Including ESG concerns in the portfolio selection process

AN MCDM APPROACH

EDVIN LUNDSTRÖM, CARL SVENSSON
Including ESG concerns in the portfolio selection process

AN MCDM APPROACH

EDVIN LUNDBERG
CARL SVENSSON

Degree Project in Applied Mathematics and Industrial Economics (15 credits)
Degree Progr. in Industrial Engineering and Management (300 credits)
Royal Institute of Technology year 2014
Supervisor at KTH was Krister Svanberg
Examiner was Johan Karlsson

TRITA-MAT-K 2014:13
ISRN-KTH/MAT/K-14/13--SE

Royal Institute of Technology
School of Engineering Sciences
KTH SCI
SE-100 44 Stockholm, Sweden
URL: www.kth.se/sci
In recent years investors in the financial markets around the globe have begun to focus on non-financial factors in their portfolio selection processes. Three main areas of concerns are: Environmental, Social and corporate Governance (ESG). Previous research has mainly focused on implementing these concerns using qualitative methods, e.g. negative screening. Our thesis integrates these concerns in a Multi-Criteria Decision-Making (MCDM) framework, making it possible for investors to view the portfolio selection as a trade-off between three criteria: Return, Risk and ESG. This extends the traditional Markowitz frontier from two to three dimensions. Companies included are the ones in the index OMXS30. Return and risk are estimated using the single-index model. The ESG criterion is implemented as a linear function and estimated using two public ESG indices.

We will use two different optimization methods, the weighted sum approach and the ε-constraint method to compute the efficient frontier. These are evaluated and we conclude that each method has its own strengths and weaknesses. We can see that integrating ESG concerns as a third objective in addition to risk and return alters the portfolio selection process. It increases the complexity of choosing a portfolio, but also yielding a better decision basis for the investor. To mitigate the increase of complexity we propose the ESG-to-variability ratio in analogy with the Sharpe ratio, effectively reducing the number of portfolios an investor should consider.
Sammanfattning


För att beräkna den effektiva fronten använder vi två optimeringsmetoder: the weighted sum approach och the $\varepsilon$-constraint method. Dessa utvärderas och vi drar slutsatsen att respektive metod har såväl styrkor som svagheter. Vi kan se att ett inkluderande av ESG som en tredje målfunktion, utöver risk och avkastning, förändrar portföljvalsprocessen. Komplexiteten vid portföljval ökar, samtidigt som investeraren får ett bättre beslutsunderlag. För att lindra ökningen av komplexitet så introducerar vi the ESG-to-variability ratio i analogi med Sharpe ratio, vilket effektivt reducerar antalet portföljer en investerare bör välja emellan.
# Contents

1 Introduction
   1.1 Background .................................................. 1
   1.2 Problem formulation ......................................... 2
   1.3 Outline of the report ....................................... 3

2 Methodology
   2.1 ESG .......................................................... 4
   2.2 Portfolio methods and evaluation .............................. 5
       2.2.1 Multiple-criteria decision-making ..................... 6
       2.2.2 Data used ................................................. 7
   2.3 Delimitations ............................................... 7

3 Some preliminaries
   3.1 Return ....................................................... 8
   3.2 Variance ..................................................... 9
   3.3 Convexity .................................................... 10
   3.4 Optimality .................................................. 10
   3.5 Multi-objective optimization ................................. 11
   3.6 Pareto optimality ........................................... 11

4 Portfolio optimization
   4.1 Introduction to MPT ......................................... 13
       4.1.1 Initial stochastic programming problem ............... 13
       4.1.2 Deterministic formulation ............................. 14
   4.2 Portfolio selection with multiple criteria ................. 16
   4.3 Estimating parameters ....................................... 17
       4.3.1 A naïve approach ....................................... 17
       4.3.2 The single-index model ................................. 18

5 Environmental, Social and corporate Governance concerns
   5.1 Introduction to ESG ......................................... 23
   5.2 Investors and ESG ........................................... 24
   5.3 Companies and ESG .......................................... 24
   5.4 Ratings ...................................................... 26
       5.4.1 Folksam’s Index for Responsible Business .......... 27
       5.4.2 STOXX ESG Index ........................................ 27
# List of Figures

3.1 Classifying the efficient solutions ................................................. 12
3.2 The ideal and nadir objective vector ............................................. 12
4.1 Hierarchical structure ................................................................. 15
4.2 Efficient frontier ............................................................................. 16
4.3 Plot of $r_{it}$ versus $r_{mt}$ .............................................................. 21
5.1 The components of an ESG rating .................................................. 26
6.1 WSM with two objective functions .................................................. 30
6.2 $\varepsilon$CM with two objective functions ......................................... 31
8.1 The efficient frontier ....................................................................... 39
8.2 Efficient frontier from different angles ............................................. 39
8.3 Efficient frontier as contour plots .................................................... 40
8.4 With or without early exits ............................................................... 43
8.5 Different ways to compare run time ............................................... 44
8.6 Comparing distances between normalized and regular WSA ............. 45
8.7 Comparing distances between WSA and $\varepsilon$CM .......................... 46
8.8 Alternative dispersion measurements .............................................. 46
8.9 Nondominated solutions generated by the different methods .......... 47
9.1 How order affects elimination ......................................................... 52

B.1 Number of stocks in portfolio along a contour curve ...................... 63
B.2 Contour lines with ESG-to-variability ratio maximizing portfolios ...... 63
B.3 Contour lines with Sharpe ratio maximizing portfolios ..................... 64
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Examples of ESG factors</td>
<td>23</td>
</tr>
<tr>
<td>6.1</td>
<td>Estimating nadir values</td>
<td>33</td>
</tr>
<tr>
<td>8.1</td>
<td>Estimated alpha, beta, mean, variance and ESG rating</td>
<td>37</td>
</tr>
<tr>
<td>8.2</td>
<td>Investor profiles</td>
<td>41</td>
</tr>
<tr>
<td>8.3</td>
<td>Investor specific portfolios</td>
<td>42</td>
</tr>
<tr>
<td>C.1</td>
<td>Ten principles in the areas of human rights, labour, the environment and anti-corruptions</td>
<td>65</td>
</tr>
<tr>
<td>C.2</td>
<td>ESG data used</td>
<td>66</td>
</tr>
<tr>
<td>C.3</td>
<td>ESG data as ranks</td>
<td>67</td>
</tr>
<tr>
<td>C.4</td>
<td>Parameters and ranks</td>
<td>68</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Background

In a groundbreaking article published in 1952, Harry M. Markowitz introduced what would become the foundation of Modern Portfolio Theory (MPT). For this he would later receive The Sveriges Riksbank Prize In Memory of Alfred Nobel, for “having developed the theory of portfolio choice” (KVA, 1990). Rubenstein (2002) goes as far as claiming that the publication in 1952 constitutes “the moment of birth of modern financial economics.”

Previous research had used Bernoulli (1954) law of large numbers, leading to the conclusion that all risk could be diversified away. Markowitz, however, claimed that the law of large numbers was not applicable to a portfolio of securities, since returns are too intercorrelated. Therefore diversification cannot eliminate all variance. Hence, there is a trade-off between return and variance. Furthermore, he assumed that investors are risk averse. This means that when choosing between two portfolios with the same expected return, the investor will choose the one with the smallest variance. Markowitz (1952), divides the portfolio selection process into two stages:

1. From observation and experience one forms beliefs about the future performances of available securities.

2. With the relevant beliefs formed in the first stage, one must choose a single portfolio.

Markowitz was primarily interested in the second stage. His main proposal is the $E - V$ rule. It states that an investor would (or should) want to select one of the efficient portfolios, i.e. those with minimum $V$ (variance) for given $E$ (expected return) or more, or put differently, those with maximum $E$ for given $V$ or less. From a given set of $(E, V)$ combinations, i.e. the efficient frontier, it is then up to the investor to choose the preferred portfolio. In the article from 1952 he presents the optimization problem associated with this rule.

During the years following his publication Markowitz’s theory was hailed and can still be found in elementary finance books as e.g. Berk and DeMarzo (2014). However, economic theory has progressed and several critics have questioned the theory, their claim is that it does not match the real world in several ways (Xidonias et al., 2012). This has led to several offsprings of MPT, one being Multi-Criteria Portfolio Theory. Steuer, Qi, et al. (2008) claims that: “There has been growing interest in how to incorporate additional criteria beyond “risk and return” into the portfolio selection process.” Wang and Lee (2011)
assert that investors often need to base their portfolio selection on subjective preferences, with several nontraditional\(^1\) criteria. The relevance of a multi-criteria approach is in part due to:

- The traditional model (MPT) has been criticized for its dependence on the estimation of parameters from historical data. By including criteria as R&D spending, ESG criteria\(^2\) or the opinions of stock market analysts the model may mitigate the imperfections associated with the estimation of the traditional parameters (Steuer, Qi, et al., 2008).

- The assumption of rational economic agents has been challenged, see e.g. Kahneman and Tversky (1979) in the field of behavioral economics. Multiple Criteria Portfolio Optimization can be seen as a way of incorporating nontraditional (and in some sense irrational) goals in a standard mathematical framework Xidonas et al. (2012).

- The increased concerns of social responsibility and environmental change can in this framework be integrated in the decision basis for investors. One survey showed that two in three of the large asset owners in the Nordic countries use ESG criteria (Novethic, 2013a) and another showed that in the UK, 57 percent of investors believe that ESG is a factor in financial performance (BNP Paribas, 2010).

Previous research has mainly been focused on the mathematics behind extending the model from a two-criteria problem to a multiple-criteria optimization problem. That is, focusing on the mathematical interpretation and the computational aspects. Little work has been done in suggesting relevant criteria, and especially how they should be quantified.

### 1.2 Problem formulation

This thesis addresses investors with additional concerns, in excess of the traditional mean-variance criteria. If the mean-variance framework does not give a good representation of investor concerns, is there a model the investor could use? The thesis will examine this under the assumption that it is a multi-criteria model.

The purpose is to examine how an ESG criterion may be integrated in a mathematical framework, thus taking the ESG concerns of investors into account in the portfolio selection process. This involves examining the following issues:

- Motivate why ESG concerns should be included in the portfolio selection process. Suggest a mathematical interpretation and find suitable (public) data to construct an ESG criterion.

- Formulate an optimization problem that includes expected return, variance and ESG. Engineer a MATLAB algorithm that solves the subsequent multi-criteria optimization problem using two methods applicable to the problem. This is preceded by the estimation of the traditional parameters (i.e. expected returns and covariances). Furthermore we will discuss how the inclusion of a ESG criterion affects the portfolios of investors.

\(^1\)We distinguish between traditional and nontraditional criteria, where the former expression represents mean and variance while the latter includes all other possible criteria

\(^2\)Environmental, Social and corporate Governance
1.3 Outline of the report

- Chapter 2, Methodology, provides a description of how the thesis was completed. We briefly introduce and describe the main theories and models we rely upon, and some rationale to why they were chosen. Furthermore, we introduce the empirical data used.

- Chapter 3, Some preliminaries, provides some notation and basic mathematical concepts needed to understand the following chapters. This chapter can be skipped by readers with a mathematical background.

- Chapter 4, Portfolio optimization, starts by presenting the original mean-variance model, and then shows how it can be extended to a setting where additional criteria are included. The chapter’s last section presents a model for estimating expected returns and covariance.

- Chapter 5, Environmental, Social and corporate Governance, presents ESG and explains why it might be of interest for investors and companies to consider ESG factors. The last section reviews the two indices used in the subsequent chapters to include ESG concerns in the portfolio model.

- Chapter 6, Optimization methods, introduces the two optimization methods used to solve the MCDM problem in this thesis: the weighted sum approach and the $\varepsilon$-constraint method.

- Chapter 7, Implementation, gives a brief explanation of how the optimization problem is set up and solved.

- Chapter 8, Results, is divided into two parts. First we evaluate how the inclusion of ESG affects the portfolio choice by solving the optimization problem. Then we evaluate the used optimization methods using two main measurements, time and dispersion.

- Chapter 9, Discussion, contains interpretation and discussion of the results.

- Chapter 10, Conclusion, summarizes the conclusions we are able to make.
Chapter 2

Methodology

2.1 ESG

To investigate why investors should be concerned about ESG and if they already are, we used second hand data. One way of investigating the latter would have been to develop a questionnaire which could be distributed to a group of investors. This approach was deemed inappropriate for two reasons: (1) The possibility of being able to sample a large and well diversified group of investors are limited for our time-frame, and (2) Our primary goal is not to say if investors are concerned, but rather if they should be. We will instead try to answer this question by reviewing previous studies regarding the subject. The former question was answered by reviewing previous academic as well as non-academic research. We have chosen to not only review academic research since a lot of research about investor concerns are financed by or facilitated at non-academic institutions, i.e. banks, rating agencies and consulting firms. Possible problems with this approach is the problem of impartiality and lack of peer reviewing. Academic research was primarily searched for in KTHB Primo and Google Scholar. Non-academic research was primarily searched for and found by searching for research published by rating agencies as Novethic and non-profits such as the Global Reporting Initiative.

A lot of the previous research regarding ESG has focused on the possibility of using it as a proxy for financial performance, see e.g. Fulton et al. (2012) which reviews several studies. This relationship should, in our model, be taken as a “bonus” since the extra expected return lies outside the mathematical framework used in the thesis. Investors wanting to use ESG in our framework to achieve higher return can of course do it, but the effects are not quantified in terms of return.

A problem that arises is that there is no unified definition of ESG or non-financial concerns. Keywords used to search were therefore not restricted to ESG but also included similar definitions as: ”Socially Responsible Investing”, ”Sustainable and Responsible Investment”, ”Corporate Social Responsibility” and ”Global Reporting Initiative”.

Since there is no unified definition of ESG, there does not exist a uniform way of quantifying it. We will not try to derive our own ESG ratings from scratch since this requires extensive research and knowledge in fields we have yet to and might never master. We have no research budget and have thus not been able to buy data. This limitation has two main implications: (1) Our sample is limited and (2) The raw data used to construct free indices are seldom made public. One should note that we do not know of any ESG indices developed by independent institutions. The ESG indices used in this thesis are
mainly based on reporting done by the companies, this combined with no access to raw
data makes it hard to evaluate if the ratings are ”correct”. We tried to mitigate this effect
by limiting our universe of stocks to the OMXS30\(^1\). It seems reasonable to assume that
larger companies are more thoroughly reviewed than smaller companies. This limitation
might mitigate the problems associated with quantifying the ESG criterion. However, this
means a limitation of portfolio choice in which, all else equal, it is better with more stocks
to choose from.

