Norwegian School of Economics

Bergen, Spring 2019

Study on Portfolio Selection with Skewness at Oslo Stock Exchange

By Isak Alexandersen and Kim Huu Luong

Supervisor: Nils Friewald

Master Thesis, Finance

NORWEGIAN SCHOOL OF ECONOMICS

Contents

List of Figures

List of Tables

Abstract

In this paper we investigate the statistical measure of skewness in a portfolio management setting at Oslo Stock Exchange (OSE). Our analysis follows earlier research on the topic of non-normal investor preferences which prices skewness as a relevant factor. We analyze distributional properties of monthly returns in individual assets and find on OSE that 1) skewness is pervasive, 2) positive skewness has a moderate level of persistence in the long term and can reasonably be predicted, and 3) diversification and skewness are negatively correlated. As a second major focal point, we form strategies which include a preference for skewness using Polynomial Goal Programming. We compare them to traditional portfolios using a traditional financial performance measure (Sharpe Ratio) and skewness. With two model specifications we find mixed results regarding skewness - the strategies are only able to produce higher portfolio skewness than the classical mean-variance portfolio in one scenario. A second finding is that we can not reject a null hypothesis of equal Sharpe Ratio between the skewness-strategies and the mean-variance portfolio. We also find that the two skewness-strategies are 1) less diversified than the other portfolios and 2) more risky as a consequence. Skewness and variance seem to be opposing goals for an individual with non-normal investor preferences.

1 Introduction

In this paper we study portfolio management at Oslo Stock Exchange (OSE) where we attempt to improve asset allocation by including skewness in addition to mean and variance. First we make an account of skewness of stock returns at OSE and investigate related distributional properties of single assets and portfolios. An overview of skewness by sector is also provided. Then we create a portfolio strategy with the first three central moments, called the Mean-Variance-Skewness (MVS) portfolio. The strategy is compared to other traditional portfolio strategies using the Sharpe Ratio in addition to assessing portfolio skewness to see if including skewness enables an investor to make significant gains. The aim is to make qualitative statements about the role of skewness in portfolio management where distributional properties of stock returns and non-normal investor preferences are the main points of consideration. In other words, we want to know if skewness is a worthwhile goal to include when forming portfolios consisting of risky assets trading at OSE.

The motivation for this analysis is based on a critical view of the weaknesses of Modern Portfolio Theory (MPT), mainly the simplistic model formulations and assumptions. MPT implicitly assumes normally distributed returns and normal investor preferences. Expanded versions of MPT could provide more explanatory power and add new intuitions; while standard deviation of returns shows how much risk an investor faces, skewness better reveals what type of risk investors choose to take. This thesis thus makes the implicit assumption that skewness is relevant for investors in a portfolio management setting. E.g. Harvey and Siddique (2000) find that skewness is in fact a priced factor which means investors are willing to trade expected return (or accept higher risk) for higher skewness. A closer examination of MVS-strategies at OSE is therefore a worthwhile endeavor.

In our research, we only extend the analysis to skewness and not to kurtosis (the fourth standardized central moment) or even higher moments. Firstly, only considering skewness as an extension of the MPT framework allows a sharper focus. Secondly, there may not exist strong behavioristic arguments for investor attitude towards kurtosis that are comparable to the first three moments (Kraus and Litzenberger, 1976), making skewness a clearer objective to research.

To cover the topic of skewness in portfolio management at OSE, we pursue two main lines of inquiry. Firstly, can we use skewness in portfolio management at OSE? This is covered in three parts: We calculate skewness in individual assets to ascertain whether normality in returns is a reasonable assumption. A brief overview of different sectors of the economy is also provided. Then we investigate whether skewness is persistent in individual assets, which is an indicator for how well we can predict future skewness. Finally we look into the effect portfolio size has on portfolio skewness which has implications on optimal asset allocation.

Secondly, does taking skewness into account in portfolio management improve out-of-sample performance using Polynomial Goal Programming? We form an optimal portfolio strategy called the Mean-Variance-Skewness portfolio and compare out-of-sample Sharpe Ratio and skewness with the Mean-Variance portfolio (MV), Global Minimum Variance Portfolio (GMV), and the equally-weighted portfolio (EW). Assessment of performance of the different portfolios are done by comparing Sharpe Ratio (SR) and Adjusted for Skewness Sharpe Ratio (ASSR).

The rest of the thesis is structured as follows. In section 2, we review relevant literature on the topic of non-normal investor preferences. In section 3 we present the data used in the analysis which include monthly stock returns from OSE and monthly returns from the Oslo Stock Exchange Benchmark Index (OSEBX). In section 4, we explain our methodology and the implementation of the MVS portfolio. We present the results and interpretation in section 5. Conclusion follows in section 6. Weaknesses and suggestions for further research are pointed out in section 7.

2 Literature Review

Since the initial research into portfolio management by Markowitz (1952), considerable effort has gone into researching skewness as a natural extension of the framework. We present relevant literature on the topic to create a brief review of work that has been done regarding non-normal investor preferences. The most important points of discussion pertain to the existence of skewness, behavioural explanations of non-normal investor preference, and how inclusion of higher moments affect optimal portfolio selection. Additionally we summarize relevant developments in portfolio optimization with higher moments to motivate our choice of method in this paper.

Throughout the literature slightly different definitions of skewness have been used for different purposes. For the sake of consistency, whenever this paper refers to skewness, it is defined as the standardized third central moment. We thus make a distinction regarding the separate (but closely related) statistical measure of the third central moment.[1] Additionally, positive skewness is characterized by a distribution of returns whose right tail is longer; for negative skewness the left tail is longer. With no skewness, returns are normally distributed with similar tails on either side. In economic terms, skewness signifies an increased possibility of extreme returns in either direction.

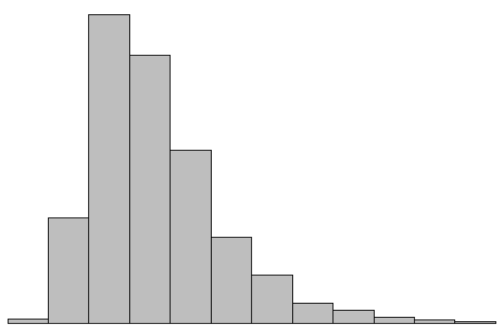

Figure 1: Illustration of a Distribution with Positive Skewness.

[1] The third central moment of a distribution is given by $E[(X - E[X])^3]$. Skewness, the standardized third moment, is on the form $\frac{E[(X-E[X])^3]}{(E[(X-E[X])^2])^{\frac{3}{2}}}$

2.1 General Insights

Early research into skewness has revealed some general results relevant to portfolio selection. Regarding the existence of skewness, Singleton and Wingender (1986) and Beedles and Simkowitz (1980) find in the US stock market that the distribution of stock returns clearly deviate from normality, and more importantly, that skewness is an abundant and persistent phenomenon. Both papers also point out another important finding, namely that positive skewness is significantly more common than negative. Beedles and Simkowitz (1980) note that twice as many stocks are positively skewed than negative and that the average skewness for all 500 stocks for all 5 year-periods from 1927 to 1976 is positive (skewness > 0). Singleton and Wingender (1986) find similar trends: roughly 50-75% of simple asset returns with dividends from 1961 to 1980 exhibit significant positive skewness (skewness > 0.3).[2] As to our knowledge, no comparable research has been done to shed light on skewness of stocks at OSE.

A well-established result regarding non-normal preferences is that investors will accept a lower expected return to obtain skewness in their portfolio all else equal. In other words: they are willing to 'trade' return for skewness. In particular, investors prefer positively skewed assets and portfolios and avoid negative skewness. This insight is supported by a host of empirical research in different markets and time periods, including papers on the New York Stock Exchange from 1936-1970 (Kraus and Litzenberger, 1976), 17 international stock indices from 1993-2000 (Prakash et al., 2003) among others (Chunhachinda et al., 1997; Harvey and Siddique, 2000; Canela and Collazo, 2007). The conclusion were drawn from two different methodological angles: (1) how well expanded models of the Capital Asset Pricing Model (CAPM) which included skewness fit empirical data, or (2) how well skewness-based portfolios performed relative to other strategies. The behavioral explanation commonly given is that an average risk averse person seeks lottery-like returns (small but frequent losses, extreme but rare gains), but dislikes the opposite situation where unlikely

[2] Classification of skewness is in this paper is based on a commonly used heuristic used by e.g. Singleton and Wingender (1968): positive skewness is $\hat{s} > 0.3$, negative skewness $\hat{s} < -0.3$, and no skewness $|\hat{s}| \leq 0.3$. For reference: a normally distributed variable has a skewness of 0 (which is a standard assumption in MPT).

but extreme losses are present.

Research also suggest that skewness does not only have statistical and theoretical significance, but can also create impacts of economic and practical importance. Lai (1991), Chunhachinda et al. (1997), Prakash et al. (2003) and Lai et al. (2006) find that including the third moment in a portfolio optimization problem causes a major change to the optimal allocation of assets, resulting in the aforementioned trading of mean for skewness. They argue that this result provides grounding for the claim that the MV portfolio is not optimal under the assumption of non-normal investor preference, and that the MVS portfolio may be a more realistic framework.

However, there may be difficulties in implementing the MVS portfolio because of a lack of predictability in skewness. Singleton and Wingender (1986) find in the US market (CRSP database) from 1961-1980, that skewness in single stocks and in portfolios does not persist over time. This is surprising given the fact that Beedles and Simkowitz (1980) also discovered that the share of skewed assets from 1927-1976 in the US market was fairly stationary and predictable. The conclusion that follows from a lack of persistence is that past skewness in assets or portfolios may not predict future skewness very well, making it more difficult to pick assets with desirable properties out-of-sample. On the other hand, Sun and Yan (2003) provides evidence from the US and Japanese markets in the period 1975-1995 that suggest the MVS portfolio may have better skewness persistence than other types of portfolios such as MV.

Another important point of note is that portfolio skewness decreases rapidly with increasing diversification (Beedles and Simkowitz, 1978; Beedles, 1979; Singleton and Wingender, 1986). The authors use portfolio size as a proxy to measure diversification. As such it is a naïve interpretation we also use in the rest of this paper: ‘one should put ones eggs (money) in several baskets (assets) to limit the loss in case there is a hole in a basket (negative development of asset price)’. In particular, an analysis by Beedles and Simkowitz (1978) find that over 92% of diversifiable skewness is eliminated in portfolios of 5 stocks. This implies that portfolio optimization with skewness actually revolves around systematic (non-diversifiable) and not total skewness. The authors argue that the diversification effect presents a tradeoff between two competing goals: one should diversify assets in order to remove unwanted risk (variance) in a portfolio, but one should also decrease diversification in order to attain desired portfolio skewness. This adds another layer of complexity on top of the original tradeoff between wanting more return but simultaneously desiring lower risk. Related to this point, Mitton and Vorkink (2007) and Beedles and Simkowitz conclude from the US market that when investor preference for skewness is introduced, the optimal portfolio tends to be relatively under-diversified. Mitton and Vorkink also discover that under-diversified investors outnumber diversified investors with a ratio of 26 to 1 among 65 562 US households which may reveal a widespread interest for skewness. Proposing non-normal preferences as an explanation for under-diversification may put to question a typical assertion of lack of rationality or otherwise imperfect capital markets.

