618 # Interior-point method to solve maximization of minimum variance
619 wvV ← optvar(Q=sM2, N=N, div=divrate, tr=turnover rate*N, oldw=vlw); vV ← variance(wvV)
620 wmvV ← optvar(Q=sM2, N=N, div=divrate, tr=turnover rate*N, oldw=mvlw); mvV ← variance(wmvV)

```

```

621 wmvsvV ← optvar(Q=sM2, N=N, div=divrate, tr=turnoverrate*N, oldw=mvslw); mvsV ← variance(wmvsvV)
622 wmvskV ← optvar(Q=sM2, N=N, div=divrate, tr=turnoverrate*N, oldw=mvsklw); mvskV ← variance(wmvskV)
623
624 # Set constraints and specify the settings for DEoptim
625 funds ← colnames(R2)
626 init.portf ← portfolio.spec(assets=funds)
627 init.portf ← add.constraint(portfolio=init.portf, type="full_investment")
628 init.portf ← add.constraint(portfolio=init.portf, type="long_only")
629 init.portf ← add.objective(portfolio=init.portf, type="risk", name="StdDev")
630 init.portf ← add.constraint(portfolio=init.portf, type="weight_sum",
631                             min_sum=0.99, max_sum=1.01)
632 if(constraint_d ==1) {init.portf ← add.constraint(portfolio=init.portf,
633                                                    type="diversification", div_target=divrate)}
634 if((i > 1)&(constraint_t == 1)) {init.portf ← add.constraint(portfolio=init.portf,
635                                                             type="turnover", turnover_target=turnoverrate)}
636
637 # Apply settings for DEoptim
638 portfolio = init.portf
639 optimize_method="DEoptim"
640 search_size=20000
641 trace=TRUE
642 momentFUN='set.portfolio.moments'
643 parallel = TRUE
644 rp=NULL
645
646 N ← length(portfolio$assets)
647 if (ncol(R) > N) {
648   R ← R[,names(portfolio$assets)]
649 }
650
651 out ← list(); weights ← NULL
652 constraints ← get_constraints(portfolio)
653 .formals ← formals(momentFUN)
654 .formals ← modify.args(formals=.formals, arglist=NULL, dots=FALSE)
655 .formals ← modify.args(formals=.formals, arglist=NULL, R=R, dots=FALSE)
656 .formals ← modify.args(formals=.formals, arglist=NULL, portfolio=portfolio, dots=FALSE)
657 .formals$... ← NULL
658
659 if(optimize_method == "DEoptim"){
660   stopifnot("package:DEoptim" %in% search() || require("DEoptim",quietly = TRUE))
661
662   if(hasArg(itermax)) itermax=match.call(expand.dots=TRUE)$itermax else itermax=N*50
663
664   #Specify parameters according to section 3.3.4
665   NP ← round(search_size/itermax)
666   # Set population size
667   if(NP < (N * 15)) NP ← N * 15
668   if(!hasArg(itermax)) {
669     itermax ← round(search_size / NP)
670     #set minimum number of generations
671     if(itermax < 20000) itermax ← 20000
672   }
673
674   # Store settings for DEoptim

```