To investigate how the two used indices could be combined into one we looked for
previous research regarding the combinations of indices or rankings. We found no relevant
research and we therefore chose a combination method based on our own arguments. We
chose this method before inspecting the data to ensure that there was no ”selection bias”
when choosing the method used for combination.

2.2 Portfolio methods and evaluation

Markowitz original work from 1952 serves as a basis for the second part of the problem
formulation. Following Markowitz’s initial endeavor in trying to, in a mathematical sense,
find optimal portfolios, a vast amount of research has been done. A good research paper
describing some of the most important findings is Elton and Gruber (1997).

In his initial publication, Markowitz did not focus on how the parameters needed for
the model should be estimated. How one should to this in the best way is still up for
debate. An early suggestion was proposed by Sharpe (1964), this so-called single-index model\(^2\) was originally formulated as: \(R_i = A_i + B_i I + C_i\). Where \(R_i\) is the return of the
asset \(i\), \(A_i\) and \(B_i\) parameters and \(C_i\) a random variable with expected value of zero. \(I\) is
the level of some index, and could according to Sharpe be “the level of the stock market
as a whole, the Gross National Product, some price index or any other factor thought to
be the most important single influence on the returns from securities”. This is the model
that we will use to estimate the needed parameters. The single-index model has received
some criticism and several extensions and modifications have been proposed. The most
common extension has been to introduce several indices, see e.g. Fama and French (2004)
and Chen et al. (1986). However, these models increases the complexity without providing
a better way to fulfill the purpose of our thesis, we settled with the single-index model.
The data we used will be discussed below.

In Markowitz (1952), no short-selling (i.e. speculating in decreasing stock prices) was
allowed. Allowing short-selling primarily affects the computational time and power needed
to compute the efficient frontier since it becomes unbounded\(^3\). If short-selling should be
included in the model is debatable since all investors cannot (or are not allowed to) engage
in it (Elton, 2014). Furthermore, a reason for us to not allow short selling is the ESG
objective function. A short position is equivalent to a negative ESG score, which implies
a flawed interpretation.

One of the assumptions behind MPT is that there are no transaction costs, i.e. you
can buy and sell shares for their market value, in reality this is not the case. We will not
go into more detail about including transaction costs and will disregard this in the rest of
the thesis. We refer the interested reader to Perold (1984).

\(^1\)consists of the thirty most traded stocks on the Stockholm Stock Exchange
\(^2\)Sharpe called it the market model
\(^3\)i.e. with short sells the expected return of a portfolio can take all values on \([-\infty, \infty]\)
Markowitz did not propose a “standard” algorithm for choosing a single portfolio from the efficient frontier, instead he stated “the investor could select the combination he preferred”. Some well know work has been done on the problem of choosing a single portfolio. The most famous is probably the Sharpe ratio introduced by Sharpe (1966), which we will discuss later on in the thesis.

As was mentioned in section 1.1 including multiple criteria (objectives) in the portfolio selection process has caught the interest of some researchers in finance and mathematics. We will not go into detail explaining research that has been done since a lot of it is outside the scope of this thesis. Steuer and Na (2003) provides a categorized bibliography on the applications of multiple criteria decision models in finance between 1955 and 2001. The paper ends with the authors saying “the research opportunities in exploring the application [boldness added] the multicriteria technologies in finance appear at this time to be particularly substantial.”

Unlike as with Markowitz development of MTP, it is difficult to find the origin of multiple-criteria portfolio optimization. Steuer, Qi, et al. (2008) claims that: “there has almost always been a slight undercurrent of multiple objectives in portfolio selection.” This paper served as our starting point, in which the authors present the underlying theory of multiple-criteria portfolio optimization. We used some of their results to formulate our optimization problem.

### 2.2.1 Multiple-criteria decision-making

While the literature mentioned above would mainly be considered as belonging to the field of finance, we must turn to the field of mathematics to actually be able to solve the optimization problem we face. Multiple-criteria decision-making1 (MCDM) are mathematical models for solving problems when several criteria are to be optimized at once. A central concept in MCDM is the decision-maker (DM), this is person or organization for whom the optimization is done. There are several ways to classify the different optimization methods used for solving MCDM optimization problems. Cohon (1985) groups the methods in two different categories:

1. Generating methods, methods that generate the set of efficient solutions, the decision maker is then presented with this set and chooses one solution. In these methods the decision maker is involved after the optimization.

2. Preference-based methods, these methods take into consideration the preferences of the decision maker before making or during the optimization. The generated solution is then the one that comes closest to satisfying these preferences.

In this thesis we are interested in two models belonging to the first category, but this classification is far from complete since at least one of the models could be used in a preference-based settings.

The first model used is the weighted sum approach (WSA) introduced in section 6.1 and the second is the ε-Constraint Method (εCM) introduced in section 6.2.

A primitive version of WSA can be seen in Zadeh (1963). As one might assume from its name, one of the problems with WSA is choosing the weighting factors, another problem is that WSA does not handle non-covex problems. Several propositions have been made

---

1 in some texts Multiple-criteria decision analysis (MCDA)
to address these problems, see e.g. Kim and Weck (2004) and Chankong and Haimes (1983). We alternated WSA by normalizing the functions, yielding increased performance.

$\varepsilon$CM was first introduced in Haimes et al. (1971). Research has mainly focused on adapting the method for different types of problems, this often involves choosing the $\varepsilon$ constraints in an efficient way. An interesting extension is the augmented $\varepsilon$-constraint method presented in Mavrotas (2009), which we used parts of.

2.2.2 Data used

As we discuss in section 8.1, we studied the effects of inclusion of ESG concerns by using a hypothetical setting. The portfolio selection would take place May 1, 2014 and the portfolio would be held during one year. To estimate the expected returns and covariances we used the single-index model. This model uses historical data to produce estimates of future market conditions. However, regarding the length of time series to be used neither researchers nor practitioners are in agreement. We followed the advice from Berk and DeMarzo (2014) and used weekly adjusted closing prices for the three years preceding the portfolio selection date. We want to point out once again that for our purpose it is enough with plausible estimates, rather than the best possible.

We started by limiting the number of stocks to the ones listed on the index OMXS30. This depends on three things: the first was mentioned in section 2.1, the second is that the burden of collecting data decreases and third, the computation times for optimization problem does not become unreasonably long. Moreover, we felt it was reasonable to reduce the sample further. Since both the A and B shares of Atlas Copco are included in OMXS30 and the ESG data we used concerned the company as a whole, we decided to only include the A share in our sample. We chose to keep the A share instead of the B share since it had a larger weight in the OMXS30. This left us with 29 different stocks. The historical prices were then retrieved from Yahoo Finance. We did not validate the data by comparing it to data provided by another source, e.g. Nasdaq OMX.

2.3 Delimitations

We will, to a large extent, build our work on standard mathematical and financial theories. We will not try to further develop or invent theories in these areas. Instead, we are limited to use some of the more basic methods provided in these fields.

When solving an optimization problem with multiple objectives, we will use methods that are primitive compared to state-of-the-art methods. Since we use somewhat inferior methods for computation we will limit our thesis to include a smaller amount of securities. However, this limitation will not interfere with the purpose of the thesis.

To summarize, while there may be alternative (and in some aspects better) ways of solving the problems considered in this thesis, we will use less advanced methods.
Chapter 3

Some preliminaries

This chapter contains some basic definitions and concepts that will be used throughout the thesis.

3.1 Return

Fundamental in this thesis is the notion of return. The return of a security is the gain or loss in a particular time period (Berk and DeMarzo, 2014). For a security \( i \) we will denote this as \( R_i \). Furthermore, denote the price of the security at time \( t \) as \( P_t \). Then the return for a non-dividend paying security over the time period \( [0, T] \) is given by:

\[
R_i = \frac{P_T - P_0}{P_0} \tag{3.1.1}
\]

For a dividend paying stock the return is:

\[
R_i = \frac{Div_T + P_T}{P_0} - 1 = \frac{Div_T}{P_0} + \frac{P_T - P_0}{P_0} \tag{3.1.2}
\]

Since the return over the time period \( [0, T] \) is unknown at time 0, this is a random variable. The outcome of the random variable \( R_i \) will be denoted \( r_i \). However, more important for the problems considered in the thesis is the expected return. We will use the notation:

\[
\mu_i = \mathbb{E}[R_i] \tag{3.1.3}
\]

For the sake of convenience, use the vector notation for investments 1 to \( n \):

\[
\mu^\top = [\mu_1, \ldots, \mu_n] \tag{3.1.4}
\]

When multiple securities are combined into a portfolio we assign each security a portfolio weight according to:

\[
x_i = \frac{\text{Value of investment } i}{\text{Total value of portfolio}} \tag{3.1.4}
\]

The composition of a portfolio can thus be summarized as a vector:

\[
x^\top = [x_1, \ldots, x_n]
\]
The (random) return of a portfolio is given by:

\[ R_P = \sum_{i=1}^{n} x_i R_i \]  

(3.1.5)

And thus the expected return of a portfolio is:

\[ E[R_P] = E \left[ \sum_{i=1}^{n} x_i R_i \right] = \sum_{i=1}^{n} x_i E[R_i] = \sum_{i=1}^{n} x_i \mu_i = \mu^\top x \]  

(3.1.6)

### 3.2 Variance

The most common measure of risk is the variance of a security or a portfolio and consequently also the standard deviation. In finance, it is common to refer to the standard deviation as the volatility. The variance is defined as:

\[ \sigma^2 = Var(R) = E[(R - E[R])^2] \]  

(3.2.1)

The standard deviation or volatility follows as:

\[ \sigma = StDev(R) = \sqrt{Var(R)} \]

It is of interest to calculate the variance of a portfolio of stocks. The derivation follows below:

\[ \sigma^2(x) = E \left[ \left( \sum_{i=1}^{n} x_i R_i - E[R_P] \right)^2 \right] = E \left[ \left( \sum_{i=1}^{n} x_i (R_i - E[R_i]) \right)^2 \right] \]

\[ = E \left[ \left( \sum_{i=1}^{n} x_i (R_i - E[R_i]) \right) \left( \sum_{j=1}^{n} x_j (R_j - E[R_j]) \right) \right] \]

\[ = E \left[ \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j (R_i - E[R_i])(R_j - E[R_j]) \right] \]

\[ = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j E[(R_i - E[R_i])(R_j - E[R_j])] = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j \sigma_{ij} = x^\top C x \]

That is:

\[ \sigma^2(x) = x^\top C x \]  

(3.2.2)

Where we introduced the notation \( \sigma_{ij} \) for the covariance between asset \( i \) and \( j \). Also note that we will use \( \sigma_{ii} = \sigma_i^2 \). \( C \) denotes the covariance matrix, i.e. a matrix holding the covariances between all available assets. The diagonal elements are the variances for each security.
3.3 Convexity

In this section and in section 3.4 some of the basic concepts in mathematical optimization utilized in this thesis are presented. For a more thorough review of these concepts see Sasane and Svanberg (2013) or Griva et al. (2009).

A set \( F \) is said to be convex if, for all elements \( x, y \in F \) and \( t \in [0, 1] \):

\[
(1 - t)x + ty \in F \quad (3.3.1)
\]

A function \( f \) is said to be convex on the set \( F \) if:

\[
f((1 - t)x + ty) \leq (1 - t)f(x) + tf(y) \quad (3.3.2)
\]

and strictly convex if changed to a strict inequality. If the inequality is flipped the function is said to be concave. From this it is easy to see that if \( f \) is concave, \(-f\) is convex.

If \( f(x) \) is a convex function and \( F \) is a convex set, the resulting optimization problem is called a convex optimization problem:

\[
\begin{align*}
\text{(CO)} & \quad \begin{cases}
\text{minimize} & f(x) \\
\text{subject to} & x \in F
\end{cases}
\end{align*}
\]

Here \( f(x) \) is called the objective function and \( F \) the feasible set.

3.4 Optimality

To be able to classify feasible points to a problem we need to define different types of optimality. A point \( x_\ast \in F \) is called a local minimizer if there is an \( \epsilon \) such that:

\[
f(x_\ast) \leq f(x) \quad \forall x \in F \quad \text{such that } ||x - x_\ast|| < \epsilon \in \mathbb{R}_+ \quad (3.4.1)
\]

and a strict local minimizer if changed to a strict inequality. \( x_\ast \) is called a global minimizer if:

\[
f(x_\ast) \leq f(x) \quad \forall x \in F \quad (3.4.2)
\]

A convex problem has the special property that a local minimizer is a global minimizer. To determine if the optimal solution to a convex problem has been found the following definitions of a feasible direction and a descent direction are useful. For a vector \( d \in \mathbb{R}^n \) and a feasible point \( x \in F \) and a scalar \( t \in (0, \epsilon) \), \( d \) is called a feasible direction if:

\[
x + td \in F \quad (3.4.3)
\]

and a descent direction if:

\[
f(x + td) < f(x) \quad \forall t \in (0, \epsilon) \quad (3.4.4)
\]

For a convex problem \( x_\ast \) is a global minimizer if and only if there is no feasible descent direction for the objective function at \( x_\ast \).

It can be shown that a strictly convex problem for which there exists an optimal solution, the optimal solution is unique. We will now show that if a solution exists, it is unique. Assume that we have found a feasible optimal solution, \( x_\ast \) and that there exists another feasible optimal solution, \( y \). First note that by 3.3.1 \((1 - t)x_\ast + ty \) is feasible, we then have by using equation 3.3.2 that:

\[
f(x_\ast + td) = f((1 - t)x_\ast + ty) < (1 - t)f(x_\ast) + tf(y) = f(x_\ast).
\]

Which according to equation 3.4.4 implies that there exists a feasible descent direction, an obvious contradiction to equation 3.4.2 with a strict inequality sign.
3.5 Multi-objective optimization

A multi-objective optimization problem can be formulated as:

\[
\begin{align*}
\text{minimize} & \quad f_1(x) \\
\text{minimize} & \quad f_2(x) \\
& \vdots \\
\text{minimize} & \quad f_n(x) \\
\text{subject to} & \quad x \in \mathcal{F}
\end{align*}
\]  

(MO)  

(3.5.1)

We call \([f_1(x), f_2(x), \ldots, f_n(x)]^\top\) the objective vector and \(f_j(x)\) the \(j\):th objective function. Whereas the optimal solution to a single-objective problem is easily defined, the concept of a unique optimal solution to (MO) is not. As is often the case in this thesis, a point that is optimal for one objective function is seldom optimal for another. A typical example is the trade-off between return and risk for a portfolio, a high return is often associated with a high risk.