2.2 Portfolio Optimization with Skewness

Expanding the MPT framework naturally increases the theoretical complexity and the computational requirements for pricing assets or constructing optimal portfolio strategies. There exists several techniques to solve portfolio optimization problems with higher moments, two of which have gained recognition in the literature. The first is called Polynomial Goal Programming (PGP) (also known as 'the primal approach'), originally developed by Tayi and Leonard (1988) and first put in use in portfolio selection by Lai (1991). It has seen application in different settings by e.g. Chunhachinda et al. (1997), Prakash et al. (2003) and Canela and Collazo (2007). The second is called 'the dual approach', a method with recent contributions made by e.g. Harvey et al. (2002), Jondeau and Rockinger (2006) and Briec et al. (2007). The two methods use different approaches: PGP is based on optimizing the distributional properties of a portfolio (mean, variance, skewness) by targeting several goals simultaneously, while 'the dual approach' applies a Taylor-series expansion with reference to the investor's utility function. PGP has seen more widespread use because of its main advantages, namely that (1) a globally optimal solution is guaranteed, (2) there is flexibility in including different investor preferences for the three moments, and (3) the computational requirements are relatively simple (Lai, 1991). The main disadvantage of PGP is that the objective function cannot be directly related to a utility function due to preference for the different moments being entirely ad-hoc (Jondeau and Rockinger, 2006). Conversely, the main advantage of 'the dual approach' is a clear relation to investor preference which harmonizes with well-established portfolio theory, but one of its drawbacks is not being able to guarantee a globally optimal solution due to inherent non-convexity (Briec et al., 2007).

Ultimately, there is no consensus on how to unify different optimization methods and consequently no 'correct' way to solve problems with higher moments, making analysis of optimal asset allocation a more difficult task than analyzing the MPT framework. We employ the PGP method for our optimization purposes. Further details on the method is found in the following section.

3 Methodology

In this section we outline the methodology underlying the implementation of portfolio optimization with skewness. The practical steps are the following. We start by separating the dataset into two categories: in-sample periods where estimation of the three moments is performed, and out-of-sample periods where the portfolio strategy is tested. We use the rolling window procedure to structure the dataset appropriately. The next step consists of specifying and implementing the method for moment estimation using the data in the in-sample periods. We use so-called robust moment estimation to calculate asset means, variance-covariance- and skewness-coskewness terms. An important aspect to note is that the robust moments found in the methodology section are only used for optimization purposes. They have no direct statistical or economic interpretation, and elsewhere in the paper we consistently use the sample third moment and sample skewness for all other purposes. In the third step we solve an optimization problem which create optimal asset weights that define the different investment strategies. Polynomial Goal Programming (PGP) is employed, a method that handles optimization in the presence of multiple conflicting goals. Finally, the resulting portfolios from the optimization are tested in all out-of-sample periods and different measures are used to assess the relative performance of each strategy.

Additionally, we have chosen to include transaction costs in our analysis to create a slightly more realistic framework. Transaction costs are used in the formulation of the PGP optimization model but they are also applied to out-of-sample testing of each strategy when they accrue returns.

3.1 Rolling Window Procedure

The rolling window procedure is a common economic forecasting technique often used in portfolio management. It imposes structure in the following way. First, a subset of length m of the data is chosen. The first l months of this "window" is the in-sample period which is used for estimating moments. After month l, optimal asset weights are calculated and the portfolio is rebalanced at month $l + 1$. Then the portfolio is held for h months, allowing it to accrue returns in the out-of-sample period $[l+1, l+h]$. The "window" is then shifted h months to obtain next subsample, and the procedure is repeated until the end of the dataset is reached.

We chose an estimation window of length $l = 120$ and holding period of length $h = 3$. As an illustration, the first estimation period is the months 1-120 and the first holding period the months 121-123. The second estimation period is the months 4-123 and the second holding period months 124-126, and so on. In practice, the choice of estimation- and holding period length is arbitrary; our choice is partly based on earlier research conducted in situations similar to ours, e.g. DeMiguel et al. (2009). With a dataset of 168 monthly observations of returns, we have 16 "windows" in total which generate $T - l = 48$ monthly out-of-sample returns from the months 121-168. Thus 10 years of returns are in-sample and a total of 4 are out-of-sample.

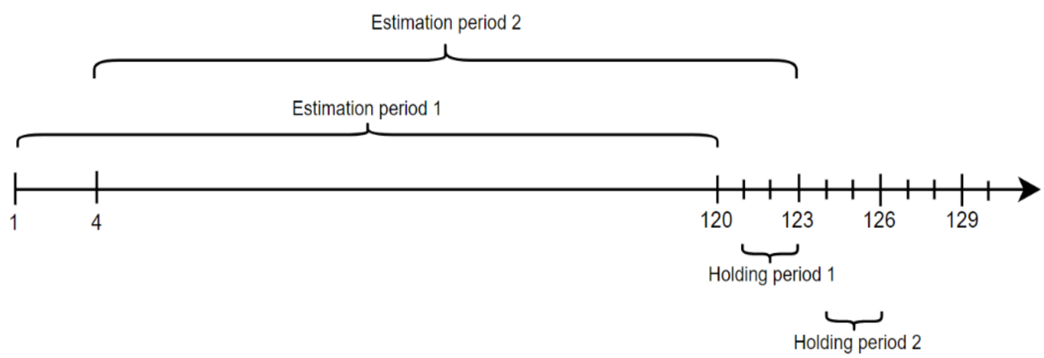

Figure 2: Illustration of Rolling Window Procedure for the First Two Periods.

3.2 Robust Moment Estimation

To calculate portfolio moments in the optimization procedure, one first needs to estimate moments for individual assets. In the MVS framework, this translates to calculating the mean, the variance-covariance matrix and the skewness-coskewness matrix of the stock returns. It is common to use historical 'plug-in' sample-moments for this purpose, where the mean and individual elements of the variance-covariance matrix and skewness-coskewness matrix are respectively

$$\hat{\mu}_i = \frac{1}{T}\sum_{t=1}^{T} r_{it} \qquad \forall i \tag{1}$$

$$\hat{\sigma}_{ij} = \frac{1}{T}\sum_{t=1}^{T}(r_{it} - \bar{r}_i)(r_{jt} - \bar{r}_j) \qquad \forall i, j \tag{2}$$

$$\hat{s}_{ijk} = \frac{1}{T}\sum_{t=1}^{T}(r_{it} - \bar{r}_i)(r_{jt} - \bar{r}_j)(r_{kt} - \bar{r}_k) \qquad \forall i, j, k \tag{3}$$

where $\bar{r}$ denotes the sample mean.

There are several commonly known problems with using sample-moments. Firstly is the curse of dimensionality: the amount of parameters one has to estimate increases exponentially with the number of stocks. With a portfolio of 20 stocks, one would need 45 years of monthly data to merely exceed the number of parameters needed to estimate portfolio variance, skewness and kurtosis (Martellini and Ziemann, 2010). A second problem is that the sample moments exhibit significant estimation error - deviations from the true values of what is being estimated.

To alleviate these shortcomings, one can replace sample estimators with robust moment estimators. Different classes of estimators have been used in financial and statistical literature, such as Maximum-likelihood-estimators (M-estimators). Another approach is called Shrinkage Estimation. The method was originally formalized by Stein (1956), adopted for estimating expected returns by Jorion (1986), further developed by Ledoit and Wolf (2003) to create a robust variance-covariance matrix, and finally extended to the skewness-coskewness and kurtosis-cokurtosis matrices by Martellini and Ziemann (2010). We use shrinkage estimation to create robust mean, variance-covariance and skewness-coskewness estimates. The method used for estimating the mean is slightly different from the two higher moments and we will give them separate treatments in the two following sections.

3.2.1 Robust Estimation of Expected Return

To compute robust estimates of expected return for each asset, we use the so-called Bayes-Stein estimator developed by Jorion (1986). The method builds upon the principles of shrinkage estimation by Stein (1956) and improves estimation of the expected return in presence of significant estimation error caused by e.g. outliers in the data. Bayes-stein estimation consists of calculating the usual sample estimates $\overline{Y}$, choosing a suitable target Y_0, and weighing the two components together to 'shrink' the sample-estimates towards the target. Y_0 can be any vector - even a vector of zeros - but the greatest gains are made when the target is closer to the true expected value. In his analysis, Jorion uses the average return of the minimum variance portfolio as the shrinkage target Y_0 and finds that robust estimation with this target consistently outperforms the sample mean. In this paper, we apply this recommended Y_0.

In general, Bayes-Stein estimation is on the form

$$\widehat{\mu}^{BS} = (1-\widehat{w})\overline{Y} + \widehat{w} Y_0 1_N \tag{4}$$

with a shrinkage intensity $\widehat{w} \in [0, 1]$. In empirical applications, Jorion suggests using the following shrinkage intensity:

$$\widehat{w} = \frac{N+2}{(N+2) + (\overline{Y} - Y_0 1_N)^T \widehat{\Sigma}^{-1} (\overline{Y} - Y_0 1_N)} \tag{5}$$

where 1_N is a vector of ones and $\widehat{\Sigma} = \frac{T-1}{T-N-2}$.

A point of note is that the shrinkage intensity $\widehat{w}$ tends to decrease with increased sample size because expected returns are then more accurately estimated by the sample estimator. Conversely, this implies that Bayes-Stein estimation is especially effective in the case of relatively few observations.

3.2.2 Robust Estimation of Variance and Skewness

Single-target shrinkage estimation of the variance-covariance- and skewness-coskewness matrix was defined by Martellini and Ziemann (2010). Using skewness estimation as an illustration (variance is conceptually similar), the method involves calculating the sample matrix $\widehat{\Phi}$, choosing a suitable target matrix $\widehat{T}$, and weighing the two components together. Single-target shrinkage estimation is on the form

$$\widehat{\Phi}^{ST} = (1-\lambda)\widehat{\Phi} + \lambda\widehat{T} \tag{6}$$

with a shrinkage intensity $\lambda \in [0,1]$. In practical applications, one needs to consider two questions: (1) what is an appropriate target matrix $\widehat{T}$, and (2) what is the optimal shrinkage intensity λ?

Among the most prominent target choices in the literature are the latent 1-factor model by Simaan (1993); the 1-factor model Martellini and Ziemann (2010); adjusted constant correlation, independent marginals, independent and identical marginals, and central-symmetric matrix by Boudt et al. (2017). It is not immediately clear what the best choice of $\widehat{T}$ is for any given situation, but contemporary research suggests that all models will decrease estimation errors significantly, especially for the skewness-coskewness matrix. As an example, Boudt et al. finds that in some applications, shrinkage estimation with the adjusted constant correlation target decreases estimation errors by up to 37% for skewness-coskewness estimates.