```

675 DEcformals ←formals(DEoptim.control)
676 DEcargs ←names(DEcformals)
677 pm ← pmatch(names(dotargs), DEcargs, nomatch = 0L)
678 names(dotargs[pm > 0L]) ←DEcargs[pm]
679 DEcformals$NP ←NP
680 DEcformals$itermax ←itermax
681 DEcformals[pm] ←dotargs[pm > 0L]
682
683 # Our specification of DEoptim parameters
684 # use DE/current-to-p-best/1 (Formula 61)
685 if(!hasArg(strategy)) DEcformals$strategy=6
686 # 1/1000000 of 1% change in objective is significant
687 if(!hasArg(reltol)) DEcformals$reltol=1/1000000
688 # number of assets times 50 tries to improve
689 if(!hasArg(steptol)) DEcformals$steptol=N*50
690 # JADE mutation parameter (set by default)
691 if(!hasArg(c)) DEcformals$c=.4
692 if(!hasArg(storepopfrom)) DEcformals$storepopfrom=1
693 }
694
695 # Add constraint for lower and upper bound (Formula 58)
696 traceDE=FALSE; DEcformals$trace ←traceDE; upper ←constraints$max; lower ←constraints$min
697
698 # Additional settings for DEoptim
699 if(hasArg(rp_method)) rp_method=match.call(expand.dots=TRUE)$rp_method else rp_method="sample"
700 if(hasArg(fev)) fev=match.call(expand.dots=TRUE)$fev else fev=0:5
701 rp ← random_portfolios(portfolio=portfolio, permutations=(NP+1),
702                       rp_method=rp_method, eliminate=FALSE, fev=fev)
703 DEcformals$initialpop ←rp
704 controlDE ←do.call(DEoptim.control, DEcformals)
705 controlDE$trace ←1000; controlDE$itermax ←20000
706
707 # include these specification of DEoptim optimization in the do.call function
708 cond ←list(lower=lower[1:N], upper=upper[1:N], control=controlDE, R=R, portfolio=portfolio,
709           env=dotargs, normalize=FALSE, fnMap=function(x) fn_map(x, portfolio=portfolio)$weights)
710
711 # Please find the codes from Listing 6
712 # The function from the sourced code is indicated as
713 # the function before cond in the do.call function
714
715 # Step 1 in PGP
716 # Formula 39 (G1c)
717 # Code 1: Maximize skewness for MVS portfolio
718 maxsmvs = try(do.call(DEoptim, c(max_skew1,cond)), silent = FALSE)
719 wmvSsk ←maxsmvs$optim$bestmem; mvSsk ←-maxsmvs$optim$bestval
720 print(paste("Max Skewness Value (MVS) and Weights:", (mvSsk))); print(wmvS,k)
721
722 # Code 2: Maximize skewness for MVSK portfolio
723 maxsmvsk = try(do.call(DEoptim, c(max_skew2,cond)), silent = FALSE)
724 wmvskSk ←maxsmvsk$optim$bestmem; mvskSk ←-maxsmvsk$optim$bestval; rm(maxsmvsk)
725 print(paste("Max Skewness Value (MVSK) and Weights:", (mvskSk)))
726
727 # Formula 40 (G1d)
728 # Code 3: Minimize kurtosis for MVSK portfolio

```

```

729   minkmvsk = try(do.call(DEoptim, c(min_kurt,cond)), silent = FALSE)
730   wmvskk ← minkmvsk$optim$bestmem; mvskk ← minkmvsk$optim$bestval
731   print(paste("Min Kurtosis Value (MVSK) and Weights:", (mvskk))); print(wmvskk)
732
733   # Step 2 in PGP
734
735   # Formula 41 (G2)
736   # Code 4: Derive weight vector by minimizing Z (MV portfolio)
737   mv = try(do.call(DEoptim, c(o_mv,cond)), silent = FALSE)
738   mvweights ← mv$optim$bestmem; mvv ← mv$optim$bestval
739   rm(mv); gc(reset = TRUE)
740
741   # Code 5: Derive weight vector by minimizing Z (MVS portfolio)
742   mvs = try(do.call(DEoptim, c(o_mvs,cond)), silent = FALSE)
743   mvsweights ← mvs$optim$bestmem; msv ← mvs$optim$bestval
744   rm(mvs); gc(reset = TRUE)
745   print(paste("MVS and Weights:", (msv))); print(mvsweights)
746
747   # Code 6: Derive weight vector by minimizing Z (MVSK portfolio)
748   mvsk = try(do.call(DEoptim, c(o_mvsk,cond)), silent = FALSE)
749   mvskweights ← mvsk$optim$bestmem; mvskv ← mvsk$optim$bestval
750   rm(mvsk); gc(reset = TRUE)
751   print(paste("MVSK and Weights:", (mvskv))); print(mvskweights)
752 }
753
754 # Weights for the holding period:
755 w_V ← round(wvV,5)
756 w_MV ← round(mvweights/sum(mvweights),5)
757 w_MVS ← round(mvsweights/sum(mvsweights),5)
758 w_MVSK ← round(mvskweights/sum(mvskweights),5)
759 }

```

---

### Listing 12: Hypothesis test

---

```

760 # Load monthly returns
761 mret ← read.csv("mrett.csv")
762 n = ncol(mret)
763
764 # Function for difference in moments
765 meandiff ← function(ret) { return(mean(ret[,1]) - mean(ret[,2]))}
766 vardiff ← function(ret) { return(log(var(ret[,1])) - log(var(ret[,2])))}
767 skewdiff ← function(ret) { return(skewness(ret[,1], method = "sample") -
768                                   skewness(ret[,2], method = "sample"))}
769 kurtdiff ← function(ret) { return(kurtosis(ret[,1], method = "sample_excess") -
770                                   kurtosis(ret[,2], method = "sample_excess"))}
771
772 # Alpha-component in SE for HAC
773 alpha ← function (V.hat){
774   dimensions = dim(V.hat)
775   T = dimensions[1]
776   p = dimensions[2]
777   o = 0; u = 0
778   for (i in (1:p)) {
779     # AR(1) with ols estimates

```