3.6 Pareto optimality

To address the problem of contradicting objective values, Pareto optimality is introduced. In words a Pareto optimal point is a point such that no objective function can get a “better” value without “worsening” another objective function. A Pareto optimal point will henceforth be called an efficient solution. The naming of different types of optimality differ in literature, here we will use the same as in Caramia and Dell’Olmo (2008). Mathematically we have that a point \(x^\ast\) is called a weak efficient solution to MO if there is no point \(x \in \mathcal{F}\) such that:

\[
f_i(x) < f_i(x^\ast) \quad \forall \quad i \in \{1, 2, \ldots, n\}
\]

(3.6.1)

and a point \(x^\ast\) is called a strictly efficient solution to (MO) if there is no \(x \in \mathcal{F}\) such that:

\[
f_i(x) \leq f_i(x^\ast) \quad \forall \quad i \in \{1, 2, \ldots, n\}
\]

(3.6.2)

with at least one strict inequality. From this follows that every strictly efficient solution to (MO) is also a weak efficient solution to (MO). We define \(\mathcal{F}_\ast\) as the set of efficient solutions:

\[
\mathcal{F}_\ast = \{x^\ast \in \mathcal{F} | x^\ast \text{ is a weakly efficient solution}\}
\]

(3.6.3)

As mentioned earlier, most MO-problems does not have a unique efficient solution \(x^\ast\), and therefore not a unique optimal value \([f_1(x^\ast), \ldots, f_n(x^\ast)]\). Instead one talks about the set of nondominated solutions. The set of nondominated solutions is formally introduced as:

\[
\mathcal{H}_\ast = \{f(x) \in \mathcal{H} | x \in \mathcal{F}_\ast\}
\]

(3.6.4)

where \(\mathcal{H} = \{f(x) \in \mathbb{R}^n | x \in \mathcal{F}\}\), i.e. the image of \(f\) under the feasible set.

In figure 3.1, the set of nondominated solutions for a problem with two objective functions is illustrated. We can also see which of the nondominated solutions that are derived from strictly efficient solutions.\(^1\)

\(^1\)Recall that all nondominated solutions correspond to weakly efficient solutions
Two useful concepts are the ideal objective vector and the nadir objective vector. The ideal objective vector is denoted as:

$$z^* = [z_1^*, z_2^*, \ldots, z_n^*]^\top$$

where $z_i^* = \underset{x \in F}{\text{minimize}} f_i(x)$. The ideal objective vector’s name obviously stems from the fact that it would be the ”best” solution and it defines the lower bound of the nondominated set. As previously discussed there is often a trade-off between functions in multi-objective problems, and therefore $z^*$ seldom belongs to $\mathcal{H}$. The nadir objective vector is denoted as:

$$z^{nad} = [z_1^{nad}, z_2^{nad}, \ldots, z_n^{nad}]^\top$$

where $z_i^{nad} = \underset{x \in F^*}{\text{maximize}} f_i(x)$, i.e. $z_i^{nad}$ is the ”worst” value of $f_i$ on the nondominated set and $z^{nad}$ defines the upper bound of the nondominated set. We can often compute the ideal objective vector but it is much harder to compute the nadir objective vector and an approximation is almost always needed.

For the same nondominated set as in figure 3.1 the ideal and the nadir objective vector are illustrated.

Figure 3.1: Classifying the efficient solutions

Figure 3.2: The ideal and nadir objective vector
Chapter 4

Portfolio optimization

4.1 Introduction to MPT

This section introduces the fundamental concepts of Modern Portfolio Theory, in a rather unconventional (but enlightening) way. The section builds mainly upon Steuer, Qi, et al. (2008). The theory is presented in a more rigorous manner compared to Markowitz (1952).

4.1.1 Initial stochastic programming problem

The problem of portfolio selection can in its most basic form be described as follows:

- A fixed sum of money to be invested
- A universe consisting of \( n \) securities
- A single holding period, with a predetermined beginning and end

We utilize the notation outlined in section 3. The portfolio weight for the \( i \):th security, \( x_i \), was defined in equation 3.1.4. Since this is a proportion of the fixed sum to be invested, we know that the \( x_i \)'s must sum to 1. This will thus be a constraint for the optimization problem. In equation 3.1.2 the random variable for the \( i \):th security’s return over the holding period was defined. Note that while the realized returns are unknown until the end of the holding period, it is assumed that all means \( \mu_i \), variances \( \sigma_{ii} \) and covariances \( \sigma_{ij} \) are known in the beginning of the holding period. How to estimate these parameters will be presented in section 4.3.

The random variable for the return of the portfolio is (as in equation 3.1.5):

\[
R_P = \sum_{i=1}^{n} x_i R_i = x^\top R
\]

If we assume that the investor is only interested in maximizing the (uncertain) return of a portfolio, the portfolio selection problem is the following optimization problem:

\[
(\text{SP}) \begin{cases} 
\text{maximize} & R_P = x^\top R \\
\text{subject to} & x \in S 
\end{cases}
\]
Where $S$ denotes the feasible region:

$$S = \{ x \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = 1, \alpha_i \leq x_i \leq \beta_i \}$$  \hspace{1cm} (4.1.1)$$

$\alpha_i$ and $\beta_i$ represent limits for the portfolio weight for each security. A notable case is $\alpha_i = 0$, i.e. the case when no short positions\(^1\) are allowed. (SP) looks like a linear programming problem, but it is in fact a stochastic programming problem since the $R_i$:s are random variables and the composition of the portfolio (i.e. the vector of portfolio weights, $x$) must be determined at the beginning of the holding period. Let (SP) be called the investor’s initial stochastic programming problem.

### 4.1.2 Deterministic formulation

The initial stochastic programming problem is not solvable with the optimization methods used in this thesis. The solution to a stochastic programming problem is not well defined, and is thus difficult to solve.

Hence, to solve (SP) one needs an interpretation and a decision. A common approach utilized in the literature is to transform the stochastic problem to an equivalent deterministic problem. Typically, these formulations make use of some characteristics of the certain random variables or some statistical characteristics. To transform (SP) into a simplified deterministic problem, Steuer, Qi, et al. (2008) reviews some historical results:

In the early 17-th century, mathematicians assumed that a gambler would be indifferent between a stochastic outcome of a gamble and receiving its expected value in cash. In the setting of portfolio selection, this assumption implies that the investor in the situation of portfolio selection is indifferent between holding a portfolio of stocks, or receiving its certainty equivalent:

$$CE = E[R_P]$$

Now, given that the investor wants to maximize $CE$, the deterministic formulation becomes:

$$\begin{cases} 
\text{maximize} & E[R_P] \\
\text{subject to} & x \in S
\end{cases}$$  \hspace{1cm} (4.1.2)$$

However, what is known as the St. Petersburg Paradox was discovered by Bernoulli. The paradox provides an example of a gamble with infinite expected value, but that a gambler (in reality) would be willing to abstain from for a small amount of money. Stated differently: according to the "old" theory a rational player would (or perhaps should) be willing to pay an infinite amount of money to participate in the gamble, but in reality this is not the case. Following this, Bernoulli suggested that instead of comparing cash outcomes, one should compare the "utilities" of different outcomes. Set the utility of a cash outcome as a function $U : \mathbb{R} \to \mathbb{R}$. Then we have:

$$U(CE) = E[U(R_P)]$$

\(^1\)a short position or a short sell is the sale of a borrowed security, expecting that the asset will fall in value, thus making the investor a potential profit
In other words, the expected utility of an uncertain portfolio equals the utility of the certainty equivalent, $CE$. For a given utility maximizing investor, the problem becomes:

\[
\begin{align*}
(UP) \quad \text{maximize} & \quad E[U(R_P)] \\
\text{subject to} & \quad x \in S
\end{align*}
\]

This problem will be referred to as the "undetermined" deterministic problem, since the utility function and its parameters are unknown, and the problem cannot be solved in its current form.

However, investors are assumed to prefer the expected value $E[R_P]$ over the uncertain outcome of $R_P$, i.e. they are assumed to be risk-averse, and thus $U$ is a concave function. Two schools of thought have emerged to deal with the undetermined nature of the utility function. One, following Roy, tries to establish details of an investor’s preference structure and then uses that information to solve (UP), directly being able to determine the optimal portfolio. The other one, following Markowitz, includes a parametrization of $U$ and then attempts to solve (UP) for all possible values of the (unknown) parameters of the utility function. Markowitz considered the following parameterized quadratic utility function:

\[
U(x) = x - (\lambda/2)x^2
\]

(4.1.3)

$U(x)$ is a normalized function such that $U(0) = 0$ and $U''(0) = 1$. This leaves just one parameter, $\lambda$, the coefficient of risk aversion. Using this parameterization, Markowitz could show that all potential optimal solutions to (UP) for a risk-averse investor can be obtained by solving the deterministic problem:

\[
(DP) \quad \begin{align*}
\text{maximize} & \quad E[R_P] \\
\text{minimize} & \quad Var(R_P) \\
\text{subject to} & \quad x \in S
\end{align*}
\]

The set of all optimal solution vectors $x \in S$ are called the efficient set, and the set of all images of the efficient points are called the nondominated set (see section 3.6). Steuer, Qi, et al. (2008) concludes that with the utility function as in equation 4.1.3, the best deterministic formulation is (DP). The process outlined in this section is displayed in figure 4.1.
Now using equation 3.1.6, 3.2.2 and 4.1.1, (DP) is reformulated as:

\[
\begin{align*}
\text{maximize} & \quad \mu^\top x \\
\text{minimize} & \quad x^\top C x \\
\text{subject to} & \quad \sum_i^n x_i = 1 \\
& \quad \alpha_i \leq x_i \leq \beta_i, \ \forall i \in \{1, \ldots, n\}
\end{align*}
\]

In this form the problem is solvable using suitable optimization methods. Doing this generates the efficient frontier\(^1\), as displayed in figure 4.2\(^2\)

![Efficient frontier](image)

**Figure 4.2: Efficient frontier**

### 4.2 Portfolio selection with multiple criteria

When investors have multiple-argument utility functions, a multiple criteria formulation of the portfolio optimization process is appropriate. Steuer, Qi, et al. (2008) names two situations when this is likely the case:

- When investors have additional concerns (in excess of portfolio return), such as maximizing the ESG rating among the stocks included in the portfolio or wanting to minimize the number of securities included in the final portfolio. Thus, instead of wanting only to maximize the stochastic objective of portfolio return, the investor wants to optimize some combination of stochastic and deterministic objectives.

- When investors are resistant to accepting the assumption that all means, variances and covariances can be considered known at the beginning of the holding period. To mitigate this, the investor might be interested of adding measures such as R&D spendings, dividends or growth in sales in the portfolio optimization process.

One could imagine a vast amount of objectives to be included in the optimization problem. However, which criteria that might be included in the portfolio selection process is a

\(^1\) We are aware that this name may cause some confusion for the reader since efficient usually refers to the variables and not the objective functions. The efficient frontier is part of the nomenclature in finance and can in this thesis be substituted salva veritate with the nondominated set.

\(^2\) Note that the frontier is “flipped” when comparing to figure 3.1, this is due to one of the objective functions (Expected Return) being maximized instead of minimized.
question for the individual investor, matching individual preferences. Note that one needs to classify objectives as stochastic or deterministic. In the case of stochastic objectives, one needs to determine a deterministic interpretation.

Some measures, e.g. R&D, could be argued to belong to either category. Indeed the R&D spendings during the holding period is stochastic, but one could claim that investments before the holding period are the most relevant, thus enabling the measure to be handled as deterministic.

However, other objectives does not have some characteristic that enables an interpretation similar to the one mentioned above. In these cases, one could employ the process outlined in section 4.1.2, that is assigning a mean-variance pair to each stochastic objective. However, for practical reasons, when variations in the stochastic objectives are of minor magnitude (or importance), one could represent the objectives by just taking the expected value (compare with equation 4.1.2). If possible, this is very advantageous, since only the means are to be estimated. Thus effectively decreasing the burden of data collection, and the associated problem of estimating eventual covariances.

Now if we denote some additional criteria \( z_1, \ldots, z_j \) and allow the exclusion of variances in these objectives, the optimization problem (to some extent analogous to (DP) above) is the following:

\[
\begin{align*}
\text{maximize} & \quad E[R_P] \\
\text{minimize} & \quad Var(R_P) \\
\text{maximize} & \quad E[z_1] \\
\quad & \vdots \\
\text{maximize} & \quad E[z_j] \\
\text{subject to} & \quad x \in S
\end{align*}
\]

After employing a suitable optimization method, one retrieves a set of efficient solutions (see section 3.6). For the special case when only one additional criteria is added, the efficient frontier in figure 4.2 extends to a surface in the three dimensional space. However, the concept of an efficient frontier can be generalized into any number of dimensions. Thus giving the efficient frontier as a hypersurface.

4.3 Estimating parameters

In this section we will review two ways of estimating the traditional parameters used in MPT. First we present an in some sense obvious way of doing this. Then we present a better-performing method, namely the single-index model.

4.3.1 A naïve approach

In most situations (and throughout this thesis) the probability distributions of stock returns are unknown. Without that knowledge, in what ways is it possible to estimate and compare risk and return? According to Berk and DeMarzo (2014) a popular approach is to extrapolate from historical data. It is claimed that this is a sensible strategy if the environment is stable and it is believed that the distribution of future returns should mirror that of past returns.
Computing historical returns

The realized return is the return of all possible returns that actually occurs over a certain time period. By using equation 3.1.2 we can compute the realized return for a security over a chosen time period. Denote the realized return for security \( i \) and time period \( j \) as \( r_{i,j} \). Then by repeating this for a \( n \) time periods we can calculate the average return of a security:

\[
\bar{r}_i = \frac{1}{n} \sum_{j=1}^{n} r_{i,j}
\]

Note that this is the balancing point of the empirical distribution. Thus, if the probability distribution is the same over time, the average return is an estimate of the expected return.