Regarding question (2): the optimal shrinkage intensity is dependent on choice of target matrix. The general idea is to use the relationship between the asymptotic variance for the sample estimator (π), mis-specification of the structured target estimator (ρ) and the asymptotic covariance between the sample estimator and target estimator (γ). The optimal shrinkage intensity is given by

$$\lambda = \frac{1}{T}\frac{\pi-\rho}{\gamma} \tag{7}$$

A point of note is that the shrinkage intensity is heavily dependent on the bias of the sample estimators. An overestimation of the bias tends to yield a large λ which puts more weight on the less biased target matrix; it is not entirely uncommon to observe λ close to 100% (Boudt et al., 2017). Additionally - and similarly to Bayes-Stein estimation - when then number of observations is large, λ tends to be smaller as the sample estimates naturally have less estimation error. We will not focus further on the derivation of single-target shrinkage estimation and refer to e.g. Martellini and Ziemann (2010) or Boudt et al. (2017) for more details on the technique.

Shrinkage estimation has been further generalized to Multi-Target Shrinkage (MTS), with Bartz et al. (2014) as one of the first contributors. As the name suggests, the method uses several target matrices $\widehat{T}$ to shrink the sample-estimates towards and is thus a straightforward extension of the single-target case. This becomes apparent when considering the general form of MTS:

$$\widehat{\Phi}^{MT} = \left(1 - \sum_{m=1}^{M} \lambda_m\right)\widehat{\Phi} + \sum_{m=1}^{M} \lambda_m \widehat{T}_m \tag{8}$$

where M is the number of target matrices and $\sum \lambda_m \leq 1$ and $\lambda_m \geq 0$. Compared to simple-target shrinkage, multi-target shrinkage creates even more robust estimates, especially in a portfolio optimization setting (Boudt et al., 2016). Some of the best estimates (i.e. yielding the smallest estimation error) can be produced by using 5 of the 6 goals mentioned earlier, leaving out Simaan's (1993) latent 1-factor model.

With these considerations in mind, our approach for estimating robust moments is the following: For the skewness-coskewness matrix, we use multi-target estimation with the 1-factor, adjusted constant correlation, independent marginals, independent and identical marginals, and the central symmetric matrix as targets. For the variance-covariance matrix we leave out the central symmetric target as the method is only defined for skewness-coskewness. The second and third moment thus use slightly different robust estimators, but in practical terms both robust estimators will yield better results than their sample counterparts.

An important note for the Martellini-Ziemann 1-factor target matrix has to be made. To use this target, an observed factor (vector of returns) which correlates with the assets under investigation (assets trading on OSE) has to be supplied. Stock indices are typically used for this purpose in empirical applications. We choose returns from the Oslo Stock Exchange Benchmark Index as the required factor as they should correlate well with our selection of stocks from the same market.

3.3 Constructing the MVS Portfolio

To obtain optimal portfolio weights for each out-of-sample period we need to solve a multi-objective optimization problem where the first and third portfolio moments (mean and unstandardized skewness) are maximized while the second (variance) is minimized. We employ Polynomial Goal Programming which was introduced by Lai (1991) for use in portfolio selection, a technique which has seen broad acceptance and use in empirical applications (e.g. Chunhachinda et al.,1997; Prakash et al., 2003). We use definitions and notation found in a later paper by Lai et al. (2006) where the method is further refined. In this paper they extend the PGP-model up to the fourth moment, but we exclude kurtosis and only consider the third moment (unstandardized skewness) in addition to mean and variance for reasons discussed in the introduction. As their framework is a generalized result and the optimization model applies for an arbitrary amount of moments, this difference in specification is therefore a trivial change.

3.3.1 Model Assumptions

In order to implement the optimization technique, our framework has to be well-defined. We base our assumptions on papers by Lai (1991) and Chunhachinda et al. (1997). They can be summarized as follows

1. Investors are risk-averse individuals who maximize the expected utility of their end-of-period wealth.
2. There are N assets and no risk-free asset.
3. All assets are marketable, perfectly divisible, and have limited liability.
4. The capital market is perfect with no taxes.
5. Transaction costs exist when buying or selling an asset.
6. Short-sale of assets is not allowed.

These assumptions differ slightly from the two previously mentioned papers. Firstly, we have included transaction costs which incur whenever one or more assets are traded. If transaction costs are not included in the optimization model itself but were instead applied only in the out-of-sample testing, we may find potentially unrealistic solutions: Intuitively, without punishment from transaction costs, it could be reasonable for the optimization procedure to make great changes of the portfolio weights from one holding period to the next, where in practice this would be undesirable (where transaction costs exist). A further simplifying assumption is that transaction costs are a constant share of the total transaction amount; we include no minimum fee for transactions which are often present in practice. For stock trading at OSE, relevant transaction costs are 0.05% of the total transaction amount, based on quotes from two major trading platforms in Norway (DNB, 2019; Nordnet, 2019). As a second note, we are not concerned with the tradeoff between wealth allocation in risky and non-risky assets in the PGP optimization itself: we only care about creating an optimal risky portfolio allocation. We exclude a risk-free asset for simplicity.

3.3.2 Portfolio Optimization with PGP

The PGP procedure consists of two steps. In the first step, one finds the portfolio weights that yield the best possible value for each of the three portfolio moments separately. To describe the model, we introduce the following mathematical notation. $X = (x_1, x_2, ..., x_N)^T$ is the portfolio weights where x_i is the percentage of wealth invested in the ith risky asset. $R = (R_1, R_2, \dots, R_n)^T$ is monthly asset returns where R_i is the monthly rate of return on the ith asset for the months 1, 2, ... , 168. $\bar{R} = (\bar{R}_1, \bar{R}_2, \dots, \bar{R}_N)^T$ is the mean asset returns, where $\bar{R}_i$ is the mean return of the ith asset. V is the $N \times N$ variance-covariance matrix of the asset returns. S is the $N \times N^2$ skewness-coskewness matrix of the asset returns.

Based on this formulation, the three first portfolio moments are defined as

Portfolio mean $\equiv R_p = X^T \bar{R}$.

Portfolio variance $\equiv V_p = X^T V X$.

Portfolio third moment $\equiv S_p = X^T S (X \otimes X)$.

where $\otimes$ denotes the kronecker product. Step 1 of the PGP procedure can then be summarized in the following three subproblems (SPs) where each moment is optimized without taking the others into account where we obtain the solution as the optimal portfolio moments R_p^*, V_p^* and S_p^*.

$$SP(1) = \begin{cases} \text{Maximize} & R_p = X^T \bar{R} \\ \text{subject to} & X^T 1_{\text{N}} = 1 \\ & X \geq 0 \end{cases} \tag{9}$$

$$SP(2) = \begin{cases} \text{Minimize} & V_p = X^T V X \\ \text{subject to} & X^T 1_{\text{N}} = 1 \\ & X \geq 0 \end{cases} \tag{10}$$

$$SP(3) = \begin{cases} \text{Maximize} & S_p = X^T S (X \otimes X) \\ \text{subject to} & X^T 1_{\text{N}} = 1 \\ & X \geq 0 \end{cases} \tag{11}$$

In the second step of PGP we find the portfolio weights that maximize the mean and third moment, and minimize the variance simultaneously. The method now introduces three distance variables: d_1, d_2 and d_3. The solutions from subproblems 1, 2 and 3 are the "best case scenario" for the three moments, and since the distance variables are assumed to always be greater than zero, they represent underachievement of the multi-goal-optimization in relation to the optimal values for each of the three moments (R_p^*, V_p^* and S_p^*). The goal of the second step is to minimize this underperformance, adjusted for the investors preference for mean, variance and skewness, denoted by λ_1, λ_2 and λ_3 respectively where $\lambda_i \geq 0$. A value of $(\lambda_1, \lambda_2, \lambda_3) = (1,\ 1,\ 0)$ defines the mean-variance portfolio, $(0,\ 1,\ 0)$ the global minimum variance portfolio, for instance. Optimal moment values, distance variables and investor preference are integrated in an objective function by applying the Minkowski distance which is given by

$$Z = \sum_{k=1}^{m} \left| \frac{d_k}{A_k} \right|^{\lambda_k} \tag{12}$$

where A_k are the optimal moments from step 1.

Now we expand the optimization model of Lai et al. (2006) by including transaction costs in the following way:

$$TC_t = (|X_t - X'_{t-h}| \times 0.0005)^T 1_N \tag{13}$$

where X_t are the optimal weights at the start of a holding period at time t and X'_{t-h} are the normalized optimal weights from the previous holding period

$$X'_{t-h} = \frac{X_{t-h} \times \prod_{i=t-h}^{t} (1 + R_i)}{(X_{t-h} \times \prod_{i=t-h}^{t} (1 + R_i))^T 1_N}. \tag{14}$$

We include transaction costs (TC) in the first restriction of the optimization problem where deviation from the optimal mean portfolio return R_p^* is determined. It should be noted that our implementation of transaction costs still allows big portfolio rebalances from one period to the next because we chose a 'weak' restriction where the optimization is merely punished for rebalancing. In that sense, the optimization problem may still find 'impractical' solutions for real life, but we have ensured model parity between different specifications (equally weighted, GMV, MV and MVS) and applied a reasonable transaction cost to limit extreme turnover.

Finally, optimization problem 2 in the PGP procedure is on the following form

$$P(2) = \begin{cases} \text{Minimize} & \left|\frac{d_1}{R_p^*}\right|^{\lambda_1} + \left|\frac{d_2}{V_p^*}\right|^{\lambda_2} + \left|\frac{d_3}{S_p^*}\right|^{\lambda_3} \\ \text{subject to} & X^T\bar{R} - TC + d_1 = R_p^* \\ & X^T V X - d_2 = V_p^* \\ & X^T S(X \otimes X) + d_3 = S_p^* \\ & X^T 1_{\text{N}} = 1 \\ & X \geq 0 \\ & d_1, d_2, d_3 \geq 0 \end{cases} \tag{15}$$

which yields the optimal weights X* for each out-of-sample period. The PGP procedure is done for all 16 sub-periods in our dataset, creating 16 sets of optimal portfolio weights.

3.4 Performance Criteria

Our performance criteria will be the Sharpe Ratio (SR) and Adjusted for Skewness Sharpe Ratio (ASSR), introduced by Zakamouline and Koekebakker (2009). The SR is a well established performance criterion, both in literature and in practice (Sharpe, 1994). However, SR can be problematic to use if the distribution of return is non-normal. Empirical evidence, both from the literature and recent data retrieved from OSE, support that the distribution of return are not normal (see, for example Brooks and Kat, 2002). On the contrary, there has also been evidence in favor of SR, a claim that although the underlying assumption of SR is somewhat violated, the performance criterion is still sufficient to rank performance on equal footing with performance criterias that do take non-normality into account (see for example Eling and Schuhmacher, 2007). Eling and Schuhmacher find that the Sharpe Ratio ranked 2,763 hedge funds virtually the same as 12 other performance measures, where some of them directly dealt with non-normality by adjusting for higher moments like skewness and kurtosis. The formula for SR is given by:

$$SR = \frac{E[R_a - R_f]}{\sigma_a} = \frac{E[R_a - R_f]}{\sqrt{var[R_a - R_f]}} \tag{16}$$

where R is return, a is the underlying asset (or portfolio of assets) and f is the risk-free component.