```

780   fit = ar(V.hat[, i], 0, 1, method = "ols")
781   rho.hat = as.numeric(fit[2])
782   sig.hat = sqrt(as.numeric(fit[3]))
783   u = u + 4 * rho.hat^2 * sig.hat^4/(1-rho.hat)^8
784   o = o + sig.hat^4/(1 - rho.hat)^4
785 }
786 return(u/o)
787 }
788
789 # Parzen kernel estimate for Psi
790 Psi.hat <-function (V.hat, moment_num) {
791   T <- length(V.hat[, 1])
792   alpha.hat <-alpha(V.hat)
793   S.star <-2.6614*(alpha.hat*T)^0.2
794   Psi.hat <-compute.Gamma.hat(V.hat, 0)
795   j <- 1
796   while (j < S.star) {
797     Gamma.hat <-compute.Gamma.hat(V.hat, j)
798     Psi.hat <-Psi.hat + kernel.Parzen(j/S.star)*(Gamma.hat + t(Gamma.hat))
799     j <- j + 1
800   }
801   Psi.hat <-(T/(T-2*moment_num))*Psi.hat
802   return(Psi.hat)
803 }
804
805 # Asymptotic standard error for Parzen kernel for each moment
806 ste.pw <-function(ret, moment){
807   ret1 <-ret[, 1]
808   ret2 <-ret[, 2]
809   T <- length(ret1)
810   mu1.hat <-mean(ret1)
811   mu2.hat <-mean(ret2)
812
813   if(moment == "mean"){
814     moment_num <-1
815     # Gradient vector
816     gradient <-rep(0, 2)
817     gradient[1] <-1
818     gradient[2] <- -1
819     T <- length(ret1)
820     V.hat <- cbind(ret1 - mean(ret1), ret2 - mean(ret2))
821     A.ls <-matrix(0,2,2)
822     V.star <-matrix(0,(T-1),2)
823     reg1 <-V.hat[1:T-1,1]
824     reg2 <-V.hat[1:T-1,2]
825     for(j in 1:2) {
826       fit <-lm(V.hat[2:T,j] ~ -1 + reg1 + reg2)
827       A.ls[j,] <-as.numeric(fit$coef)
828       V.star[,j] <-as.numeric(fit$resid)
829     }
830     svd.A <-svd(A.ls); d <-svd.A$d; d.adj <-d
831     for(i in 1:2){
832       if(d[i] > 0.97){d.adj[i] <-0.97}
833       else if(d[i] < -0.97){d.adj[i] <- -0.97}

```

```

834   }
835   A.hat ← svd.A$u%%diag(d.adj)%%t(svd.A$v)
836   D ← solve(diag(2) - A.hat)
837   reg.mat ← rbind(reg1, reg2)
838   for(j in 1:2) {
839     V.star[,j] ← V.hat[2:T,j]-A.hat[j,]%reg.mat
840   }
841 } else if(moment == "var"){
842   moment_num ← 2
843   gamma1.hat ← mean(ret1^2)
844   gamma2.hat ← mean(ret2^2)
845   # Gradient vector
846   gradient ← rep(0, 4)
847   gradient[1] ← -2*mu1.hat/(gamma1.hat - mu1.hat^2)
848   gradient[2] ← 2*mu2.hat/(gamma2.hat-mu2.hat^2)
849   gradient[3] ← 1/(gamma1.hat-mu1.hat^2)
850   gradient[4] ← -1/(gamma2.hat-mu2.hat^2)
851   T ← length(ret1)
852   V.hat ← cbind(ret1 - mean(ret1), ret2 - mean(ret2),
853                 ret1^2 - mean(ret1^2), ret2^2 - mean(ret2^2))
854   A.ls ← matrix(0,4,4)
855   V.star ← matrix(0,(T-1),4)
856   reg1 ← V.hat[1:T-1,1]
857   reg2 ← V.hat[1:T-1,2]
858   reg3 ← V.hat[1:T-1,3]
859   reg4 ← V.hat[1:T-1,4]
860   for(j in 1:4){
861     fit ← lm(V.hat[2:T,j] ~ -1 + reg1 + reg2 + reg3 + reg4)
862     A.ls[j,] ← as.numeric(fit$coef)
863     V.star[,j] ← as.numeric(fit$resid)
864   }
865   svd.A ← svd(A.ls)
866   d ← svd.A$d; d.adj ← d
867   for(i in 1:4){
868     if(d[i] > 0.97) {d.adj[i] ← 0.97}
869     else if(d[i] < -0.97) {d.adj[i] ← -0.97}
870   }
871   A.hat ← svd.A$u%%diag(d.adj)%%t(svd.A$v)
872   D ← solve(diag(4) - A.hat)
873   reg.mat ← rbind(reg1, reg2, reg3, reg4)
874   for(j in 1:4){
875     V.star[,j] ← V.hat[2:T,j] - A.hat[j,]%reg.mat
876   }
877 } else if(moment == "skew"){
878   moment_num ← 3
879   gamma1.hat ← mean(ret1^2)
880   gamma2.hat ← mean(ret2^2)
881   s1.hat ← mean(ret1^3)
882   s2.hat ← mean(ret2^3)
883   # Gradient vector
884   gradient ← rep(0,6)
885   gradient[1] ← (-3*gamma1.hat^2+3*mu1.hat*s1.hat)/((gamma1.hat-mu1.hat^2)^2.5)
886   gradient[2] ← -(-3*gamma2.hat^2+3*mu2.hat*s2.hat)/((gamma2.hat-mu2.hat^2)^2.5)
887   gradient[3] ← (-3*s1.hat+3*mu1.hat*gamma1.hat)/(2*(gamma1.hat-mu1.hat^2)^2.5)

```