Estimation of covariance from historical data

We are interested in finding the risk of a portfolio. Thus we need to know to which degree the securities in the portfolio face common risks and how their returns move together. The covariance is a measure that allows us to measure this. If the probability distribution for each security was known we could use equation 3.2.2. However, since this is not the case, we want to use historical data to estimate the covariance matrix. Hence we use the following:

\[
\text{Cov}(r_i, r_j) = \frac{1}{T-1} \sum_{t} (r_{i,t} - \bar{r}_i)(r_{j,t} - \bar{r}_j) \quad (4.3.1)
\]

Which is done for all combinations \( i, j \) of securities. We will refer to this as the historical covariance matrix.

4.3.2 The single-index model

The method presented in this section is based on Elton (2014), where the authors assert that it is a widely used technique. Furthermore, it is claimed that the covariance matrix estimated using the single-index model performs better as an estimation of future covariances compared to the historical covariance matrix (equation 4.3.1).

By just observing stock prices, one sees that individual stock prices (in general) tend to increase when the market goes up, and decrease when the market goes down. This suggests that correlation between stocks to some extent stems from their correlation with the market. A useful measure of this correlation can be obtained by relating the stock return to the market return. We can write the return on a stock as:

\[
R_i = a_i + \beta_i R_m \quad (4.3.2)
\]

where \( a_i \) is a random variable, the component of security \( i \)'s return that is independent of the performance of the market. \( R_m \) is the return of the market index (random variable). \( \beta_i \) a constant measuring the expected change in \( R_i \) given a change in \( R_m \).

The variable \( \beta_i \) is the in finance well known beta, a measure of a stocks sensitivity to changes in the market\(^1\). With \( \beta_i = 1.5 \) the return for a given stock is expected to increase (decrease) by 1.5% if the market index increases (decreases) by 1%. The variable \( a_i \) is the part of the return of the stock that is independent of the development of the market. We

\(^1\)Note that beta is sometimes defined in terms of excess return, i.e. the return that exceeds the risk-free interest rate
can decompose $a_i$ into two terms. Set $\alpha_i = E[a_i]$ and let $e_i$ be the random element of $a_i$ with $E[e_i] = 0$. We have:

$$a_i = \alpha_i + e_i$$

and thus we can write the return of a stock as:

$$R_i = \alpha_i + \beta_i R_m + e_i$$  \hspace{1cm} (4.3.3)

Note that $e_i$ and $R_m$ are random variables. Denote their standard deviations $\sigma_{e_i}$ and $\sigma_m$ respectively. So far, we have not made any simplifying assumptions. However, it is convenient to have that $e_i$ and $R_m$ are uncorrelated, that is:

$$Cov(e_i, R_m) = E[(e_i - 0)(R_m - \mu_m)] = 0$$  \hspace{1cm} (4.3.4)

If this holds, then how well equation 4.3.3 describes the return of a given security is independent of the market return. With regression analysis, we can ensure that $e_i$ and $R_m$ will be uncorrelated (at least over the period of the data fitted to the equation). A regression analysis also gives estimates of $\alpha_i$, $\beta_i$ and $\sigma_{e_i}^2$.

Up to this point, all the equations are definitions or can be made to hold by construction. Now we need to make an assumption. The assumption is that $e_i$ and $e_j$ are independent for all $i$ and $j$. That is:

$$E[e_i e_j] = 0 \ \forall \ i, j \ (i \neq j)$$

This implies that the only reason for two stocks to move together is because of their common movement with the market. Thus we ignore that there may be other effects (e.g. industry effects) that gives comovement among stocks. There is nothing in the normal regression method that forces this assumption to hold. It is merely a simplifying assumption representing an approximation of reality. Thus the performance of the single-index model depends (in part) of how well this assumption holds.

Now we will derive the expected return, volatility and covariance when the single-index model is used as a representation of the comovement of stocks. First, the expected return\footnote{as before we write $\mu_i = E[R_i]$ and $\mu_m = E[R_m]$} of a security (using equation 4.3.3):

$$\mu_i = E[\alpha_i + \beta_i R_m + e_i] = E[\alpha_i] + E[\beta_i R_m] + E[e_i]$$

$\alpha_i$ and $\beta_i$ are constants, $E[e_i] = 0$. Therefore:

$$\mu_i = \alpha_i + \beta_i \mu_m$$  \hspace{1cm} (4.3.5)

The expected return of a security has two sources, corresponding to the two terms in equation 4.3.5. The first term, $\alpha_i$, is the firm specific (unique) return. The second term, $\beta_i \mu_m$, is the systematic (or index driven) return. Now consider the variance of a security which is given by (as in equation 3.2.1):

$$\sigma_i^2 = E[(R_i - \mu_i)^2]$$

Now, using equation 4.3.3 and 4.3.5:

$$\sigma_i^2 = E[(\alpha_i + \beta_i R_m + e_i - \alpha_i - \beta_i \mu_m)^2] = E[(\beta_i (R_m - \mu_m) + e_i)^2]$$
Furthermore, expanding the square gives:

$$\sigma_i^2 = \beta_i^2 E[(R_m - \mu_m)^2] + 2\beta_i E[e_i(R_m - \mu_m)] + E[(e_i)^2]$$

Note that we assumed that $e_i$ and $R_m$ are uncorrelated (equation 4.3.4), and by definition $E[(e_i)^2] = \sigma_{e_i}^2$. Hence we conclude that:

$$\sigma_i^2 = \beta_i^2 E[(R_m - \mu_m)^2] + E[(e_i)^2] = \beta_i^2 \sigma_m^2 + \sigma_{e_i}^2 \tag{4.3.6}$$

In the same way as for the return of an asset, the variance can be divided into two parts. The first term of equation 4.3.6 is the firm specific (unique) variance. The second term is the systematic (or index driven) variance. Furthermore, by definition, the covariance between the returns of two assets is:

$$\sigma_{ij} = E[(R_i - \mu_i)(R_j - \mu_j)]$$

Using equation 4.3.3 and 4.3.5:

$$\sigma_{ij} = E[\left(\alpha_i + \beta_i R_m + e_i - \alpha_i - \beta_i \mu_m\right)\left(\alpha_j + \beta_j R_m + e_j - \alpha_j - \beta_j \mu_m\right)]$$

$$= E \left[\beta_i(R_m - \mu_m + e_i)(\beta_j(R_m - \mu_m) + e_j)\right]$$

$$= \beta_i \beta_j E[(R_m - \mu_m)^2] + \beta_i E[e_j(R_m - \mu_m)] + \beta_j E[e_i(R_m - \mu_m)] + E[e_i e_j]$$

By assumption the last three terms are equal to zero, thus yielding:

$$\sigma_{ij} = \beta_i \beta_j \sigma_m^2 \tag{4.3.7}$$

The equations 4.3.5, 4.3.6, 4.3.7 can be used to determine the input data needed to perform the portfolio optimization. However, before we can do that, we need to compute estimates of $\alpha_i$, $\beta_i$ and $\sigma_{e_i}^2$ for all $i$, and also $\sigma_m^2$ and $\mu_m$. We will do this in the subsequent section.

**Estimation using historical data**

To use the single-index model, estimates of the beta of each stock is needed. These could be provided by analysts (thus subjective estimates). Another approach is to estimate betas from historical data, and use these historical betas as an estimate of future betas. Berk and DeMarzo (2014) argues that this approach makes sense since the betas of stocks appears to be relatively stable over time. Now recall equation 4.3.3:

$$R_i = \alpha_i + \beta_i R_m + e_i$$

The equation is expected to hold at each time instant, although the values of the parameters might vary over time. When observing historical data, we cannot directly observe either of the parameters. What we can observe is the return on the individual securities and the market. Since equation 4.3.3 should hold at each moment in time, there exists a straightforward procedure to estimate $\alpha_i$, $\beta_i$ and $\sigma_{e_i}^2$.

In fact, 4.3.3 is just a straight line. Thus we can estimate the parameters by performing a linear regression. Denote the return on security $i$ observed over time period $t$ as $r_{it}$, and correspondingly for the return on the market as $r_{mt}$. Then if one plots the stock return versus the market return, we get $\beta_i$ as the slope of the best fitting line and $\alpha_i$ as the intercept, see figure 4.3 for an example.
By best-fitting, we mean the values of $\alpha_i$ and $\beta_i$ that minimizes the sum of squares of the vertical distance for each data point to the line. That is, we compute the OLS\textsuperscript{1} estimate. In the regression, the dependent variable is the returns of security $i$. The independent variable is the market returns $r_{mt}$. Let $\hat{\beta}_i$ be the OLS estimate of $\beta_i$ and correspondingly $\hat{\alpha}_i$ for $\alpha_i$. Then the estimated residuals are given by:

$$\hat{e}_{it} = r_{it} - (\hat{\alpha}_i + \hat{\beta}_i r_{mt})$$

Say that we use $n$ data points, that is use data from $n$ time periods $t$. Then we compute $\sigma^2_{ei}$ as the sample variance of the residuals $\hat{e}_{it}$. Let $\bar{e}_i$ denote the sample mean of the residuals. Hence:

$$\hat{\sigma}^2_{ie} = \frac{1}{n-1} \sum_{t=1}^{n} (\hat{e}_{it} - \bar{e}_i)^2$$  \hspace{1cm} (4.3.8)

Now, we only need to compute $\sigma^2_m$ and $\mu_m$. A straightforward way to do this is estimate $\sigma^2_m$ as the sample variance and $\mu_m$ as the mean of the historical returns. That is:

$$\hat{\sigma}^2_{m} = \frac{1}{n-1} \sum_{t=1}^{n} (r_{mt} - \bar{r}_m)^2$$  \hspace{1cm} (4.3.9)

and

$$\hat{\mu}_m = \frac{1}{n} \sum_{t=1}^{n} r_{mt}$$  \hspace{1cm} (4.3.10)

Another way of doing this is to use estimates provided by analysts. That is, use the analysts’ views of future development of the market. In this thesis however, we will historical data to determine the estimates.

Now we have determined how to estimate $\alpha_i$, $\beta_i$, $\sigma^2_{ei}$, $\sigma^2_m$ and $\mu_m$. Thus we can compute the input data used in MPT. Note that we need to run a linear regression for each security $i = 1, \ldots, N$. This gives $\hat{\alpha}_i$, $\hat{\beta}_i$, $\hat{\sigma}^2_{ei}$ for all $i$. Utilize the following matrix notation:

$$\hat{\alpha}^\top = [\hat{\alpha}_1, \ldots, \hat{\alpha}_N]$$

$$\hat{\beta}^\top = [\hat{\beta}_1, \ldots, \hat{\beta}_N]$$

$$\hat{\sigma}^{2\top}_{e} = [\hat{\sigma}^2_{1e}, \ldots, \hat{\sigma}^2_{Ne}]$$

\textsuperscript{1}Ordinary Least Squares
Then the vector of expected returns is given by (using equation 4.3.5):

$$
\mu = \begin{bmatrix} 
\mu_1 \\
\vdots \\
\mu_N 
\end{bmatrix} = \begin{bmatrix} 
E[R_1] \\
\vdots \\
E[R_N] 
\end{bmatrix} = \begin{bmatrix} 
\hat{\alpha}_1 \\
\vdots \\
\hat{\alpha}_N 
\end{bmatrix} + \begin{bmatrix} 
\hat{\beta}_1 \\
\vdots \\
\hat{\beta}_N 
\end{bmatrix} \times \hat{\mu}_m = \hat{\alpha} + \hat{\beta} \times \hat{\mu}_m 
$$ (4.3.11)

And the covariance matrix (using equation 4.3.6 and 4.3.7):

$$
C = \begin{bmatrix} 
\hat{\sigma}_1^2 & \cdots & \hat{\sigma}_{1N} \\
\vdots & \ddots & \vdots \\
\hat{\sigma}_{N1} & \cdots & \hat{\sigma}_N^2 
\end{bmatrix} = \begin{bmatrix} 
\beta_1^2 \hat{\sigma}_m^2 + \hat{\sigma}_{e1}^2 & \cdots & \beta_1 \beta_N \hat{\sigma}_m^2 \\
\vdots & \ddots & \vdots \\
\beta_N \hat{\beta}_1 \hat{\sigma}_m^2 & \cdots & \beta_N^2 \hat{\sigma}_N^2 + \hat{\sigma}_{eN}^2 
\end{bmatrix} =
$$

$$
= \begin{bmatrix} 
\beta_1^2 \hat{\sigma}_m^2 & \cdots & \beta_1 \beta_N \hat{\sigma}_m^2 \\
\vdots & \ddots & \vdots \\
\beta_N \hat{\beta}_1 \hat{\sigma}_m^2 & \cdots & \beta_N^2 \hat{\sigma}_N^2 
\end{bmatrix} + \begin{bmatrix} 
\hat{\sigma}_{e1}^2 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & \hat{\sigma}_{eN}^2 
\end{bmatrix}
$$

Which can be compactly written as:

$$
C = \hat{\beta} \hat{\beta}^T \hat{\sigma}_m^2 + D(\hat{\sigma}_e^2) 
$$ (4.3.12)

where $D(\hat{\sigma}_e^2)$ is a $n \times n$ diagonal matrix whose $i$:th diagonal element is $\hat{\sigma}_{ei}^2$.

In summary: perform linear regressions to obtain alpha and beta values for each stock. Estimate the expected variance and mean of the market return. Use equation 4.3.11 and 4.3.12 to obtain the vector of expected returns and the covariance matrix.
Chapter 5

Environmental, Social and corporate Governance concerns

5.1 Introduction to ESG

Environmental, Social and corporate Governance (ESG) refers to three broad dimensions of corporate behavior. ESG criteria is a catch-all term for criteria used in Socially Responsible Investing (SRI) \(^1\) (Eurosif, 2012). Berry and Junkus (2013) claims that most definitions of SRI includes: “Integrating personal values and societal concerns with investment decisions”. ESG factors are the performance metrics used to evaluate a company’s performance in the different dimensions. Within the three dimensions, a broad set of factors used to evaluate the sustainability and ethical impact of a corporation are included. Galbreath (2013) asserts that ESG has become key indicators of risk management, management competence and non-financial performance. Some of the ESG concerns commonly addressed are presented in table 5.1 (AP1, 2014 and Galbreath, 2013).