We also include a performance criterion, the ASSR, that can evaluate portfolio performance adjusted for skewness. This way we can compare the MVS with our benchmark portfolios with a performance criterion for both scenarios of normal and non-normal investor preferences. ASSR is SR adjusted for portfolio skewness, where a positive (negative) skewness will lead to a higher (lower) ASSR value. This means positive skewness will be rewarded, and vice versa, negative skewness will be punished. The formula for ASSR is given by

$$ASSR = SR\sqrt{1 + b_3 \frac{Sk}{3} SR} \tag{17}$$

where Sk is the portfolio skewness and b_3 is the investor preference for skewness. It is expressed as

$$b_3 = \frac{\rho + 1}{\rho} \tag{18}$$

where ρ is the investor's degree of risk aversion. A lower risk aversion yields a higher demand for positive skewness and a stronger dislike of negative skewness ceteris paribus. Zakamouline and Koekebakker (2009) argue for different levels of this parameter, and conclude that a high level of risk aversion and thus $b_3 = 1$ may be reasonable for a representative investor. Note that preference for skewness in ASSR is different from the investor preference in the optimization problem, λ_3.

3.5 Measuring Portfolio Concentration

As explained in the literature review, empirical evidence suggests that diversifying a portfolio may lead to lower portfolio skewness, and conversely, that portfolios with higher skewness should be less diversified. Thus it may be of interest to characterize the asset allocation of the different strategies by calculating portfolio concentration. It should be noted that diversification in the broad sense encompasses a variety of methods and performance measures, and the efficacy of diversification relies on factors such as volatility and correlations of assets. Our goal is to characterize the strategies by concentration, assessing what's often referred to as 'weight diversification' (Richard and Roncalli, 2015). We use the Diversification Index (DI) for this purpose which is defined as the complement of the concentration, a common practice in the Industrial Organization literature (see e.g. Woerheide and Persson, 1993). The (weight) diversification index is defined as

$$DI \;=\; 1 - \sum_{i=1}^{N} X_i^2 \tag{19}$$

where X_i is the portfolio weight of asset i. DI ranges from 0 to (1 - $\frac{1}{N}$), where a higher score signifies a portfolio weight lower concentration (higher weight-diversification). The EW portfolio where total wealth is perfectly spread among all assets will by definition represent the lowest amount of concentration possible and yields the highest DI. A portfolio consisting of a single stock (100% weight) will naturally have the minimum possible DI of 0.

4 Data

We consider a dataset of 80 randomly chosen stocks trading at OSE over the period 2004 to 2017. The data are used for two different purposes: For the descriptive analysis, all 80 asset returns are examined to gain insight into the properties of the Norwegian stock market. The data for the 80 stocks are retrieved from 'Børsprosjektet NHH', an initiative at the Norwegian School of Economics which gets returns directly from OSE market data. Of the 80 firms, 21 operate in the Industrial sector, 19 in Finance, 15 in Energy, 9 in Information Technology, and 16 Other.[3] The stock returns are monthly observations, adjusted for dividends and special events (such as stock-splits) for consistency. Missing data due to e.g. stocks not being traded the entire period account for $< 1\%$ of observations. This is handled by inserting a value of 0, representing no return for the investor for the relevant period.

Due to the fact that skewness is sensitive to portfolio size (see e.g. Beedles and Simkowitz, 1978; Beedles, 1979), we include different specifications for consideration in portfolio optimization. In this paper we investigate portfolios which contain 5 and 15 assets. Ideally, we would like a broader range of configurations but we find computing power to be a limiting factor. The two sets of stocks are randomly chosen. The 15 stocks are a subset of the 80; the 5 are a subset of the 15. Despite not including a larger selection of stocks for asset allocation, it may be enough to gain some qualitative insights on the role of skewness in portfolio management. An overview of the relevant stocks used in optimization is provided in table 1.

As mentioned earlier, we use robust moment estimates. This requires estimation of a 1-factor model where an observed factor needs to be provided. For this purpose, we use returns from Oslo Stock Exchange Benchmark Index (OSEBX). We retrieve the daily divided-adjusted index price from the Oslo Børs website (Oslo Børs, 2019) and manually calculate monthly returns.

[3]The 'Other' category include Real estate, Health care, Consumer staples, Communication services and Materials.

Table 1: Overview of Stocks Used in Optimization

Asset name	Sector	Return	Standard deviation	Skewness
DNO	Energy	0.0263	0.1749	1.0028
Ekornes	Consumer Discretionary	0.0083	0.0728	-0.0316
Fred. Olsen Energy	Energy	0.0150	0.1779	2.8418
Jinhui Shipping and Transportation*	Industrials	0.0193	0.1859	0.5885
Kongsberg Gruppen	Industrials	0.0154	0.0764	0.3656
Lerøy Seafood*	Consumer Staples	0.0230	0.0990	0.0583
Norske Skogindustrier	Materials	-0.0117	0.2164	1.1497
Orkla	Consumer Staples	0.0141	0.0744	-0.5164
Petroleum Geo-Services*	Energy	0.0081	0.1353	-0.5026
SAS AB	Industrials	0.0018	0.1643	1.3745
Sparebank 1 SMN*	Finance	0.0146	0.0729	-0.1899
Statoil	Energy	0.0109	0.0648	0.1630
Subsea 7*	Energy	0.0205	0.1112	-0.1916
Tomra Systems	Industrials	0.0127	0.0884	-0.0449
Veidekke	Industrials	0.0200	0.0832	0.1526

Numbers are monthly figures. Data are from 2004-2017. All assets in the table are used in the 15-stock optimization; assets with stars (*) are used in 5-stock optimization.

5 Results

In this section we present the results of our analysis. Consistent with the research questions, there are two main points of consideration. First we recount our findings on the existence and persistence of skewness for assets at OSE. Following this is a look at how portfolio size affects portfolios in relation to skewness. These three points are addressed in sections 5.1.1, 5.1.2 and 5.1.3 respectively. The findings provide important context and implications for portfolio management with higher moments. Regarding the second research question, we present the results from the asset allocation procedure with the MVS portfolio in focus. The goal is to see whether the MVS portfolio is able to outperform the traditional MV portfolio at OSE by assessing end-of-period wealth, portfolio moments, and a classical performance measure (Sharpe Ratio) during the out-of-sample period spanning 2014-2017. These results are found in section 5.2.

5.1 Skewness of Stock returns at Oslo Stock Exchange

5.1.1 Existence of Skewness in Asset Returns

We now provide insights into the distributional properties of individual assets at OSE and discuss some of the implications. We consider development of OSE over time using all 80 stocks of our dataset. Sample skewness is calculated using 4-year rolling estimates from 2006 to 2017, leaving out the first two years of our dataset. Skewness for the entire 12-year period is included as well.

Table 2: Percent of Assets with Skewness in Different Periods

Time period	2006-09	2010-13	2014-17	2006-17
Positive Skew	40	49	68	58
Negative Skew	31	8	11	15
No Skew	29	43	21	27

All entries in percent.

For the 80 asset returns, 58% had positive skew, 15% negative, and 27% no significant skewness using data from the period 2006 to 2017. The fact that there are roughly twice as many assets with positive skewness than negative is a significant discovery, which is also in line with comparable numbers in the US stock market (Singleton and Wingender, 1986). Concerning the three sub-periods, we note that the proportion of positively skewed stocks from period 1 to 3 grows from 40% to 68%, and the proportion of negatively skewed shrinks from period 1 to 3 from 31% to 11%. There may be several explanations for why we observe a preponderance of assets with positive skewness rather than negative. One is survivorship bias: the fact that our stock selection only include firms that survived for the entire 2004-2017 period means that firms which went bankrupt (and may in advance have experience extreme negative losses) have been excluded from our sample.

As expected, the period where the market exhibited the lowest amount of positive skewness and the highest amount of negative was around the peak of the financial crisis (2008). The economic story is quite clear: when the market as a whole is experiencing great sustained devaluation, we would then expect to see this reflected in individual stocks as more extreme losses. The Norwegian stock market also had a smaller crisis around 2015 caused by the sharp decline in the oil price, but curiously this is not reflected in the skewness: moving from the second to the third 4-year period, the market share of positively and negatively skewed stocks both grew (the share of negative only slightly).

To follow up on the investigation on distributional properties of stocks at OSE, we confirm that returns are not normally distributed by performing the Shapiro-Wilk test (Shapiro and Wilk, 1965). The null hypothesis of the test asserts that a given distribution is normal, while the alternative hypothesis specifies deviation from normality. Based on the test results (not included in the paper for brevity), we note three things: First, a total of 89% of stocks reject the null of normality at the 5% significance level. Secondly, all assets with positive or negative skewness ($|\hat{s}| > 0.3$) reject the null hypothesis. Thirdly, some assets reject the null hypothesis despite having no significant skewness. The first point establish the main finding, namely that the overwhelming majority of individual assets returns at OSE are not normally distributed. The second and third point together reveal another result: skewness is a sufficient but not necessary condition for non-normality. I.e. if a stock has skewed returns, it is not normal, but it can also be non-normal without significant skewness. Several causes can be proposed: (1) the rule-of-thumb value of 0.3 is not an accurate measure of skewness, (2) there are impacts of kurtosis or even higher moments which skewness does not inform us of, (3) outliers in the data impact the Shapiro-Wilk test more than the measure of sample skewness, and (4) like any statistical test, the Shapiro-Wilk test is prone to being overly-sensitive when the sample size is large which causes rejection of normality even when the distribution is reasonably normal. At any rate, we conclude that asset returns on OSE are generally non-normal, and some subset of these are non-normal specifically because of the presence of skewness.

It is useful to further characterize stocks at OSE by analyzing skewness by sector. This gives a more detailed view of the exchange. Table 3 shows average 4-year rolling skewness for individual assets over time by 4 sectors.

Table 3: Average Skewness by Sector

	Industrials	Finance	Energy	Information Technology	Others	All
2006-2009	0.0892	0.0602	-0.3661	0.8771	0.4551	0.1588
2010-2013	0.6899	0.4356	0.2844	0.7858	0.3806	0.5024
2014-2017	1.1269	0.1825	0.7306	1.5688	0.7228	0.7972

Number of assets: Industrials = 21, Finance = 19, Energy = 15, Information Technology = 9, Others = 16, All = 80. The 'Others' category include Real estate, Health care, Consumer staples, Communication services and Materials.

As is apparent from the table, average skewness of all 80 stocks have increased significantly over time. This is consistent with table 2 which states that the amount of assets with positive skewness has increased as well. The rising trend is reflected in all sectors. All sectors have exhibited positive average skewness for the entire time frame except for the energy sector in the years 2006-2009. Possibly, this sector experienced more extreme negative returns during the financial crisis than the others. Industrials, Finance and Energy have low average skewness during the financial crisis, with Information technology and 'Others' as an exception. Industrials and Information Technology have the highest average skewness overall; technology firms are particularly famous for experiencing rapid or extreme growth, contributing to high positive skewness.