```

888 gradient[4] ← -(-3*s2.hat+3*mu2.hat*gamma2.hat)/(2*(gamma2.hat-mu2.hat^2)^2.5)
889 gradient[5] ← 1/((gamma1.hat-mu1.hat^2)^1.5)
890 gradient[6] ← -1/((gamma2.hat-mu2.hat^2)^1.5)
891 T ← length(ret1)
892 V.hat ← cbind(ret1 - mean(ret1), ret2 - mean(ret2),
893             ret1^2 - mean(ret1^2), ret2^2 - mean(ret2^2),
894             ret1^3 - mean(ret1^3), ret2^3 - mean(ret2^3))
895 A.ls ← matrix(0,6,6)
896 V.star ← matrix(0,T-1,6)
897 reg1 ← V.hat[1:T-1,1]
898 reg2 ← V.hat[1:T-1,2]
899 reg3 ← V.hat[1:T-1,3]
900 reg4 ← V.hat[1:T-1,4]
901 reg5 ← V.hat[1:T-1,5]
902 reg6 ← V.hat[1:T-1,6]
903 for(j in 1:6){
904   fit ← lm(V.hat[2:T,j] ~ -1 + reg1 + reg2 + reg3 + reg4 + reg5 + reg6)
905   A.ls[j,] ← as.numeric(fit$coef)
906   V.star[,j] ← as.numeric(fit$resid)
907 }
908 svd.A ← svd(A.ls)
909 d ← svd.A$d; d.adj ← d
910 for(i in 1:6){
911   if(d[i] > 0.97){d.adj[i] ← 0.97}
912   else if(d[i] < -0.97){d.adj[i] ← -0.97}
913 }
914 A.hat ← svd.A$u%*%diag(d.adj)%*%t(svd.A$v)
915 D ← solve(diag(6)-A.hat)
916 reg.mat ← rbind(reg1,reg2,reg3,reg4,reg5,reg6)
917 for(j in 1:6){
918   V.star[,j] ← V.hat[2:T,j] - A.hat[j,]%*%reg.mat
919 }
920 } else if(moment == "kurtosis"){
921   moment_num ← 4
922   gamma1.hat ← mean(ret1^2)
923   gamma2.hat ← mean(ret2^2)
924   s1.hat ← mean(ret1^3)
925   s2.hat ← mean(ret2^3)
926   k1.hat ← mean(ret1^4)
927   k2.hat ← mean(ret2^4)
928   # Gradient vector
929   gradient ← rep(0, 8)
930   gradient[1] ← (12*mu1.hat*gamma1.hat^2-12*(mu1.hat^2)*s1.hat+
931               4*mu1.hat*k1.hat-4*gamma1.hat*s1.hat)/((gamma1.hat-mu1.hat^2)^3)
932   gradient[2] ← -(12*mu2.hat*gamma2.hat^2-12*(mu2.hat^2)*s2.hat+
933                4*mu2.hat*k2.hat-4*gamma2.hat*s2.hat)/((gamma2.hat-mu2.hat^2)^3)
934   gradient[3] ← (-6*(mu1.hat^2)*gamma1.hat+8*mu1.hat*s1.hat
935                -2*k1.hat)/((gamma1.hat-mu1.hat^2)^3)
936   gradient[4] ← (-6*(mu2.hat^2)*gamma2.hat+8*mu2.hat*s2.hat
937                -2*k2.hat)/((gamma2.hat-mu2.hat^2)^3)
938   gradient[5] ← -4*mu1.hat/((gamma1.hat-mu1.hat^2)^2)
939   gradient[6] ← -4*mu2.hat/((gamma2.hat-mu2.hat^2)^2)
940   gradient[7] ← 1/((gamma1.hat-mu1.hat^2)^2)
941   gradient[8] ← -1/((gamma2.hat-mu2.hat^2)^2)

```