<table>
<thead>
<tr>
<th>Environmental concerns</th>
<th>Social concerns</th>
<th>Governance concerns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Climate change</td>
<td>Human rights</td>
<td>Board independence</td>
</tr>
<tr>
<td>Energy and water use</td>
<td>Labor rights</td>
<td>Corruption &amp; bribery</td>
</tr>
<tr>
<td>Carbon emissions</td>
<td>Product safety</td>
<td>Reporting &amp; disclosure</td>
</tr>
<tr>
<td>Recycling</td>
<td>Gender equality</td>
<td>Shareholder protection</td>
</tr>
<tr>
<td>Product design</td>
<td>Health and safety</td>
<td>Business Ethics</td>
</tr>
<tr>
<td></td>
<td>Fair trade</td>
<td>Executive compensation</td>
</tr>
<tr>
<td></td>
<td>Community involvement</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.1: Examples of ESG factors

A distinction must be made, the use of ESG criteria can be analyzed from (at least) two different perspectives. One can view it from a company perspective or from an investor perspective. (This thesis focuses on the investor and therefore the focus will be on the latter.) To clarify the difference, a company might ask itself ”How can our company work with and evaluate ESG criteria?” whereas an investor might ask ”How would I benefit from the use of ESG criteria in the portfolio selection process?”. Still the company perspective

\(^1\)There is some confusion regarding this acronym, SRI is also known as “Sustainable and Responsible Investment” etc.
is of importance since the regular investor is limited to public information and therefore confined by how and what companies report.

5.2 Investors and ESG

Surveys have showed that an increasing amount of investors care about ESG criteria. With one survey showing that two in three of large asset owners in the Nordic countries use ESG criteria (Novethic, 2013a). In the UK, 57 percent of investors believe that ESG is a factor in financial performance (BNP Paribas, 2010). The UN Principles for Responsible Investment (UNPRI) has 1200 signatories managing in total assets worth $34 trillion. The signatories agree to abide by six principles, where the first one is: "We will incorporate ESG issues into investment analysis and decision-making processes" (UNPRI, 2014c).

In light of this, the crucial question is: Are ESG issues relevant to investors? In this section we will argue that this is the case, using two arguments:

- Decreased risk, increased return
- Responsibility towards society

UNPRI (2014b) asserts that the answer is an unambiguous yes. Their argument falls within the first category. As evidence, the financial impacts on corporations after scandals (e.g. BP and the Deepwater Horizon spill, Barclays after the Libor scandal) are stressed. Thus, the argument is that corporations monitoring ESG issues are associated with decreased risk. Eccles et al. (2012) found in a study of US companies, that High Sustainability companies significantly outperform their counterparts (i.e. Low Sustainability) in the long run, in terms of stock market performance. An overview of previous research showed that there is some evidence that the inclusion of ESG factors in the portfolio selection does not impose a penalty on the performance (i.e. return) of the portfolio (UNEP FI, 2007). However, it should be noted that in this thesis the main purpose of including ESG factors is to reflect investor concerns, not as an indicator or proxy of future return.\footnote{Thus, this relationship should be taken as a “bonus” since the extra expected return lies outside the mathematical framework used in the thesis. Investors wanting to use ESG in this framework to achieve higher return can of course do it, but the effects are not quantified. (Since the ESG effects on mean, variance and covariance cannot be estimated using the methods we have used.)}

In the global public perception a number of real and present dangers have emerged. To mention a few: climate change, fragile ecosystems and social impacts of urban and industrial pollution (UNEP FI, 2007). Having concerns about these and similar issues falls within the first category. The arguments for why investors should care about these issues are almost philosophical in nature. The question for investors is then to which extent these concerns should affect the portfolio selection.

5.3 Companies and ESG

As previously mentioned there has been an increasing interest in ESG factors. A 2010 UN Global Compact survey found that of 100 global CEOs and CFOs 93 percent believed that sustainability would be critical for their business in the future (Lacey et al., 2010). Increasing investor and regulatory demand are driving forces but also assumed correlation between financial and ESG performance.
A regulatory push for ESG criteria can interestingly be seen in emerging markets such as South Africa, India and Brazil, which have all started to demand primarily sustainability reporting from their listed companies (Jung, 2013). In Sweden the government has adopted guidelines demanding that state-owned companies issue sustainability reports in accordance with GRI’s guidelines. Exceptions are possible and are based on the principle “comply or explain”. Nasdaq OMX, the owners of the largest stock exchange in Sweden, also owns the Copenhagen exchange and there enforces a similar “comply or explain” principle for corporate-governance (Martin, 2012).

Demand for ESG reporting is also increasing with surveys as (Novethic, 2013a) showing that two thirds of large asset owners in the nordic countries use ESG criteria. According to Bean (2013) “ESG-conscious investors are important for ensuring a continued source of working capital”. The same authors argue that in an environment where the number of investors and tools used for evaluating investments are increasing, the importance of high ESG ratings is increasing as well. One should note that even though awareness of the increased importance of ESG ratings, only 2-3 percent of all public conference calls done at the time of quarterly reporting in the USA mentioned sustainable development (UNEP FI and WBCSD, 2010).

The last “why factor” is the correlation between financial performance and “good” ESG factors. Fulton et al. (2012) reviews several studies which have examined the correlation and finds that 100 percent of the reviewed studies agrees that companies with high ESG (and/or CSR) ratings have lower costs of capital in debt and equity. The report also finds that 85 percent of the reviewed studies show that firms with high ESG ratings shows higher accounting based performance. One should note that a vast majority of the reviewed studies only focused on one factor, i.e. E, S or G.

How companies should work with ESG performance is at large a question that is company specific. UNEP FI and WBCSD (2010) found through a series of workshops with companies and investors that there should be clear links between ESG factors and strategy. There are studies showing general things that companies can try to excel in, e.g. Edmans (2011) found that companies with high levels of employee satisfaction outperform their counterparts in the long run.

A more general question is how ESG should be reported. Formal corporate reporting outside the required (under the Accounting Standards) financial reporting is commonly referred to as ESG reporting. The number of international corporations that publish standalone non-required reports is increasing. CSR reporting has become the norm instead of the exception within the world’s largest corporations. Between 2005 and 2013, companies issuing CSR reports increased from 52 to 93% (Boer et al., 2013). This increase in ESG reporting can be interpreted as that ESG issues are increasingly important for corporations (Murphy and McGrath, 2013).

In some financial areas a common language has been adopted to define certain terms. An example is the International Financial Reporting Standards (IFRS), where terms as asset, liability and equity are defined. Resulting in stakeholders being able to evaluate a company using certain performance metrics and then being able to compare different companies. Currently there is no global standardized approach stakeholders can use for evaluating companies ESG performance. In the social and environmental accounting (SEA) literature and in practice the reports used to report ESG work and ESG factors in-

---

1The Global Reporting Initiative is a not-for-profit driven organization promoting the use of sustainability reporting. Which in part is done by providing reporting guidelines
clude but are not limited to: sustainability reports, corporate social responsibility (CSR) reports, global reporting initiative (GRI) reports, corporate social disclosure (CSD) reports and triple bottom line (TBL) reports. The 2010 UN Global Compact Survey found that 49 percent of the interviewed respondents cited the complexity of implementation as the most significant barrier for ESG reporting. In the response to problems voiced in this and other surveys several reporting frameworks have been created. According to Cuesta and Valor (2013) the GRI guidelines can be considered as a good instrument for reporting ESG criteria. Another valuable guideline is the “KPI’s for ESG” published by DVFA 1. These can be incorporated in the GRI framework and provide general as well as sector specific ESG factors DVFA (2010).

5.4 Ratings

Increasing interest in ESG combined with a lack of standardization has led to the formation of several indices and rating agencies. The French rating agency Novethic has analyzed the ESG rating market two times (Novethic, 2013b) and (Novethic, 2011). Novethic has found that the ESG rating market has seen a consolidation in terms of agencies, but the market is still dynamic and the offering is becoming more and more diverse. In broad strokes the market for ESG data for investors can be divided into two parts: companies providing ratings and companies providing raw data. Raw data providers give investors the possibility of creating their own ratings by providing data. Companies providing ratings have quantified the raw data to measurable key performance indicators (KPI) or rankings. In this thesis we have limited our scope to the latter category and specifically data that is free of charge. The rating agencies usually combine a company’s ESG policies and the implementation, performance and reporting of said policies. This is illustrated in figure 5.1.

![Figure 5.1: The components of an ESG rating](image)

During our research we have found two different rankings/indices that can be accessed and used free of charge for all investors. The two are: Folksam’s Index for Sustainable Business and STOXX ESG Index. The methodologies and scopes of these two differ in some aspects, which we will try to describe below. Both indices use the UN Global

---

1 Deutsche Vereinigung für Finanzanalyse und Asset Management. German abbreviation for: Society of Investment Professionals in Germany
Compact Ten Principles in some way, these principles can be seen in appendix C.1. In the sections below the data given from one provider constitutes an index and one of the different grades a company receives in an index constitutes a rating. E.g. the index is Folksam’s Index for Responsible Business, in which each company receives two ratings: (1) Environmental and (2) Social.

5.4.1 Folksam’s Index for Responsible Business

This section will build on Lundberg Markow and Bönnelyche (2013). Folksam is one of Sweden’s largest insurance companies and institutional owners. Folksam’s corporate governance rules states that the companies in which Folksam has an ownership share should actively work towards a sustainable and ethical world. The purpose of this is to create more value for its customers. This has led to the publication of the Folksam’s Index for Responsible Business, which is a ranking of the companies listed on the Stockholm Stock Exchange. The ranking uses a 0-7 grading scale. Each company is evaluated using a scheme similar to that of figure 5.1. The criteria being evaluated are derived from Global Compact, OECD and GRI guidelines with data from GES Investment Services. The ranking evaluates the listed companies capacity to handle:

- Environmental aspects, divided in two main categories: (1) A general assessment of the environmental guidance system and (2) governance of direct or indirect environmental impact.
- Social aspects, divided in three main categories: (1) Employee rights, (2) The company in society and (3) Human rights in the supply-chain (excluded for some service companies).

Folksam has divided the companies into sixteen different sectors. There are both sector specific and general criteria related to environmental and social concerns. How these are weighted is not made public. It is also noteworthy that environmental concerns are one rating but social and governance concerns are lumped together to one rating. How this affects the ratings is hard to say since specific KPIs are not made public. It seems as if social concerns are a bigger part of the second rating then governance. Besides the ranking Folksam has found that an increasing amount of companies work actively with environmental and social responsibilities. The OMXS30 companies in general have higher then average ratings. Of the companies listed on the OMXS30 all companies except Nokia Oyj are included\(^1\). The OMXS30 companies have a mean score of 4.16 in environmental rating and 4.12 in the social rating. As can be seen in appendix C.3, amongst the OMXS30 companies the best rated company is Volvo in the environmental rating and Ericsson in the social rating\(^2\). The lowest rated in the environmental category is Investor and Tele2 has the lowest social rating of the OMXS30 companies.

5.4.2 STOXX ESG Index

This section builds on STOXX (2014). STOXX is an index provider owned by Deutscher Börse Group and SIX Swiss Exchange. STOXX provides several specialized indices; of interest in this thesis is the STOXX Global ESG Leaders. The companies are selected

\(^1\)We suspect that this is due to Nokia Oyj having its primary listing on the Helsinki Stock Exchange
\(^2\)These two are actually best of all companies listed on the Stockholm Stock Exchange
from the STOXX Global 1800 Index, containing 600 European, 600 American and 600 Asia/Pacific region stocks. Excluded are companies not complying with the Global Compact Principles and companies dealing with “controversial” weapons. The used KPIs are provided by the ESG rating firm Sustainalytics and are closely correlated with the DVFA KPIs mentioned in section 5.3. Companies are divided into 42 different industries, where each industry has sector specific KPIs and for which the general KPIs are weighted differently. The KPIs come from three different categories: (1) Environmental, (2) Social and (3) Corporate Governance, these three are also the given ratings.

The scores are normalized and each company receives a value between 0-100 in each rating. Of the OMXS30 companies all but SSAB were reported. The mean scores for the OMXS30 companies were 72.2 in the environmental rating, 74.3 in the social rating and 68.3 in the corporate governance rating. As can be seen in appendix C.3, the highest rated companies were: Getinge in the first, Atlas Copco in the second and SEB in the third. The lowest rated were in the same order as before: Investor, Tele2 and Telia Sonera. Noteworthy is that Investor had the lowest environmental rating in the Folksam’s Index for Responsible Business as well and Tele2 which had the lowest social ranking in Folksam’s also had the lowest here. Tele2 does not have the lowest ranking in the Corporate Governance rating, which strengthens our theory of the Folksam Index having a greater emphasis on Social rather then Corporate Governance. Volvo which had the best environmental rating in the Folksam Index is eclipsed by twenty companies in this ranking.

All in all the difference in rating between the two indices strengthens our decision to use two indices rather than one:

- As can be seen in appendix C.3, of the top five from each index there was almost no overlap
- There was one company missing in each index.
- Different methods were used when constructing the indices
- Indices were created from a different number of available shares
Chapter 6

Optimization methods

In section 2.2.1 we used a classification method from Cohon (1985) to classify methods used for solving multi-objective optimization problems:

1. Methods that generate the set of efficient solutions, the decision maker is then presented with this set and chooses one solution. In these methods the decision maker is involved after the optimization.

2. Preference-based methods, these methods take into consideration the preferences of the decision maker before making or during optimization. The generated solution is then the one that comes closest to satisfying these preferences.

Note that for the methods that generate the set of efficient solutions we still need a basic preference, i.e. one needs for every objective function if it is to be maximized or minimized. As was mentioned in section 2.2.1 the two methods presented below are usually said to belong to the first category. For a thorough introduction see e.g. Miettinen (1998) or Ehrgott (2005). In the following sections we will assume, without loss of generality, that all objective functions shall be minimized\(^1\).

### 6.1 Weighted sum approach

There are several ways one can compute the efficient solutions to (MO), i.e. the set \(\mathcal{F}_s\). The theoretically easiest method is the so-called weighted sum approach. The weighting problem to (MO) is defined as:

\[
(WP) \begin{cases} 
\text{minimize} & \sum_{i=1}^{n} \lambda_i f_i(x) \\
\text{subject to} & x \in \mathcal{F} 
\end{cases}
\]  

Where \(\sum_{i=1}^{n} \lambda_i = 1\) and \(\lambda_i \geq 0\) Solutions to (WP) are weak efficient solutions, this is proved using a contradiction. Let \(x_\star\) be a solution (WP) and assume that \(x_\star\) is not a weak efficient solution to (MO). According to equation 3.6.1 there then exists a point \(\tilde{x}\) such that \(f_i(\tilde{x}) < f_i(x_\star) \forall i \in \{1, 2, \ldots, n\}\) which implies that: \(\sum_{i=1}^{n} \lambda_i f_i(\tilde{x}) < \sum_{i=1}^{n} \lambda_i f_i(x_\star)\) since the weights are non-negative. This is an obvious contradiction, and \(x_\star\) has to be a weak efficient solution to (MP).