We now have an indication that it is possible to construct positively skewed portfolio returns as roughly 40% to 70% of stocks have positive skewness over the period 2006-2017. At the very least there exists trivial portfolios consisting of a single stock with positive skewness. A description of individual assets, however, does not adequately describe the portfolio management setting at OSE. One has to be mindful of the synergistic nature of a portfolio, i.e. how the attributes of a portfolio of stocks differ from the individual stocks themselves. In particular, we find that the effect of portfolio diversification is significant. This is explored in section 5.1.3. First, we examine whether future skewness can be predicted from past skewness.

5.1.2 Persistence of Skewness

Empirically it is well established that the assumption of constant mean and variance does not always hold, especially in turbulent periods such as financial crises; times when it is the most important for a given framework to be robust. We extend this train of thought to consider persistence of skewness, similar to the analysis by Singleton and Wingender (1986). We calculate rolling 4-year-estimates of sample skewness over our 14-year time period (leaving out the first two years) for all 80 stocks and note whether the asset has the same type of skewness (positive or negative) in two adjacent sub-periods.

Table 4: Percent of Assets with Persistent Skewness in Adjacent Periods

Time period	1-2	2-3	1-2-3
Positive Skew	56	69	6
Negative Skew	12	0	0

Periods: 2006-2009 = 1, 2010-2013 = 2, 2014-2017 = 3. All entries in percent.

The results reveal moderate levels of skewness persistence for individual assets: 56% have positive skewness in both period 1 and 2, increasing to 69% in period 2 and 3. In other words: if one were to predict positive skewness in the future, it is reasonable to base oneself on past skewness. Although a sizeable portion of stocks aren't persistent, at the very least one could reasonably expect to beat a coin-toss. For negative skewness the conclusion is different. Only 12% of stocks showed persistence from the first to the second 4-year period, dropping to 0% for the second and third. This is a strikingly low persistence, implying that negative skewness may be incredibly difficult to predict. Focusing on all three periods, barely any stocks show persistence of any kind. Only 6% of stock returns with positive skewness in period 1 kept their skewness throughout three consecutive 4-year periods, while none kept negative skewness for the same three periods.

Overall then, we find that skewness of stocks at OSE is fairly persistent for assets with positive skewness, but not with negative. It should be noted that this finding may be sensitive to the length of the estimation period for skewness, and our conclusion holds only for the long term (4 year periods). There is no guarantee of more or less persistence in the short term. The prediction horizon is important to keep in mind for the portfolio selection done in this paper: we predict skewness 1-3 months into the future based on data from the preceding 120 months. Proving a moderate persistence in positive skewness in the long term bodes well for investors with a preference for skewness.

5.1.3 Effect of Portfolio Size on Portfolio Skewness

As mentioned earlier, diversification has implications on portfolio management with skewness. It is a well known result that the risk (standard deviation) of a portfolio falls with increased diversification. Regarding skewness, previous research indicate that increasing diversification also causes a reduction in portfolio skewness (e.g. Beedles, 1979; Singleton and Wingender, 1986). To investigate the claim that skewness is reduced with increased diversification in our dataset, we follow Beedles' research (1979). Beedles uses portfolio size as a naïve approach to diversification. As such, the definition (which we adopt for use in this section) is slightly restrictive and does not encompass all facets of the term in the classical sense. Nevertheless, we are interested in an 'all else equal' effect on skewness, which Beedles argues for. The method is outlined by the following:

1. Sort all individual assets from lowest to highest risk
2. Choose a portfolio size of m assets
3. Create an equally weighted portfolio of the m first assets, then the subsequent m assets and so on
4. Compute skewness for each portfolio[4]
5. Report the average of the aforementioned statistic, for each m

We choose m = 1, 5, 10, 20, 40 and 80. Statistics for portfolio mean and standard deviation are included for completeness. Note that the mean is the same for all portfolio sizes; it is mathematically guaranteed with equally weighted portfolios.

[4]Beedles presented the cube root of the third moment in his paper while we calculate sample skewness. Although the magnitude of the numbers are different, the qualitative result remains the same. We report sample skewness for consistency in this paper.

Table 5: Effect of Diversification on Portfolio Moments

Portfolio size	#Portfolios	Mean (%)	SD (%)	Skewness
1	80	1.10	12.11	0.922
5	16	1.10	7.05	0.053
10	8	1.10	6.07	-0.381
20	4	1.10	5.52	-0.618
40	2	1.10	5.22	-0.718
80	1	1.10	5.00	-0.879

Data is based on monthly returns of 80 assets from the period 2004-2017.

For single stocks the average skewness is positive at 0.922, while increasingly diversified portfolios have much lower - even negative - skewness. At the 80 stocks equally-weighted portfolio, we calculate a skewness of -0.879. This suggests that unsystematic (diversifiable) skewness is quickly diversified away, and at 20 stocks it seems further diversification has a small impact. We also confirm the well known fact that portfolio risk rapidly decreases with portfolio size; most of the relevant diversification is achieved at 10 stocks. After this point, additional diversification has diminishing returns. These results show that risk and skewness are in fact goals that point in opposite directions: desire of low risk incentivizes the portfolio manager to diversify, but seeking highly positively skewed returns should ideally lead to undiversified portfolios.

The (under)diversification effect of skewness can also be explained from a purely analytical standpoint. Using robust moment estimators, we find that the vast majority of coskewness-terms are negative. Similar to how positive covariance terms increases portfolio variance with diversification, having negative coskewness-terms will cause portfolio skewness to decrease with diversification. This observation is consistent with an analysis by Albuquerque (2012) where the author finds that while individual stocks may exhibit positive skewness, the market skewness (a portfolio of all stocks) is almost always negative. As an addition to this point, we measure the skewness of the OSEBX to be -1.25 using data for the period 2004-2017.

The investor with a preference for positive skewness should therefore abstain from passive index investment to attain their desired portfolio (although there are other reasons to invest passively). This motivates an inquiry into optimal asset allocation with differently sized portfolios which we have done by using 5 and 15 assets.

5.2 Asset Allocation with Skewness

We perform PGP optimization on the assets from OSE to form optimal portfolio strategies. The implementation will be briefly explained and then the results are presented for the different portfolios.

Implementation of the EW-portfolio strategy is straightforward: in each rebalancing period we simply reallocate resources so each stock has equally as much weight in the portfolio. The GMV portfolio is constructed using PGP by setting the investor preference for mean and skewness equal to zero. More specifically we apply $(\lambda_1, \lambda_2, \lambda_3) = (0,\ 1,\ 0)$ with λ_1 being preference for expected return, λ_2 variance, and λ_3 the third moment. The MV portfolio is constructed by setting the preference for skewness equal to zero: (1, 1, 0). Similarly, the MVS portfolio is defined by equal preference for the three moments: (1, 1, 1). In addition to these portfolios we also include a model specification were the investor has a greater preference for skewness. This provides insight into how different skewness preferences may affect the performance of the portfolio allocation. The Modified MVS (MMVS) portfolio is given by (1, 1, 3). Note that the EW and GMV strategies are strictly speaking reference portfolios and the most pertinent comparison is between MV, MVS and MMVS when assessing the efficacy of skewness in portfolio management.

To characterize the outcomes of the strategies, we will in section 5.2.3 include a consideration of the performance measures mentioned in the methodology section, namely Sharpe Ratio and Adjusted for Skewness Sharpe Ratio. Of note is the fact that we annualize in-sample and out-of-sample moments to more clearly differentiate the portfolios while the performance measures is based on monthly data to allow for better statistical inference.

As mentioned earlier, there are two different scenarios under consideration: one where the investor has 15 assets available, and one where they have 5. We first present the 15-asset scenario, then the 5-asset one.

5.2.1 In-sample Results

15 stocks

Before commenting on the out-of-sample results we first present the in-sample solutions. In-sample moments are obtained when solving step 2 of the PGP procedure for the different specifications listed above.[5] A table summarizing the average values of the in-sample moments is presented below. It should be noted that in-sample, the PGP optimized for the third moment and not skewness (standardized third moment). We report this measure in the table to accurately represent the solution of the PGP procedure according to what is optimized.

Table 6: Average Value of Robust In-Sample Moments. 15 stocks

Strategy	Return	Standard deviation	3rd moment (x1000)
MV	0.1135	0.1785	-0.1445
MVS	0.1129	0.1793	-0.1332
MMVS	0.1117	0.1766	-0.1284
GMV	0.1098	0.1711	-0.1372
EW	0.1080	0.1785	-0.1282

All numbers are annualized.

[5] All in-sample moments mentioned in this section are robust estimates; out-of-sample moments are sample estimates. Robust moments do not have the same economic interpretation as sample moments and even tiny differences can be significant.

We find that the GMV portfolio predictably yield the lowest average in-sample standard deviation (17.11%) and also second lowest return (10.98%). The MV portfolio is characterized by the highest average return of 11.35%, but also the second highest average standard deviation of 17.85%. The explanation is straightforward: to get high returns, one may need to also take risks. The MVS portfolio attains the third highest average in-sample third moment of 0.1332, but curiously also the highest average standard deviation of 17.93%. The MMVS portfolio has the second highest in-sample third moment at -0.1284 and lower standard deviation (17.66%) than both MV and MVS. Finally, EW performs the worst in regards to return (10.80%) and has similar (high) standard deviation (17.85%) to the MV portfolio. Remarkably, the EW portfolio has the highest third moment of all the portfolio strategies (-0.1282), even higher than MMVS. Although one would expect MMVS to attain better third moment, it needs to be emphasized that the strategy also prioritizes other goals where the EW clearly falls short. Thus the MVS can be said to have better 'return-risk' adjusted third moment. Ultimately we see that the MVS portfolio(s) accept lower return, but also manages to acquire a higher third moment compared to the MV portfolio. Later, in section 5.2.3, we use performance measures to evaluate the portfolio strategies and in particular the MVS portfolio(s). The goal is to see if they make a significant trade of the first two moments in order to acquire a higher third moment.

5 stocks

For the results on the 5 stock-scenario, we again start with a comment on in-sample moments. The average annual values of in-sample moments are enclosed in table 7.

Table 7: Average Value of Robust In-Sample Moments. 5 stocks

Strategy	Return	Standard deviation	3rd moment (x1000)
MV	0.1640	0.2511	-0.0613
MVS	0.1639	0.2525	-0.0598
MMVS	0.1639	0.2545	-0.0571
GMV	0.1639	0.2511	-0.0615
EW	0.1635	0.2568	-0.0593

All numbers are annualized.