```

942 T ← length(ret1)
943 V.hat ← cbind(ret1-mean(ret1), ret2-mean(ret2), ret1^2-mean(ret1^2),
944             ret2^2-mean(ret2^2), ret1^3-mean(ret1^3), ret2^3-mean(ret2^3),
945             ret1^4-mean(ret1^4), ret2^4-mean(ret2^4))
946 A.ls ← matrix(0,8,8)
947 V.star ← matrix(0,T-1,8)
948 reg1 ← V.hat[1:T-1,1]
949 reg2 ← V.hat[1:T-1,2]
950 reg3 ← V.hat[1:T-1,3]
951 reg4 ← V.hat[1:T-1,4]
952 reg5 ← V.hat[1:T-1,5]
953 reg6 ← V.hat[1:T-1,6]
954 reg7 ← V.hat[1:T-1,7]
955 reg8 ← V.hat[1:T-1,8]
956 for(j in 1:8) {
957   fit ← lm(V.hat[2:T,j] ~ -1+reg1+reg2+reg3+reg4+reg5+reg6+reg7+reg8)
958   A.ls[j,] ← as.numeric(fit$coef)
959   V.star[,j] ← as.numeric(fit$resid)
960 }
961 svd.A ← svd(A.ls)
962 d ← svd.A$d; d.adj ← d
963 for(i in 1:8){
964   if(d[i] > 0.97){d.adj[i] ← 0.97}
965   else if(d[i] < -0.97){d.adj[i] ← -0.97}
966 }
967 A.hat ← svd.A$u%%diag(d.adj)%%t(svd.A$v)
968 D ← solve(diag(8)-A.hat)
969 reg.mat ← rbind(reg1, reg2, reg3, reg4, reg5, reg6, reg7, reg8)
970 for(j in 1:8) {
971   V.star[,j] ← V.hat[2:T,j]-A.hat[j,]%%reg.mat
972 }
973 }
974 ##Compute standard.error based on the chosen moment
975 Psi.hat ← D%%Psi.hat(V.star, moment_num)%%t(D)
976 se ← as.numeric(sqrt(t(gradient)%%Psi.hat%%gradient/T))
977 return(se)
978 }
979
980 # P-value for difference test w/HAC-standard errors
981 hac ← function(returns, moment, digits = 3) {
982   ret1 = returns[, 1]
983   ret2 = returns[, 2]
984
985   if(moment == "mean") {
986     # Difference between each portfolio's mean
987     Delta.hat ← meandiff(returns)
988     # Standard error
989     se.pw ← ste.pw(ret=returns, moment = "mean")
990   } else if(moment == "var") {
991     # Difference between each portfolio's log var
992     Delta.hat ← vardiff(returns)
993     # Standard error
994     se.pw ← ste.pw(ret=returns, moment = "var")
995   } else if (moment == "skew") {

```



```
996   # Difference between each portfolio's skew
997   Delta.hat ←skewdiff(returns)
998   # Standard error
999   se.pw ←ste.pw(ret=returns, moment = "skew")
1000 } else if (moment == "kurtosis") {
1001   # Difference between each portfolio's excess kurtosis
1002   Delta.hat ←kurtdiff(returns)
1003   # Standard error
1004   se.pw ←ste.pw(ret=returns, moment = "kurtosis")
1005 }
1006 # Two-sided P-value
1007 PV.pw = 2*pnorm(-abs(Delta.hat)/se.pw)
1008 return(data.frame("Diff" = Delta.hat, "PV.pw" = PV.pw))
1009 }
```

---