\(^1\)Since maximizing a function \(f(x)\) is the same as minimizing \(-f(x)\)
If we in equation 6.1.1 set \( \lambda_i > 0 \), a solution to (WP) is a strictly efficient solution to (MO). By the same assumption as above but now using equation 3.6.2 we have that: 
\[
\sum_{i=1}^{n} \lambda_i f_i(\tilde{x}) < \sum_{i=1}^{n} \lambda_i f_i(x^*)
\]
This implies that: \( \sum_{i=1}^{n} \lambda_i f_i(\tilde{x}) < \sum_{i=1}^{n} \lambda_i f_i(x^*) \), again the contradiction is obvious since all weights are strictly positive, and \( x^* \) has to be a strictly efficient solution to (MP).

In the case of two objective functions the weighted sum approach can be pictured in an intuitive fashion. We try to find \( y \) such that the line just touches the image of the objective functions, moving outward from the origin. The weighting parameters \( \lambda_1 \) and \( \lambda_2 \) defines the line’s slope, and determines which nondominated solution that will be found. The nondominated set is then generated by solving (MO) for different pairs of \( \lambda_1 \) and \( \lambda_2 \). The result can be seen below.

![Figure 6.1: WSM with two objective functions](image)

### 6.2 \( \varepsilon \)-Constraint Method

#### The method

The \( \varepsilon \)-Constraint Method is another simple method used for computing efficient solutions to (MO). If we have \( n \) objective functions the \( j \):th \( \varepsilon \)-constrained problem is defined as:

\[
\text{(EP}_j\text{)} \left\{ \begin{array}{l}
\text{minimize} \quad f_j(x) \\
\text{subject to} \quad f_k(x) \leq \varepsilon_k, \forall k \neq j \\
x \in \mathcal{F}
\end{array} \right.
\]  

(6.2.1)

i.e. the \( \varepsilon \)-constrained method amounts to solving \( \text{EP}_j \) for \( j = \{1, \ldots, n\} \). If \( x^* \) is a solution to the \( \varepsilon \)-constrained method then it is weak efficient solution, again this is proved using a contradiction. Assume that \( x^* \) is optimal to equation 6.2.1 but not a weak efficient solution. Then according to equation 3.6.1 there exists a solution to equation 6.2.1, such that \( f_k(x) < \varepsilon_k \), \( k = \{1, \ldots, n\} \), \( k \neq j \), i.e. \( x \) is feasible and \( f_j(x) < f_j(x^*) \), contradicting the optimality of \( x^* \).

Next we show that a strictly efficient solution, \( x^* \), has to be a solution to \( \varepsilon \)-constrained method with \( \varepsilon_k = f_k(x^*) \) for \( k = \{1, \ldots, n\} \). If this does not hold there exists a feasible solution, \( x \), such that \( f_j(x) < \varepsilon_j \) and \( f_k(x) \leq \varepsilon_j \) for all \( k \neq j \) for some \( j \). This obviously contradicts \( x^* \) optimality according to equation 3.6.2.
A useful property of the $\varepsilon$-constrained method is that a unique solution to it, has to be a strictly efficient solution. Assume that $x_*$ is a unique solution to equation 6.2.1 for some $j$, but not a strictly efficient solution. But since $x_*$ is unique, there cannot exist a solution $x$ such that $f_i(x) \leq f_i(x_*)$ for all $i$ and $f_m(x) < f_m(x_*)$ for at least one $m$. Hence $x_*$ is a strictly efficient solution. Now note that this simplifies the $\varepsilon$-constrained method a lot. In this thesis we will only be working with problems where at least one function (the covariance function) is strictly convex. In subsection 3.4, we mentioned that if a strictly convex problem has an optimal solution, it is unique. So if we find a solution to equation 6.2.1, the solution has to be unique, implying two things:

1. We only have to solve one optimization problems ($\text{EP}_j$) to generate an efficient solution. The $j$ we solve for can then be chosen such that we only solve the simplest ($\text{EP}_j$).

2. Since the solution is unique, the solution is strictly efficient.

The nondominated set is then computed by solving an ($\text{EP}_j$) for different values of the $\varepsilon_i$’s. The method is illustrated in the two-objective case below. The intuition is that one fixes one objective value and optimizes the other.

![Image of the objective functions](image)

Figure 6.2: $\varepsilon$CM with two objective functions

**Early exits with the $\varepsilon$-constraint method**

When using the weighted sum approach (for convex objective functions) if weights are chosen in a way such that $w_i \geq 0$ and $\sum_{i=1}^n w_i = 1$ every problem will have a solution. With the $\varepsilon$-constraint method this is not guaranteed since when all $\varepsilon_i$’s are small there might not be a feasible optimal solution. The resulting problem is doing a lot of unnecessary iterations that lack solutions. To eliminate this problem Mavrotas (2009) proposes one should start with large $\varepsilon_i$’s (where we are almost guaranteed to have a feasible optimal solution) and gradually decrease them. When a solution became unfeasible we brake out of that loop, since if for a fixed $\varepsilon_2$ there exists no feasible solution for $\varepsilon_3^{(1)}$ there will not exist a feasible solution for $\varepsilon_2$ and $\varepsilon_3^{(2)}$, where $\varepsilon_3^{(2)} < \varepsilon_3^{(1)}$ to the problem.
\[
\begin{align*}
\text{minimize} & \quad f_1(x) \\
\text{subject to} & \quad x \in \mathcal{F} \\
& \quad f_2(x) \leq \varepsilon_2 \\
& \quad f_3(x) \leq \varepsilon_3
\end{align*}
\]

6.3 Computing the ideal objective vector and approximating nadir values

One of the practical problems the weighted sum approach has is choosing the step size for each weight, i.e. how much should a specific weight \( w_i \) change between two iterations. In general it is not recommended to have the same step size for each objective functions since the magnitude of their individual ranges differ. One way to handle this problem is to normalize the individual objective functions. Miettinen (1998) proposes that this could be done in the following way:

\[
f_{i}^{\text{norm}}(x) = \frac{f_i(x) - z_i^*}{z_{i}^{\text{nad}} - z_i^*}
\]

(6.3.1)

This implies that \( f_{i}^{\text{norm}}(x) \in [0, 1] \forall x \in \mathcal{F}_* \). The nadir objective vector and the ideal objective vector are also needed when using the \( \varepsilon \)-constraint method since we need lower and upper bounds for the \( \varepsilon_i \)'s.

In section 3.6 we showed how the ideal objective vector can be computed, we also mentioned that the nadir objective vector almost always has to be approximated. Mavrotas (2009) uses a lexicographic approach for approximating the nadir objective vector for multi-objective problems. We use the same method for our problem:

1. Compute the ideal objective value for function \( i \):

\[
\begin{align*}
\text{minimize} & \quad f_i(x) \\
\text{subject to} & \quad x \in \mathcal{F}
\end{align*}
\]

The answer is the ideal objective value \( z_i^* \)

2. Compute the optimal value of another function \( j \), \( j \neq i \) when adding \( f_i(x) = z_i^* \) as a constraint:

\[
\begin{align*}
\text{minimize} & \quad f_i(x) \\
\text{subject to} & \quad x \in \mathcal{F} \\
& \quad f_i(x) = z_i^*
\end{align*}
\]

We denote the optimal value of this problem \( z_i^{*i} \)

3. Compute the optimal value of the last function function \( k \), \( k \neq i, j \) when adding \( f_i(x) = z_i^* \), \( f_i(x) = z_i^* \) and \( f_j(x) = z_j^{*j} \) as a constraint:

\[
\begin{align*}
\text{minimize} & \quad f_i(x) \\
\text{subject to} & \quad x \in \mathcal{F} \\
& \quad f_i(x) = z_i^* \\
& \quad f_j(x) = z_j^{*i}
\end{align*}
\]
We denote the optimal value of this problem $z^*_{kij}$

4. Repeat 1-3 for different values of $i$, $j$ and $k$.

The results are then put in a table where each row is the results from the optimizations above and each column the results from all optimizations above for an objective function. The ideal objective value for function $i$ is then the minimum value in column $i$ (the diagonal) and the nadir objective value for function $i$ is approximated as the maximum of column $i$. The idea behind this approximation is that the nadir values will be found where other functions reach their "best" values. This might not be the case and in the implementations we have added a small factor to the approximated nadir values. It should be noted that the order in which the optimization is done can affect the approximated values. An example of this type of table is shown in table 6.1.

<table>
<thead>
<tr>
<th>Function $i$</th>
<th>$f_1$</th>
<th>$f_2$</th>
<th>$f_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>min $f_1$</td>
<td>$z^*_1$</td>
<td>$z^*_1'$</td>
<td>$z^*_1''$</td>
</tr>
<tr>
<td>min $f_2$</td>
<td>$z^*_2$</td>
<td>$z^*_2'$</td>
<td>$z^*_2''$</td>
</tr>
<tr>
<td>min $f_3$</td>
<td>$z^*_3$</td>
<td>$z^*_3'$</td>
<td>$z^*_3''$</td>
</tr>
</tbody>
</table>

Table 6.1: Estimating nadir values
Chapter 7

Implementation

7.1 Parameter estimation

The first step in the process of determining the efficient frontier is to estimate the parameters used in the model. Expected return (a vector denoted \( \mu \)) and the covariance matrix (denoted \( C \)) will be estimated using the single-index model described in 4.3.2. By performing linear regressions using MATLAB’s function `regress`, we obtain the alpha and beta values for each stock. Combined with the residuals from the regressions and some estimate of the future market return and volatility, we can calculate \( \mu \) and \( C \).

To include ESG concerns in the portfolio selection one needs to argue for how it is to be done. In section 4.2 we presented some theory regarding how to include additional criteria in the portfolio optimization. Since we want to maximize the ESG performance of companies during the holding period we face the same problem as for return, i.e. the problem is stochastic. One could assign a mean-variance pair for ESG (in the same way as return is implemented in Markowitz’s model) for which one would want to maximize \( E[ESG_P] \) and minimize \( Var(ESG_P) \). Where \( ESG_P \) denotes the ESG value of the portfolio.

This, however, becomes very difficult since it requires us to estimate both the future variance and covariance for the ESG ratings of the stocks. Moreover, we argue that variations in the ESG rating are of less importance (and of minor magnitude) compared to return, since it is reasonable to assume that the ESG performance of a company is relatively stable over time. This enables us to only use the expected value of the ESG objective.

Let \( ESG_i^{(t)} \) be the ESG performance of company \( i \) during time period \( t \). Let time period \( t + 1 \) be the holding period, which the portfolio optimization precedes. Now, if we assume that the ESG rating of a company possess the martingale property, then the expected ESG value given all the previous values equals to the most recent. That is:

\[
E[ESG_i^{(t+1)}|ESG_i^{(1)}, \ldots, ESG_i^{(t)}] = ESG_i^{(t)}
\]

Thus as the expected value of the ESG value of a company during the holding period, we use its most recent rating. In accordance with the argument above, ESG is implemented as a linear function in the optimization problem. Thus, we need a vector holding the ESG-scores of all individual stocks. A detailed explanation of how these values are determined is provided in section 8.1.1. We denote this vector \( \gamma \).

\[
\gamma^T = \left[ E[ESG_1^{(t+1)}], \ldots, E[ESG_n^{(t+1)}] \right] = \left[ ESG_1^{(t)}, \ldots, ESG_n^{(t)} \right]
\]
7.2 The optimization problem

The problem is a multi-objective optimization problem with three objective functions: return, covariance and ESG. This results in the following problem:

\[
\begin{align*}
\text{maximize} & \quad E[R_P] = \mu^\top x \\
\text{minimize} & \quad \text{StDev}(R_P) = x^\top Cx \\
\text{maximize} & \quad E[ESG_P] = \gamma^\top x \\
\text{subject to} & \quad \sum_{i}^{n} x_i = 1 \\
& \quad 0 \leq x_i \leq 1, \forall i \in \{1, \ldots, n\}
\end{align*}
\]

Where the constraints \(0 \leq x_i \leq 1\) follows from us not allowing short sales.

To solve the following multi-objective optimization problem we use the models presented in chapter 6. The approximated nadir objective vector and the ideal objective vector are computed using the lexicographic method described in section 6.3. We want to be able to use the MATLAB functions \texttt{linprog} and \texttt{quadprog}. This means that for iterations where we first find the ideal objective value of a linear function, the second function to be optimized is the other linear function with the first linear functions ideal objective value as a constraint and finally we optimize the quadratic function with the two previous values as constraints. Also note that when we find the ideal objective value of the quadratic function no more optimizations with it as a constraint are needed. The covariance matrix is positive definite, proved in appendix A, i.e. finding the ideal objective value of the covariance function is a strictly convex optimization problem. This means that if an optimal solution exits it is also a unique global minimizer, i.e. if we would have set this as a constraint the same minimizer would have been found.

Both optimizations methods are solved using MATLABs \texttt{quadprog} function, for the weighted sum approach this means solving:

\[
\begin{align*}
\text{minimize} & \quad -\lambda_1 \mu^\top x + \lambda_2 x^\top Cx - \lambda_3 \gamma^\top x \\
\text{subject to} & \quad \sum_{i}^{n} x_i = 1 \\
& \quad 0 \leq x_i \leq 1, \forall i \in \{1, \ldots, n\}
\end{align*}
\]

where \(\lambda_1 + \lambda_2 + \lambda_3 = 1\) and \(0 \leq \lambda_i\). Different nondominated solutions are found by varying the weighting parameters.

The \(\varepsilon\)-constraint method is set up such that we solve:

\[
\begin{align*}
\text{minimize} & \quad x^\top Cx \\
\text{subject to} & \quad \sum_{i}^{n} x_i = 1 \\
& \quad 0 \leq x_i \leq 1, \forall i \in \{1, \ldots, n\} \\
& \quad -\mu^\top x \leq -\varepsilon_1 \\
& \quad -\gamma^\top x \leq -\varepsilon_2
\end{align*}
\]

Different nondominated solutions are found by varying \(\varepsilon_1\) and \(\varepsilon_2\) in the ranges determined by the lexicographic optimizations.
Chapter 8

Results

This chapter consists of two parts. In the first section we will evaluate how the inclusion of ESG criteria affects the portfolio selection process and in the second we evaluate the optimization methods. In section 8.2, both "generated portfolios" and "generated non-dominated solutions" are sometimes used in a similar manner. Recall that a generated portfolio corresponds to a generated nondominated solution. When we are increasing the amount of shares to choose from, we are increasing the computational complexity.