Of note is the fact that average annual return for all strategies are very similar - around 16.4%. The MV portfolio manages a marginally better return than MVS, MMVS and GMV. EW performs the worst in this regard. The reason for the nearly identical returns between the strategies is twofold: First, robust estimators create very small difference in expected return for the 5 assets due to shrinkage toward a common goal with a big shrinkage intensity. Secondly, one asset is heavily weighted ($> 50\%$) in every portfolio strategy due to being better than all other assets in all three moments. Assessing risk, the EW strategy produces the highest annual average standard deviation at 25.68%, followed by the MMVS portfolio at 25.45%. With an equal prioritization of the three moments and less emphasis on the third moment, the MVS strategy manages a slightly lower standard deviation at 25.25%. MV and GMV have the lowest risk and performed similarly, both with an equal standard deviation of 25.11%. With the third moment, differences are even more pronounced. MMVS obtains the highest value of -0.0571. Like in the 15 stock scenario, EW performs very well in-sample with the third moment at -0.0593, only beat by the MMVS. MVS performs similarly to EW at -0.0598. Despite this fact, MVS still outperforms EW in the two other aspects, yielding better overall performance. MV and GMV achieve almost identical third moments of -0.0613 and -0.0615 respectively, emphasizing how similar the outcomes of these two strategies are.

Ultimately, in-sample results give insight into the PGP optimization results, and how the solutions are characterized. The results show the predictions for the portfolio moments the model makes for the unknown out-of-sample period. For 15 stocks, the procedure is able to formulate strategies with clear distinctions. For 5 stocks, the robust portfolio returns are very similar, and strategies are only able to be distinguished based on risk and the third moment. Thus, the in-sample results for 15 stocks give clear expectations for the out-of-sample results, while it is more ambiguous for 5 stocks. This helps explain the out-of-sample results found in the next section.

5.2.2 Out-of-Sample Results

15 stocks

To review out-of-sample performance of the different strategies, we first assess the cumulative return to see how much end-of-period wealth the investor achieves. Returns are calculated by using the weights found by using the PGP procedure where transaction costs are subtracted when rebalancing the portfolio each quarter.

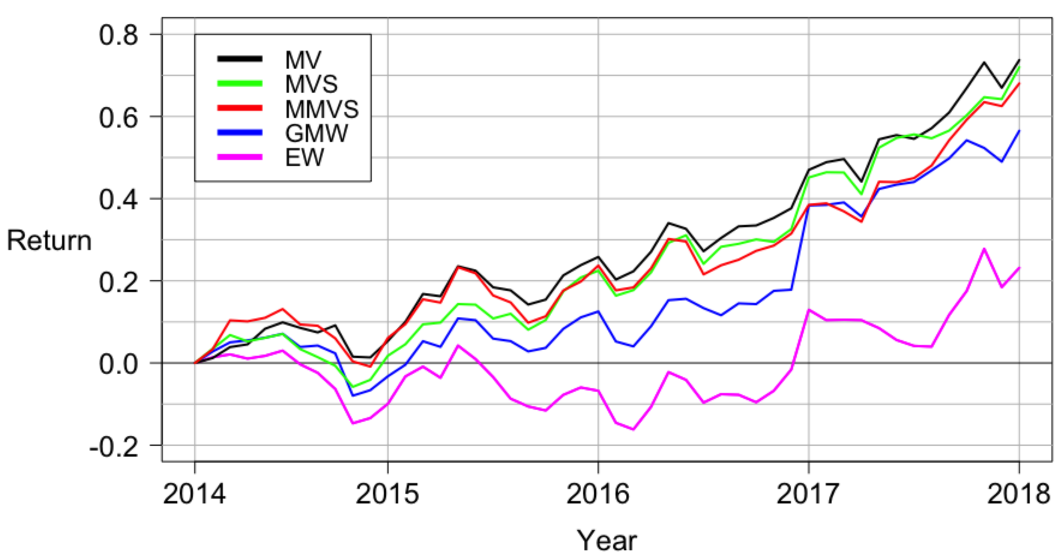

Figure 3: Monthly Cumulative Out-of-Sample Returns. 15 Stock Portfolios.

The MV and MVS portfolio end up with the highest cumulative return, achieving 74% and 72% respectively. Despite the similar outcomes at the end, MV consistently outperforms MVS throughout the entire 4-year period. The MMVS portfolio which puts relatively less weight on return, follows the two strategies closely at 68% end-of-period return - a 6 and 4 percentage point return deficit respectively over the 4 year out-of-sample period. Furthermore, we note that GMV performs significantly worse than the three aforementioned strategies. It ends up at 57% cumulative return and thus lags 11 percentage points behind the MMVS portfolio. Finally, the 'naïve' EW portfolio greatly underperforms, achieving only 23% cumulative return over 4 years. Notably, the strategy produces negative cumulative returns the first 3 years before improving in 2017.

To further assess the strategies, we consider mean, standard deviation and skewness in addition to a measure of weight-diversification, DI. The results are summarized in table 8.

Table 8: Out-of-Sample Results. 15 stocks

Strategy	Return	Standard deviation	Skewness	DI
MV	0.1481	0.1085	-0.0838	0.8829
MVS	0.1452	0.1123	0.0296	0.8627
MMVS	0.1385	0.1196	-0.0191	0.8614
GMV	0.1185	0.1370	0.2878	0.8828
EW	0.0535	0.1681	0.1533	0.9333

Numbers are annualized averages.

Considering average out-of-sample return, the MV and MVS portfolios perform similarly at 14.81% and 14.52% respectively, beating the other strategies. At the same time, they attain the lowest risk (standard deviation) at 10.85% and 11.23% respectively. MMVS gains a slightly lower average return (13.85%), likely due to a higher prioritization of skewness. It is also slightly more risky than the two aforementioned portfolios with an annualized standard deviation of 11.96%. Focusing on the reference portfolios, both GMV and EW generate poor returns which is also evidenced in figure 3. With 5.35% average annualized return, EW produces under half the return of GMV (11.85%). They are also the riskiest portfolios, with EW having close to 17% standard deviation compared to 11-12% for the MV, MVS and MMVS. The GMV strategy has the second highest risk at 13.70%, which is surprising since its sole purpose is to minimize risk. It turns out both EW and GMV are significantly affected by an outlier in the data - this is discussed in section 5.2.3.

Moving on to skewness, we first note that no strategy is strictly speaking significant in this regard ($\hat{s} > |0.3|$). Due to the fact that all portfolios diversify away most unsystematic skewness at 15 stocks, we are mainly left with systematic skewness where even small differences between strategies can be of interest despite no individual strategy having significant skewness on their own. While MV has a skewness of -0.0838, the MVS portfolio achieves 0.0296 - a 0.1134 difference. MMVS achieves only -0.0191 skewness which is slightly worse than MV. This contradicts what we would expect based on: (1) the MMVS portfolio has a higher preference for skewness than MVS, and (2) the portfolio manages to acquire a higher in-sample third moment. In-sample results seemingly does not persist out-of-sample. Remarkably, GMV has the highest skewness and is close to the rule-of-thumb of 0.3 at 0.2878. The EW portfolio manages a skewness of 0.1533; while not enough to be significant on its own, it is higher than the three main strategies.

Regarding the level of portfolio concentration, we find that the MVS and MMVS strategies are more concentrated than the MV portfolio because of their attempt to achieve higher skewness out-of-sample. They obtain DI values of 0.8627 and 0.8614 respectively, against 0.8829. The EW is by definition minimally concentrated with a DI of 0.9333 ($1 - \frac{1}{15}$). This observation aligns with the literature and shows one strategy to achieve higher skewness: choose more concentrated portfolios. EW and GMV seemingly counter this result by having higher average DI but also higher skewness than the other 3 strategies. But as pointed out earlier there is one outlier that greatly affects the two reference portfolios with respect to skewness which we will discuss in section 5.2.4. Nonetheless, in the 15-stock scenario, we affirm the tendency of decreased weight-diversification when adding skewness as a concern in portfolio management.

5 stocks

Now we switch focus to out-of-sample performance of the portfolio strategies. First we consider cumulative returns.

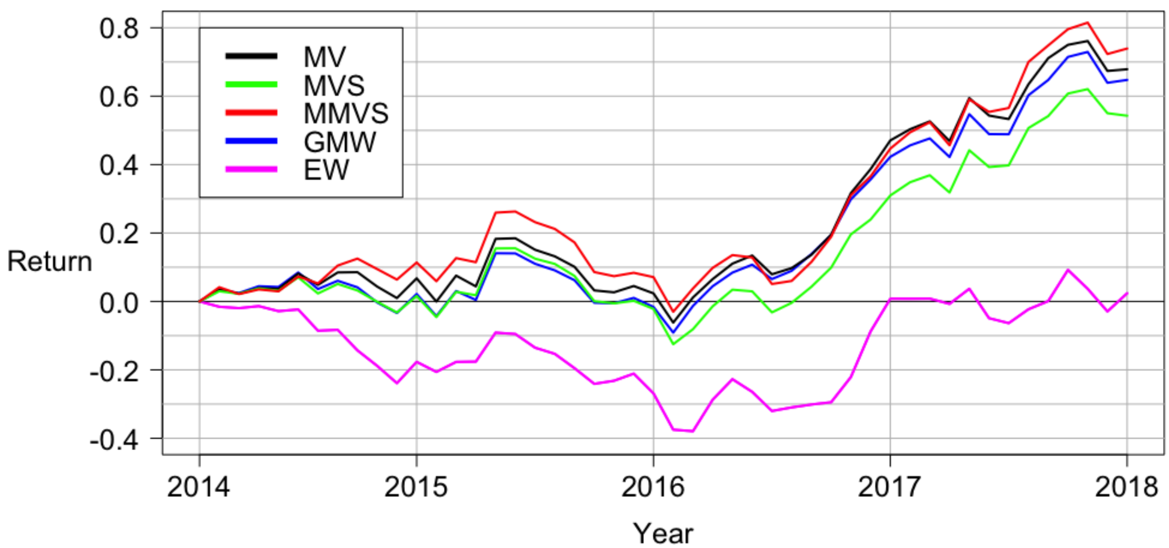

Figure 4: Monthly Cumulative Out-of-Sample Returns. 5 Stock Portfolios.

We find that the MMVS portfolio produces the highest cumulative return at 74%, followed by the MV (68%), the GMV (65%) and the MVS (54%) portfolio. This ranking is somewhat surprising because we would initially expect that the MV portfolio should be able to gain the highest return out-of-sample; still it performs reasonably well. The EW strategy yields the lowest cumulative return of 2% over the 4 years - virtually indistinguishable from 0. Like in the 15 stock scenario, EW produces negative cumulative returns in 3 of 4 years. In the graph for cumulative return we note that the 5-stock portfolios seem to perform worse than the 15 stock portfolio in the period 2014 to 2016. One contributing factor for this is that the

5 stock portfolio has a higher share of companies that are directly related to the oil industry, and were therefore negatively affected by the oil price crisis in 2015.

We will now further assess the different strategies by presenting the portfolio mean, standard deviation, skewness and the measure for portfolio concentration DI. The results are presented in table 9.

Table 9: Out-of-Sample Results. 5 stocks

Strategy	Return	Standard deviation	Skewness	DI
MV	0.1383	0.1612	0.0762	0.6296
MVS	0.1145	0.1655	0.0435	0.5781
MMVS	0.1484	0.1652	0.0370	0.5260
GMV	0.1329	0.1607	0.0912	0.6250
EW	0.0059	0.2206	0.1499	0.8000

Numbers are annualized averages.

Consistent with the results for cumulative returns, the MMVS portfolio dominates with an average annual return of 14.84% while being the third most risky with a standard deviation of 16.52%. MV also performs well: it has the second highest return of 13.83% in addition to having the second lowest standard deviation of 16.12% which is on par with GMV. The MVS portfolio performs poorly with return of 11.45% and standard deviation of 16.55%. These results make MVS worse than GMV which posts an average return of 13.29% and standard deviation of 16.07% - a surprising result given that GMV does not prioritize return in contrast to MVS. Similar to the trend observed with cumulative returns, the EW portfolio severely underperforms on other metrics as well. While having the lowest average return of 0.59%, it has by far the highest risk of 22.06%.

Regarding skewness, we find that no strategy manages to surpass the limit of 0.3 in order to achieve significant skew. We also note that no skewness is below 0, in contrast to the 15-stock case. The two reference portfolios - GMV and EW - have the highest skewness of 0.0912 and 0.1499 respectively. MV has the third highest skewness of 0.0762 with the MVS following with 0.0435. The MMVS strategy which obtains the highest third moment in-sample, has the lowest skewness out-of-sample at 0.0370. The ranking of these three strategies are the opposite of what we expect based on different investor preferences for skewness. This may be an indication that in-sample results did not predict out-of-sample results well with respect to skewness.

The DI of the different portfolios is as we would expect. MMVS has the lowest DI of 0.5260, making it the most concentrated portfolio. This is followed by MVS with an DI of 0.5781. GMV and MV are similar with 0.6250 and 0.6296 respectively. EW has an DI of 0.8000, which by definition is the highest possible value. These observations are consistent with the trend we saw with 15 stocks: the strategies with a focus on skewness (MVS, MMVS) produce more concentrated portfolio weights on average. Note that although the portfolios in question did not actually obtain higher skewness out-of-sample, the weights are based on in-sample data where they did outperform the remaining portfolios.

5.2.3 Performance Measures

In financial literature it is common to present performance measures to evaluate the relative efficacy of a given portfolio strategy. We consider the Sharpe Ratio and the Adjusted for Skewness Sharpe Ratio to be representative for normal investor preferences and non-normal investor preferences, respectively. A hypothesis test is conducted on the SR to obtain a proper statistical inference about relative performance. However, to our knowledge, no such test has been developed for the ASSR and we thus include this second measure as a point of reference with the intent to see if taking skewness into account markedly changes the measure, or if ranking of the strategies change.

The hypothesis test for the SR is developed by Ledoit and Wolf (2008) where robust inference about the measure is the goal. The null hypothesis states that there are no difference between the SRs of two given strategies a and b; the alternative hypothesis asserts that there is

$$H_0: \frac{\mu_a}{\sigma_a} - \frac{\mu_b}{\sigma_b} = 0 \qquad H_A: \frac{\mu_a}{\sigma_a} - \frac{\mu_b}{\sigma_b} \neq 0. \tag{20}$$

We use the MV portfolio as a natural benchmark which every other strategy is compared against. A low (high) P-value means that the Sharpe Ratio of strategy b is significantly (not significantly) different from the Sharpe Ratio of the MV portfolio. Further details on the test can be found in the cited paper.

15 stocks

Table 10: Performance Measures. 15 stocks.

Strategy	Sharpe Ratio	ASSR $b_3 = 1$	ASSR $b_3 = 3$	P-value
MV	0.3847	0.3775	0.3626	-
MVS	0.3662	0.3685	0.3731	0.8078
MMVS	0.3315	0.3303	0.3279	0.4614
GMV	0.2558	0.2665	0.2866	0.0843
EW	0.1130	0.1141	0.1163	0.0013

*Using skewness prefrence $b_3 = 1$. Figures are based on monthly data. P-values are related to whether the Sharpe Ratios of a given strategy is equal to the MV strategy.

Considering 15 stocks, we see that MV, MVS and MMVS are very close in regards to the SR, with only 0.05 disparity between MV (0.3847) and MMVS (0.3315). The null is not rejected as a result of high P-values (0.81 and 0.46 for MVS and MMVS), supporting the statement that the skewness-strategies did not give up a statistically significant amount of return (or add significant amounts of risk) in the pursuit of higher skewness. As seen in the out-of-sample results, the GMV portfolio fares less favorably than MV, resulting in a SR of only 0.2558 which is a 0.13 deficit. This is a significant difference at the 10% confidence level, implying that one may be reasonably suspicious about the null. The EW strategy clearly underperforms with a SR of only 0.1130 with a clear rejection of the null of equality with MV at even the 1% level. While there exist some literature that support the claim that the 'naïve' EW strategy is competitive with more complex strategies (e.g. DeMiguel et al., 2009), this notion does not seem to be supported with our findings.

With a skewness preference of $b_3 = 1$, ASSR of the different strategies does not change from SR in any appreciable way. The strategies which achieve positive skewness (MVS, GMV, EW) improve at most by 0.01, and those with negative (MV, MMVS) worsen at most by -0.007. The ranking remains the same as with SR in this case. Investors with a higher preference for skewness (and thus lower risk aversion), $b_3 = 3$, see a slightly greater change. By the numbers, MVS now ranks higher than MV at 0.3731 and 0.3636 respectively, but it is still reasonable to claim they are equal. GMV which has the highest skewness, sees the biggest change from SR by ∼0.03; a negligible amount.

5 stocks

Table 11: Performance Measures. 5 stocks.

Strategy	Sharpe Ratio	ASSR $b_3 = 1$	ASSR $b_3 = 3$	P-value
MV	0.2556	0.2584	0.2641	-
MVS	0.2131	0.2142	0.2165	0.2118
MMVS	0.2661	0.2676	0.2706	0.7631
GMV	0.2477	0.2509	0.2572	0.7162
EW	0.0383	0.0384	0.0387	0.0202

*Using skewness prefrence $b_3 = 1$. Figures are based on monthly data. P-values are related to whether the Sharpe Ratios of a given strategy is equal to the MV strategy.

For 5 stocks, we note that the performance measures reflect the out-of-sample results. MMVS is highest ranked according to SR, with MV slightly behind (0.2661 and 0.2556). MVS, due to poor return, falls behind GMV with 0.2131 against 0.2477 respectively. EW produces a SR of 0.0383 and comes in last by a wide margin. Compared to MV, we note that MVS, MMVS and GMV all fail to reject the null, and as such they can be said to have the same SR. Only EW seems to be significantly worse than MV at any reasonable confidence level.

Adjusting SR for skewness yield results similar to the 15-stock case, namely that there are only marginal impacts. Both with $b_3 = 1$ and $b_3 = 3$, ASSR barely budges from the SR value and the ranking of the portfolios remain the same. The small impact of skewness on ASSR stems from a low initial SR value, causing a low leveraging effect of skewness. This is shown in the EW figures: despite having the highest (annualized) out-of-sample skewness of 0.1499, this strategy due to a low SR (0.0383) experienced the lowest delta of any ASSR against SR.

5.2.4 Sensitivity to Outliers - 15 stocks

In the results for the 15-stock scenario, we mention that one outlier has a significant effect on the performance of the GMV and EW portfolios. The asset that causes this outlier (Fred. Olsen Energy/Dolphin Drilling) has an abnormally high return (135%) in a single month, which yields a significantly higher average return, standard deviation and skewness for the two portfolios in question. The outlier month can be seen in figure 3, toward the end of 2017 for the EW and GMV portfolios. MV, MVS and MMVS were not affected by the outlier because they did not invest significantly in this asset. To understand why we look at the in-sample moments. The asset in question has a low variance, but also very low return and skewness. The GMV strategy - which only prioritizes minimizing risk - thus includes the asset in the portfolio, while the MV, MVS and MMVS strategies which prioritize other moments, avoid this asset. EW was affected by the outlier because the portfolio by design has to invest some wealth in every asset no matter what. We include in table 12 the results where the one outlier month is removed for the GMV and EW strategies. The 3 other strategies are not included as they did not change significantly.

Table 12: Out-of-Sample Results for GMV and EW. 15 stocks

Strategy	Return	Standard deviation	Skewness
GMV	0.1185	0.1370	0.2878
GMV*	0.0745	0.1097	-0.2096
EW	0.0535	0.1681	0.1533
EW*	0.0177	0.1532	0.0553

Numbers are annualized averages.

As the table suggests, GMV and EW perform significantly worse with the outlier removed although the overall ranking of the 5 strategies do not change. The strategies experience a decline in average return of roughly 3.6 and 4.4 percentage points (67% and 37% lower returns, respectively). In the absence of the extreme month, GMV and EW achieved a lower standard deviation, 2.7 and 1 percentage points respectively. In particular, the GMV portfolio is now second only to the MV portfolio, with a negligible difference between the two, indicating that the risk-minimizing strategy works quite well when taking the outlier into account. Removing the outlier changes the skewness of GMV and EW by -0.5 and -0.1 respectively - a very significant change. This exemplifies the fact that skewness is a very sensitive measure, as a change in only 1 out of 48 observations (∼2% of the out-of-sample dataset) has a drastic impact.

6 Conclusion

In this thesis we pursue two main lines of inquiry. Firstly, can we use skewness in portfolio management at Oslo Stock Exchange (OSE)? Secondly, does taking skewness into account in portfolio management improve out-of-sample performance using Polynomial Goal Programming (PGP)?

To answer the first question we investigate the ubiquity of skewness, the persistence of skewness and the effect of portfolio size on the average portfolio skewness for 80 randomly chosen individual asset returns. With regards to ubiquity of skewness in the market, we find that 58% of stocks have positively skewed return, 15% negative, and 27% no significant skewness using data from the period 2006 to 2017. In addition, we find that the average skewness of stocks at OSE has increased in the same time frame, and that there are significant differences between sectors of the economy. This provides motivation for us to further study non-normal investor preferences. Afterwards, we count the number of assets that has persistent skewness in two following time periods. We find that 56% to 69% of assets returns at OSE are positively skewed in adjacent periods; comparable numbers are 12% and 0% for negative skewness. This means that it is to a certain extent possible to predict future (positive) skewness for individual assets. Lastly, we see that skewness and amount of stocks are negatively correlated.

With the first main point as a basis, we answer the second main point by creating a portfolio that takes skewness into account. We do this by performing the PGP procedure to create two strategies that incorporates skewness: the Mean-Variance-Skewness (MVS) and the Modified Mean-Variance-Skewness (MMVS) portfolio. Out-of-sample results are compared with the traditional Mean-Variance (MV), the Global-Minimum-Variance (GMV) and the Equally-Weighted (EW) portfolio strategies. When evaluating all of the portfolios we find that the MV portfolio manages to obtain the highest Sharpe Ratio for the scenario of 15 stocks and the second highest for 5 stocks. However, when we adjust of skewness, we find that the MVS portfolio is able to obtain the highest out-of-sample Adjusted for Skewness Sharpe Ratio (ASSR) with a high preference for skewness ($b_3 = 3$) for 15-stocks. The MMVS obtains the highest out-of-sample ASSR for 5-stocks with both low and high preference for skewness. On the other hand, in the 5-stock scenario, MVS performs worse than the MV portfolio with regards to ASSR. Results are therefore mixed in the two different cases.