8.1 Portfolio selection

The results presented in this section will all be based on a hypothetical setting. Consider the following situation:

- The portfolio selection takes places at a fixed date
- The composition shall remain unchanged for the whole holding period, whose length is predetermined

With these two characteristics in common, we will then present the optimal portfolios for a number of hypothetical investors. To be able to do this, we need to estimate the expected returns, the covariance matrix and the ESG index. This will be done in section 8.1.1. Then we will present the result of the portfolio selection process in section 8.1.2

8.1.1 Estimation of parameters

The initial step in the portfolio selection process is the estimation of input data. The data needed is the expected return for each stock, the covariances between all securities, and in our case an index for ESG. To estimate the first two, we will utilize the process explained in section 4.3.2. Afterward, we will show how the ESG index is computed.

Computing $\mu$ and $C$

Here we will use the Single-index model. An intermediary result is the alpha and beta values for each security. By using these we can then compute the means $\mu$ and the covariance matrix $C$. As mentioned before the beginning of the holding period is set to a fixed date. We set it to 2014-05-01. The holding period is set to one year.
When estimating alpha’s and beta’s using past returns, there is a trade-off made when choosing which time horizon to use. A too short time horizon gives a unreliable estimate. Old data may unrepresentative of the current state of the company. Berk and DeMarzo (2014) recommends using at least two years of weekly data or five years of monthly data. It should be noted that there is no correct or established way of how to choose data.

To compute the alpha and beta values we used the historical weekly closing prices from the three preceding years, giving 156 data points for each security. The data used was extracted from Yahoo Finance. The closing prices are adjusted, i.e. cleared from dividends and splits. Thus the historical monthly returns were calculated using equation 3.1.1, leaving weekly returns for 155 weeks. As a proxy for the market portfolio we used the index OMXSB, which is a broader index compared to OMXS30 and is constructed to mirror the overall development of the Stockholm Stock Exchange. Now after performing the linear regression described in section 4.3.2 for each security we arrive at the result displayed in column two and three of table 8.1.

<table>
<thead>
<tr>
<th>Security</th>
<th>( \hat{\alpha}_i )</th>
<th>( \hat{\beta}_i )</th>
<th>( \mu_i )</th>
<th>( \hat{\sigma}^2_i )</th>
<th>ESG rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABB</td>
<td>-0.0014</td>
<td>0.8812</td>
<td>0.0947</td>
<td>0.0556</td>
<td>0.7917</td>
</tr>
<tr>
<td>ALFA</td>
<td>0.0005</td>
<td>1.0815</td>
<td>0.1162</td>
<td>0.0702</td>
<td>0.7173</td>
</tr>
<tr>
<td>ASSA B</td>
<td>0.0029</td>
<td>0.9562</td>
<td>0.1027</td>
<td>0.0474</td>
<td>0.6167</td>
</tr>
<tr>
<td>AZN</td>
<td>0.0029</td>
<td>0.3101</td>
<td>0.0333</td>
<td>0.0322</td>
<td>0.7354</td>
</tr>
<tr>
<td>ATCO A</td>
<td>-0.0005</td>
<td>1.1455</td>
<td>0.1231</td>
<td>0.0726</td>
<td>0.7499</td>
</tr>
<tr>
<td>BOL</td>
<td>-0.0028</td>
<td>1.5573</td>
<td>0.1673</td>
<td>0.1273</td>
<td>0.7709</td>
</tr>
<tr>
<td>ELUX B</td>
<td>0.0002</td>
<td>1.1063</td>
<td>0.1189</td>
<td>0.1053</td>
<td>0.7953</td>
</tr>
<tr>
<td>ERIC B</td>
<td>-0.0013</td>
<td>0.6990</td>
<td>0.0751</td>
<td>0.0816</td>
<td>0.8355</td>
</tr>
<tr>
<td>GETI B</td>
<td>0.0003</td>
<td>0.7736</td>
<td>0.0831</td>
<td>0.0673</td>
<td>0.6659</td>
</tr>
<tr>
<td>HM B</td>
<td>0.0002</td>
<td>0.8475</td>
<td>0.0911</td>
<td>0.0474</td>
<td>0.7269</td>
</tr>
<tr>
<td>INVE B</td>
<td>0.0020</td>
<td>1.0018</td>
<td>0.1076</td>
<td>0.0449</td>
<td>0.2914</td>
</tr>
<tr>
<td>LUPE</td>
<td>0.0027</td>
<td>0.8820</td>
<td>0.0948</td>
<td>0.1633</td>
<td>0.5229</td>
</tr>
<tr>
<td>MTG B</td>
<td>-0.0037</td>
<td>1.1489</td>
<td>0.1234</td>
<td>0.1159</td>
<td>0.4925</td>
</tr>
<tr>
<td>NOKI SEK</td>
<td>0.0001</td>
<td>1.1360</td>
<td>0.1221</td>
<td>0.3316</td>
<td>0.6517</td>
</tr>
<tr>
<td>NDA SEK</td>
<td>0.0002</td>
<td>1.2390</td>
<td>0.1331</td>
<td>0.0726</td>
<td>0.7811</td>
</tr>
<tr>
<td>SAND</td>
<td>-0.0032</td>
<td>1.4684</td>
<td>0.1578</td>
<td>0.1116</td>
<td>0.6733</td>
</tr>
<tr>
<td>SCV B</td>
<td>0.0020</td>
<td>0.2453</td>
<td>0.0264</td>
<td>0.1272</td>
<td>0.6591</td>
</tr>
<tr>
<td>SEB A</td>
<td>0.0017</td>
<td>1.2892</td>
<td>0.1385</td>
<td>0.0854</td>
<td>0.8017</td>
</tr>
<tr>
<td>SECU B</td>
<td>0.0002</td>
<td>0.9026</td>
<td>0.0970</td>
<td>0.0661</td>
<td>0.4425</td>
</tr>
<tr>
<td>SKA B</td>
<td>0.0005</td>
<td>1.0186</td>
<td>0.1094</td>
<td>0.0548</td>
<td>0.7428</td>
</tr>
<tr>
<td>SKF B</td>
<td>-0.0018</td>
<td>1.1917</td>
<td>0.1280</td>
<td>0.0780</td>
<td>0.8181</td>
</tr>
<tr>
<td>SCA B</td>
<td>0.0035</td>
<td>0.6877</td>
<td>0.0739</td>
<td>0.0428</td>
<td>0.7472</td>
</tr>
<tr>
<td>SSAB A</td>
<td>-0.0057</td>
<td>1.6722</td>
<td>0.1797</td>
<td>0.1701</td>
<td>0.5479</td>
</tr>
<tr>
<td>SHB A</td>
<td>0.0021</td>
<td>0.9504</td>
<td>0.1021</td>
<td>0.0657</td>
<td>0.5724</td>
</tr>
<tr>
<td>SWED A</td>
<td>0.0019</td>
<td>1.2599</td>
<td>0.1354</td>
<td>0.1061</td>
<td>0.7649</td>
</tr>
<tr>
<td>SWMA</td>
<td>0.0006</td>
<td>0.2641</td>
<td>0.0284</td>
<td>0.0452</td>
<td>0.6173</td>
</tr>
<tr>
<td>TEL2 B</td>
<td>-0.0028</td>
<td>0.5604</td>
<td>0.0602</td>
<td>0.1716</td>
<td>0.3066</td>
</tr>
<tr>
<td>TLSN</td>
<td>-0.0004</td>
<td>0.6946</td>
<td>0.0746</td>
<td>0.0355</td>
<td>0.5085</td>
</tr>
<tr>
<td>VOLV B</td>
<td>-0.0023</td>
<td>1.3818</td>
<td>0.1485</td>
<td>0.0981</td>
<td>0.6975</td>
</tr>
</tbody>
</table>

Table 8.1: Estimated alpha, beta, mean, variance and ESG rating
With these values we can use equation 4.3.11 and 4.3.12. However, since we used monthly data we need to adjust the results by multiplying the expected means by the number of time periods we estimate for (i.e. a factor 52). The same is performed for the covariance, all elements in $C$ are multiplied with a factor 52. The vector of expected returns, $\mu$, is presented in column four in table 8.1 and variances of each security (the diagonal elements of $C$) in column five. The complete covariance matrix has been omitted, due to its size.

**Constructing the ESG-index**

In section 5.4 we examined two ESG indices, with five ratings in total. The data assembled from these indices is provided in appendix C.2. Unfortunately, some data is missing (for further information see section 5.4), which we denote #N/A.

To combine the five scores for each stock we utilized the following approach:

- Divide each individual score with the maximum score of the rating in question (e.g. for the Folksam rating, divide by 7). Thus each score will lie in $[0, 1]$.
- Calculate the mean of the five scores. If data is missing, calculate the mean of the available data.

After doing this, a composite ESG index has been constructed. The result is displayed in the last column of table 8.1.

**8.1.2 Portfolio choice**

Since the starting dates and the holding periods are the same for all investors, the efficient frontier will be the same for all investors. By using the data in table 8.1 with both optimization methods described in section 6, the efficient frontier can be calculated. The result is shown in figure 8.1. The same result is displayed in figure 8.2, now viewed from three different angles. Note once again that this frontier is the same for all investors with the same holding period.
Since the efficient frontier as depicted in 8.1 is a surface in three dimensions, it can be
represented as a number of contour lines. Then each contour line corresponds to a fixed value for some objective, while the investor faces a trade-off in the two other objectives. E.g. if one fixes a value for the ESG-index, the investor faces the traditional choice on a two-dimensional efficient frontier as first introduced by Markowitz (1952). This situation is displayed in figure 8.3a, where the number next to each contour line represents the ESG-value that all portfolios on the same line have in common. Another way to approach the trade-off between the three objectives is to plot the contour lines for different levels on the expected return on the portfolio. This is displayed in figure 8.3b.

![Diagram](image.png)

(a) $E[R_P]$ vs $\text{StDev}(R_P)$

(b) ESG vs $\text{StDev}(R_P)$

Figure 8.3: Efficient frontier as contour plots

To clarify how the inclusion of ESG criteria in the portfolio selection affects the composition of the portfolio (i.e. the vector of portfolio weights, $x$), we will show how $x$ differs for different types of investors. Table 8.2 presents a number of different investors for which we will compute the optimal portfolio. Note that every investor profile is specified in the same way. Each criteria can either be: a percentage, Min/Max or -. With this type of specification, each hypothetical investor will have a unique optimal portfolio. This is fundamental for us to in detail study the effects of including ESG criteria. As shown in figures 8.1 and 8.2, the nondominated set has been determined. Thus we can extract the intervals in which the values of the objective functions for all portfolios lie:

- $E[R_P] \in [0.0472, 0.1797]$
- $\text{StDev}(R_P) \in [0.1183, 0.4124]$
- $\text{ESG} \in [0.5479, 0.8355]$

These intervals should not be confused in such a way that one believes that any portfolio is feasible if the objective values falls within these intervals. For example: it is not possible to construct a portfolio with both $E[R_P] = 0.1797$ and $\text{StDev}(R_P) = 0.1183$.

With a percentage in table 8.2 we refer to a portfolio with some objective value fixed
at the X:th percentile.\footnote{While one may argue that an investor never would state his or her preferences in this way, we do it in this way since it is impossible to \textit{a priori} know which values that are feasible for each criteria} E.g. if an investor has 50% in the Expected Return column:

\[ E[R_P] = 0.0472 + 0.5 \times (0.1797 - 0.0472) = 0.1135 \]

This reduced the portfolio selection to a trade-off on a given contour line, as described above. On this reduced form the investor can either set a fixed percentage, or just choose between minimizing or maximizing. After this choice, a single portfolio remains, and this is the optimal one for the specific investor. In other words: each hypothetical investor has a fixed preference for one of the criteria, then a preference for Min/Max for a second criterion. A preference for the third and final criterion is redundant, since specifying two criteria uniquely determines a portfolio. The criterion that is left to fit the two other is denoted -.

Note that the order is crucial, since if maximizing (or minimizing) after one criterion is fixed is a different problem compared to first maximizing (or minimizing). In table 8.2 all hypothetical investors first sets a fixed percentage for one criterion, then minimizes/-maximizes a second and the third must be left without preference. Each investor profile described in table 8.2 uniquely determines one point on the efficient frontier which corresponds to a unique portfolio vector, \( x \). In table 8.3 these portfolio vectors are presented for each investor profile. The first three rows contains the objective values, then the following rows holds the portfolio weights and the last row contains the number of stocks in each portfolio.