Our findings suggests skewness is difficult to accurately predict in portfolio management. For the 5-stock portfolio, we find that MVS and MMVS does not achieve a higher skewness than the MV portfolio. For the 15-stock portfolios, both the MVS and the MMVS obtain higher skewness than the MV portfolio. This could indicate that it is possible to predict systematic skewness for a portfolio size of 15. Additionally, MVS consistently outperforms MMVS in regards to out-of-sample skewness despite having a lower skewness preference. Greater skewness preference does not predict out-of-sample skewness.

We find two noteworthy points regarding diversification. First, the strategies obtain slightly higher portfolio skewness on average in the 5-asset scenario than the 15-asset scenario. Secondly, the two strategies with a preference for skewness are less diversified than the alternatives. As a consequence, they also has higher standard deviation (risk). The implication is that (1) a preference for skewness leads to less diversified portfolios and (2) skewness and risk are opposing goals.

By performing a sensitivity analysis, we find that one outlier had a significant effect on portfolio performance with regards to skewness. In the 15-stock scenario, skewness of the GMV portfolio changes from 0.29 to -0.21 when removing the outlier; for EW the numbers are 0.15 and 0.06. Investors with skewness preference thus need to be particularly careful about ex-ante performance measurement.

7 Weaknesses and Further Research

We chose firms based on the criteria that they traded the entire period 2004-2017 which naturally excludes companies that went bankrupt during this time. This introduces a survivorship bias into our analysis, which may affect skewness of returns. Possibly, firms that go bankrupt are overrepresented in extreme negative returns before they dissolve. To remove the bias, one should not have a barrier towards firms that went bankrupt (or even got delisted).

On a technical note, we used Polynomial Goal Programming to create optimal strategies. An inherent weakness of the method is the fact that preferences are not derived from the investor's utility function, making the magnitude of the preference parameters entirely arbitrary. This makes ex-post comparisons of MVS to other strategies more difficult. There are alternative methods which may yield comparable or better results which include techniques like Taylor-approximation ('the dual approach'). Additionally, estimating moments remains a difficult task due to significant estimation error.

For further research on skewness in portfolio management, one could take a look at the effect of different company sizes and how they might have an effect on overall skewness. We would expect that if one discriminates assets based on market capitalization, there could be a trend for skewness. For example, we hypothesize that smaller sized companies are more prone to large one-month return which can have a sizeable effect on the overall portfolio skewness, as explained in our sensitivity analysis. Another topic of interest may be investigating high vs low growth firms. Yet another is a deeper analysis on how securities from different sectors of the economy are differentiated in regards to skewed returns, and more importantly why.

8 References

Albuquerque R., 2012, Skewness in Stock Returns: Reconciling the Evidence on Firm Versus Aggregate Returns, *The Review of Financial Studies Volume 25 Issue 5*, pp 1630–1673, doi 10.1093/rfs/hhr144

Bartz D., Höhne J., Müller K.R., 2014, Multi-Target-Shrinkage, *arXiv preprint*, arXiv: 1412.2041

Beedles W.L., Simkowitz W.A., 1978, Diversification in a Three-Moment World, *The Journal of Financial and Quantitative Analysis Volume 13 No.5*, pp 927-941, doi 10.2307/2330635

Beedles W.L., 1979, Return, Dispersion, and Skewness: Synthesis and Investment Strategy, *The Journal of Financial Research Volume 2 issue 1*, pp 71-80, doi 10.1111/j.1475-6803.1979.tb00018.x

Beedles W.L., Simkowitz M.A., 1980, Morphology of Asset Asymmetry, *Journal of Business Research Volume 8 Issue 4*, pp 457-468, doi 10.1016/0148-2963(80)90018-1

Boudt K., Cornilly D., Verdonck T., 2017, Supplementary Appendix to: A Coskewness Shrinkage Approach for Estimating the Skewness of Linear Combinations of Random Variables, doi 10.2139/ssrn.2970015

Boudt K., Cornilly D., Verdonck T., 2016, A Coskewness Shrinkage Approach for Estimating the Skewness of Linear Combinations of Random Variables, *Journal of Financial Econometrics, Forthcoming*, doi 10.2139/ssrn.2839781

Briec W., Kerstens K., Jokung O., 2007, Mean-Variance-Skewness Portfolio Performance Gauging: A General Shortage Function and Dual Approach, *Management Science Volume 53 no.1*, pp 135-149, doi 10.1287/mnsc.1060.0596

Brooks C., Kat H.M., 2002, The Statistical Properties of Hedge Fund Index Returns and Their Implications for Investors, *The Journal of Alternative Investments Volume 5 issue*

2, pp 26-44, doi 10.3905/jai.2002.319053

Canela M.A., Collazo E.P., 2007, Portfolio selection with skewness in emerging market industries, *Emerging Markets Review Volume 8 Issue 3*, pp 230-250, doi 10.1016/j.ememar.2006.03.001

Chunhachinda P., Dandapani K, Hamid S, Prakash A.J, 1997, Portfolio selection and skewness: Evidence from international stock markets, *Journal of Banking & Finance Volume 21 Issue 2*, pp 143–167, doi 10.1016/S0378-4266(96)00032-5

DNB. (retrieved 20. January 2019), Prisliste aksjehandel, retrieved from https://www.dnb.no/privat/sparing-og-investering/aksjer/merinfo/priser.html

DeMiguel V., Garlappi L., Uppal R, 2009, Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?, *The Review of Financial Studies, Volume 22 issue 5*, pp 1915–1953, doi 10.1093/rfs/hhm075

Eling M., Schuhmacher F., 2007, Does the Choice of Performance Measure Influence the Evaluation of Hedge Funds?, *Journal of Banking & Finance Volume 31 issue 9*, pp 2632-2647, doi 10.1016/j.jbankfin.2006.09.015

Tayi G.K, Leonard P.A, 1988, Bank Balance-Sheet Management: An Alternative Multi-Objective Model, *The Journal of the Operational Research Society Volume 39 No. 4*, pp. 401-410, doi 10.2307/2582120

Harvey, C.R., Liechty, J.C., Liechty, M.W., Muller, P., 2002, Portfolio selection with higher moments, *The Journal of Quantitative Finance Volume 10 issue 5*, pp 469-485, doi 10.1080/14697681003756877

Harvey C.R., Siddique A., 2000, Conditional Skewness in Asset Pricing Tests, *The Journal of Finance Volume 55 Issue 3*, pp 1263-1295, doi 10.1111/0022-1082.00247

Jondeau E., Rockinger M., 2006, Optimal Portfolio Allocation under Higher Moments, *European Financial Management Volume 12 no.1*, pp 29-55, doi 10.1111/j.1354-7798.2006.00309.x

Jorion F., 1986, Bayes-Stein Estimation for Portfolio Analysis, *The Journal of Financial and Quantitative Analysis Volume 21 No.3*, pp. 279-292, doi 10.2307/2331042

Kraus A., Litzenberger R.H., 1976, Skewness Preference and the Valuation of Risk Assets, *Journal of Finance Volume 31 Issue 4*, pp 1085-1100, doi 10.2307/2326275

Lai K.K., Yu L., Wang S., 2006, Mean-Variance-Skewness-Kurtosis-based Portfolio Optimization, *In First International Multi-Symposiums on Computer and Computational Sciences (IMSCCS'06)*, doi 10.1109/IMSCCS.2006.239

Lai T.Y., 1991, Portfolio selection with skewness: A multiple-objective approach, *Review of Quantitative Finance and Accounting l*, pp. 293-305, doi 10.1007/BF02408382

Ledoit O., Wolf M., 2003, Improved estimation of the covariance matrix of stock returns with an application to portfolio selection, *Journal of Empirical Finance volume 10 issue 5*, pp 603 – 621, doi 10.1016/S0927-5398(03)00007-0

Ledoit O., Wolf M., 2008, Robust Performance Hypothesis Testing with the Sharpe Ratio, *Journal of Empirical Finance Vol 15, Issue 5*, pp. 850-859, doi 10.1016/j.jempfin.2008.03.002

Markowitz, H., 1952, Portfolio Selection, *The Journal of Finance Volume 7 no.1*, pp 77-91, doi 10.2307/2975974

Martellini L., Ziemann V., 2010, Improved Estimates of Higher-Order Comoments and Implications for Portfolio Selection, *The Review of Financial Studies Volume 23 Issue 4*, pp 467–1502, doi 10.1093/rfs/hhp099

Mitton T., Vorkink K., 2007, Equilibrium Underdiversification and the Preference for Skewness, *The Review of Financial Studies Volume 20 Issue 4*, pp 1255-1288, doi 10.1093/revfin/hhm011

Nordnet. (retrieved 20. January 2019), Prisliste Nordnet, retrieved from https://www.nordnet.no/tjenester/prisliste.html

Oslo Børs. (retrieved 8. March 2019), Hovedindeksen, retrieved from https://www.oslobors.no/

markedsaktivitet/#/details/OSEBX.OSE/overview

Prakash A.J., Chang C.H., Pactwa T.E., 2003, Selecting a portfolio with skewness: Recent evidence from US, European, and Latin American equity markets, *Journal of Banking & Finance Volume 27 Issue 7*, doi 10.1016/S0378-4266(02)00261-3

Richard J-C., Roncalli T., 2015, Smart Beta: Magaging Diversification of Minimum Variance Portfolios, *Risk-Based and Factor Investing*, pp 31-63, doi 10.1016/B978-1-78548-008-9.50002-2

Shapiro S.S. and Wilk M.B., 1965, An Analysis of Variance Test for Normality (Complete Samples), *Biometrika Volume 52 No. 3/4*, pp. 591-611, doi 10.2307/2333709

Sharpe W.F., 1994, The Sharpe Ratio, *The Journal of Portfolio Management Volume 21 Issue 1*, pp 49-58, doi 10.3905/jpm.1994.409501

Simaan Y., 1993, Portfolio Selection and Asset Pricing—Three-Parameter Framework, *Management Science Volume 39 no.5*, doi 10.1287/mnsc.39.5.568

Singleton J.C., Wingender J.R., 1986, Skewness Persistence in Common Stock Returns, *Journal of Financial and Quantitative Analysis*, pp 335-341, doi 10.2307/2331046

Stein C., 1956, Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution, *Proc. Third Berkeley Symp. on Math. Statist. and Prob. Volume. 1 (Univ. of Calif. Press, 1956)*, pp 197-206, URL https://projecteuclid.org/ euclid.bsmsp/1200501656

Sun Q., Yan Y., 2003, Skewness persistence with optimal portfolio selection, *Journal of Banking & Finance Volume 27 Issue 6*, pp 1111-1121, doi 10.1016/S0378-4266(02)00247-9

Woerheide W., Persson D., 1993, An Index of Portfolio Diversification, *Financial Services Review. Volume 2, Issue 2*, pp. 73–85, doi 10.1016/1057-0810(92)90003-U

Zakamouline V., Koekebakker S., 2009, Portfolio performance with generalized Sharpe ratios:

Beyond the mean and variance, *Journal of Banking & Financ Vol. 33*, pp 1242-1254