<table>
<thead>
<tr>
<th>#</th>
<th>Description</th>
<th>Expected Return</th>
<th>Volatility</th>
<th>ESG-index</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>High return</td>
<td>80%</td>
<td>Min</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>High ESG</td>
<td>-</td>
<td>Min</td>
<td>75%</td>
</tr>
<tr>
<td>3</td>
<td>Low risk</td>
<td>-</td>
<td>5%</td>
<td>Max</td>
</tr>
<tr>
<td>4</td>
<td>Medium return</td>
<td>40%</td>
<td>Min</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 8.2: Investor profiles
### Table 8.3: Investor specific portfolios

<table>
<thead>
<tr>
<th>Profile</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp Return</td>
<td>0.153</td>
<td>0.055</td>
<td>0.057</td>
<td>0.100</td>
</tr>
<tr>
<td>Volatility</td>
<td>0.284</td>
<td>0.129</td>
<td>0.133</td>
<td>0.186</td>
</tr>
<tr>
<td>ESG</td>
<td>0.714</td>
<td>0.764</td>
<td>0.773</td>
<td>0.775</td>
</tr>
<tr>
<td>ABB</td>
<td></td>
<td>0.073</td>
<td>0.064</td>
<td>0.103</td>
</tr>
<tr>
<td>ALFA</td>
<td></td>
<td></td>
<td>0.039</td>
<td></td>
</tr>
<tr>
<td>ASSA B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AZN B</td>
<td></td>
<td>0.316</td>
<td>0.322</td>
<td>0.080</td>
</tr>
<tr>
<td>ATCO A</td>
<td></td>
<td></td>
<td>0.068</td>
<td></td>
</tr>
<tr>
<td>BOL</td>
<td></td>
<td>0.246</td>
<td></td>
<td>0.042</td>
</tr>
<tr>
<td>ELUX B</td>
<td></td>
<td></td>
<td>0.043</td>
<td></td>
</tr>
<tr>
<td>ERIC B</td>
<td></td>
<td>0.110</td>
<td>0.128</td>
<td>0.061</td>
</tr>
<tr>
<td>GETI B</td>
<td></td>
<td></td>
<td></td>
<td>0.072</td>
</tr>
<tr>
<td>HM B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INVE B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LUPE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTG B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOKI SEK</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NDA SEK</td>
<td></td>
<td>0.171</td>
<td></td>
<td>0.136</td>
</tr>
<tr>
<td>SAND</td>
<td></td>
<td>0.170</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCV B</td>
<td></td>
<td>0.085</td>
<td>0.090</td>
<td>0.021</td>
</tr>
<tr>
<td>SEB A</td>
<td></td>
<td>0.055</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SECU B</td>
<td></td>
<td>0.057</td>
<td>0.054</td>
<td>0.082</td>
</tr>
<tr>
<td>SKA B</td>
<td></td>
<td></td>
<td>0.036</td>
<td>0.056</td>
</tr>
<tr>
<td>SKF B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCA B</td>
<td></td>
<td>0.253</td>
<td>0.284</td>
<td>0.140</td>
</tr>
<tr>
<td>SSAB A</td>
<td></td>
<td>0.108</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHB A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWED A</td>
<td></td>
<td>0.060</td>
<td></td>
<td>0.038</td>
</tr>
<tr>
<td>SWMA</td>
<td></td>
<td></td>
<td>0.106</td>
<td>0.058</td>
</tr>
<tr>
<td>TEL2 B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLSN</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOLV B</td>
<td></td>
<td>0.148</td>
<td></td>
<td>0.011</td>
</tr>
</tbody>
</table>

* # stocks in port. 8 7 7 15

### 8.2 Optimization methods

In this section we will compare the performance of the $\varepsilon$-constraint method and the weighted sum approach by measuring run time and dispersion of the nondominated solutions in several different ways.
8.2.1 Run time

Early exits with the $\varepsilon$-constraint method

We compared the $\varepsilon$-constraint method with and without the early exits discussed in section 6.2 by increasing the amount of generated portfolios and noting run time per portfolio. The results can be seen in figure 8.4.

![Figure 8.4: With or without early exits](image)

Since the difference is noteworthy we will henceforth use this modified implementation of the $\varepsilon$CM method.

Comparing the $\varepsilon$-constraint method and the weighted sum approach

We have measured and compared the run times for the $\varepsilon$-constraint method with the weighted sum approach in three different ways:

- By increasing the included amount of shares for a fixed amount of portfolios and measuring total run time.
- By increasing the included amount of shares and measuring run time per portfolio.
- By increasing the amount of computed portfolios for a fixed amount of securities and measuring run time per portfolio.
In all of the tests shown in figure 8.5 the weighted sum approach is computationally faster than \( \varepsilon \)-constraint method. That the line for the \( \varepsilon \)-constraint method in figure 8.5a is jagged is in part due to it being hard to control the amount of generated portfolios. With the weighted sum approach we control the amount of generated portfolios by choosing the amount of iterations. Since every iteration will yield a nondominated solution increasing the included amount of shares will not affect the amount of generated nondominated solutions. For the \( \varepsilon \)-constraint method this need not be the case since when the included amount of shares increase, the early exits will probably decrease since the range of the frontier increases. This means that for a fixed step size of the \( \varepsilon_i \)'s we will not get the same amount of nondominated solutions. Since it is hard to calibrate the method to yield a given number of solutions, the generated nondominated solutions of the \( \varepsilon \)-constraint method where not always 210, but between 201 and 218.

### 8.2.2 Dispersion

One of the weaknesses of computational time as a measurement is that it does not show if we converge towards the “true”\(^1\) efficient frontier. This means that we might increase the amount of generated points without getting closer to the true efficient frontier. The proposed methods compute the efficient portfolios before involving the investor. Since we want to maximize the available information for the investor one could argue that it would be better to have a greater dispersion, rather than having many nondominated solutions.

---

\(^1\)What true means is a bit abstract, but one could imagine generating points on a line in \( \mathbb{R} \). If the line takes values between \([x, y]\) and we only generate points between \([x, z]\) where \(z < y\). It does not really matter if we increase the amount of generated points, since we are in some sense not getting closer to the “true” line.
within a close proximity. The somewhat arbitrary part is saying what "good dispersion" is.

One way of measuring the dispersion is by increasing the allowed minimum Euclidean distance between nondominated solutions and removing nondominated solutions that are too close. This was done by computing the Euclidean distances between all generated nondominated solutions and then removing nondominated solutions whose distance to another nondominated solution was too small.

**Normalizing the weighted sum approach**

With the approximated nadir objective values and computed ideal objective vector we normalized the objective functions using equation 6.3.1. The effect on run time is negligible (if there is one), since in effect we are only multiplying each objective function with a constant, instead we measured the difference in dispersion with the measurement presented above. In figure 8.6, we show how the amount of nondominated solutions decrease when the minimum allowed Euclidean distance between nondominated solutions increases. One can clearly see that the dispersion (when measured in this way) is greater when we use a normalized weighted sum approach. Coming tests will therefore be done with the normalized version of the weighted sum approach.

**Comparing the ε-constraint method and the weighted sum approach**

We then wanted to compare the performance of the different methods. Figures 8.7 show that the dispersion when measured by increasing the minimum allowed Euclidean distance seems to be better for the weighted sum approach for small distances. For larger distances the difference is barely noticeable. We can also note that in figure 8.7b, the largest distance between two nondominated solutions seems to be about the same for both methods.
The measurement proposed above possesses some weaknesses, which will be discussed in section 9.2.2. We therefore use two additional measurements of dispersion:

- By computing each nondominated solution’s distances to all other nondominated solutions and finding which is the smallest. We do this for every nondominated solution and then take the mean of these distances.

- A weighted mean of all distances

The weighted mean is here defined as $\frac{\sum_{i=1}^{C} d_i}{\sum_{i=1}^{C} i}$. Where $C = m(m-1)/2$, is the amount of pairwise distance comparisons if we generate $m$ portfolios, i.e. we only include the distance between two nondominated solutions once. $d_i$ is the $i$:th smallest pairwise comparison. This measurement is expected to decrease when we increase the amount of generated nondominated solutions, and will decrease faster if there is a lot of clustering. We computed these measurements by increasing the amount of generated nondominated solutions. The results can be seen in figure 8.8. Two things can be noted. In figure 8.8a we see that the mean minimum distance for portfolios generated by the $\varepsilon$-constraint method is larger than for portfolios generated by the weighted sum approach. The same holds in 8.8b, the difference is that the difference between the two methods seems to be increasing.

Finally we plot the nondominated set generated by each method in figures 8.9 These results will be discussed in section 9.2.2.
(a) Generated by the $\varepsilon$-constraint method  
(b) Generated by the weighted sum approach  
(c) The two methods together

Figure 8.9: Nondominated solutions generated by the different methods
Chapter 9

Discussion

9.1 ESG

There are several possible ways in which an investor could include ESG concerns as a factor in the investment process. In this thesis we chose to do this by assigning each company a quantitative score and then using multi-objective optimization methods to generate an efficient frontier. Data used to develop scores where gathered from two ESG indices and combined as presented in 8.1.1. During the period we worked with this thesis we discussed and thought about several ways one could combine the two indices to get a unified score. One proposition was to convert the indices to rankings and thereafter combining these. We chose not to do this since we think that cardinal measures for ESG are more intuitive for the investor when used in a model where a majority of efficient portfolios contain more than one stock. To solve the problem of converting two different measurement scales (Folksam’s 0-7 scale and STOXX’s 0-100 scale) we pondered which combination method that would be preferred. The discussed methods were:

1. Dividing each score by the maximum possible score and then taking the average of these scores
2. Dividing each score by the maximum score obtained by the included companies and taking the average of these scores
3. Dividing the results in 2. by the maximum score obtained in 2.
4. Normalizing the rows by subtracting the minimum obtained score and dividing by the maximum obtained score minus the minimum obtained score

We have not found any literature in which the combinations of ratings is discussed. Instead we chose the first method since it is the only one taking into account that there might be stocks not included that have a higher or lower ranking, i.e. the scores of the included companies are compared with the maximum and minimum scores. This also means that if an index increased its included stocks we would not need to recompute that specific index. We chose this model before looking at the data since we wanted to minimize the bias that could occur otherwise we could have made a selection based on our data set. When in fact we want to present a framework investors could use for other indices (and at other times for the same indices as well).
Since choosing how to combine the two indices is not an exact science, a motivation for why we chose to use two indices in the first place is needed. As was discussed in section 5.3 there does not exist a standard for quantifying ESG. The ranks of each share can be seen in appendix C.3. Using these we found that:

- There was no overlap between the five best ranked companies in Folksam’s index and STOXX’s index
- Of the five best ranked companies in Folksam’s social rating only one (H&M) could be found in the five best ranked in STOXX’s social rating and one (SEB) of the five best ranked in the corporate governance rating.
- Of the five worst ranked companies in the environmental ratings four were the same. When comparing Folksam’s social with STOXX’s social three were the same and when comparing with STOXX’s corporate governance two were.
- We compared the ranking in Folksam’s social rating with STOXX’s social rating, Folksam’s social rating with STOXX’s corporate governance rating and Folksam’s environmental rating with STOXX’s environmental rating. For seven companies the difference in rank was over fifteen places in at least one of the three ratings.

Our conclusion is that there seems to exist a consensus for the lower rankings but not for the higher ones. For an investor considering ESG as a factor in the portfolio selection process this might imply that negative screening is a possible and maybe a better option. Negative screening means that the investor in advance eliminates companies below some threshold and then uses a preferred portfolio selection model. This elimination could be done using indices or a qualitative analysis of each company. The drawback with this approach is that the trade-offs displayed in figure 8.3 are not taken into account. Using negative screening might therefore be an alternative for investors wanting two eliminate the worst ESG performing companies, whilst staying indifferent for companies who perform over a predetermined threshold. The multi-criteria optimization approach taken in this thesis, is according to us still a more feasible approach for investors interested in the made trade-offs. A third approach would be to first use negative screening after which our proposed model is used. This would allow the investor the possibility of eliminating the worst performing companies and being able to monitor the trade-offs between the objectives. We chose not to use this last approach since we limited the number of available shares to OMXS30, if we were to eliminate, say, the five lowest rated companies we would have removed about 17 percent of all shares. This approach is according to us more feasible when the number of available shares is a lot larger.

This leads to another critical subject, what happens when the number of available shares is increased? As is often the case for any financial reporting, larger companies report more and are more thoroughly examined. The STOXX ESG Index contains 40 Swedish companies and Folksam’s Index for Responsible Business contains 250 Swedish companies. Both indices combine data reported by the companies with data from other sources. One can question how reliable data is from smaller companies, not because they have any malice, but because they do not have the time and capital or face the same stakeholder pressure larger companies do. It is also reasonable to assume that external reporting and auditing is lacking in quality. For an investor relying on public (and in this

\footnote{Recall that this was a combination of Social and Corporate Governance ratings used in STOXX’s index}
case free) data this can turn out to be a major problem. Models are only as good as their inputs allow them to be and more insecurity in the inputs might make the model a lot less useful.

Another question that arises is if ESG should be treated as one or several different criteria. To a large extent this might be driven by the portfolio model, quality of data and the investor. When using the type of models discussed in this thesis, increasing the amount of criteria to a total of four or more, will make it considerably harder to choose a portfolio from the frontier. This because, it is no longer possible to illustrate the complete efficient frontier at once. Still, it is possible to argue that the three components of ESG differ and companies might excel in one criteria whilst performing badly in another. It is then up for discussion if the loss of visualization is worth the increased amount of information. Unfortunately this question lacks a general answer and depends on the portfolio model, quality of data and the investor. Since neither Folksam nor STOXX provide the raw data used to compute the different ratings, we and other investors using them, have a hard time evaluating if the ESG rating of one company is "valid" or based on bad data and speculation. It is worth pointing out that for our data, all correlation coefficients were positive. One might therefore use the analogy of the law of large numbers. Even though five data points might (and probably do) not satisfy this condition in a strict mathematical sense, it might be possible to use it as an analogy. The law of large numbers basically says that if we perform the same experiment a lot of times, the mean of the measured values will converge towards the true value. The argument then goes as follows: the compounding of different and several values of E, S and G might "converge" towards a value that illustrates how well a company handles non-financial issues. For an investor willing to pay, compounding a few indices containing several ratings for each company might yield a better representation of E, S and G than one would get if they were looked at separately.

Investors interested in ESG can have a positive outlook and as their numbers increase so will the accessible data. There are already a lot of third-party agencies offering raw data as well as standardized indicators for investors willing to pay.

9.2 Optimization methods

The two main methods, the weighted sum approach and the $\varepsilon$-constraint method, were evaluated using two measurements: time and dispersion. These methods are here used to generate the efficient frontier, after which the investor chooses a portfolio. In contrast, there are several preference-based methods, e.g. methods taking into consideration the preferences of the investor before (or during) the optimization. For preference-based methods where the investor states his or her preferences before the optimization is done, we wonder how preferences rightfully can be expressed before the optimization is done. What return should an investor aim for if he does not know the range of the possible returns? This is also complicated by the other criteria, what is an acceptable standard deviation and how much of it is the investor willing to sacrifice for a higher return? In this thesis the ESG criterion was created from indices which had aggregated several measurements of ESG, again the problem is saying what an acceptable ESG factor is before the optimization and stating acceptable trade-offs. A preference-based model where the investor is involved during the process. E.g. by defining the worst acceptable solution, generating that part of the efficient frontier and then again stating the worst acceptable solution from
this part. This approach solves some of the aforementioned problems but some remain. When generating parts of the efficient frontier information that might benefit the investor is discarded.

9.2.1 Computation time

The rationale for choosing a preference-based method is sometimes that it lowers the needed computational time. This was one reason for measuring the different methods run time. Even though we do not say which portfolio on the frontier is the best, someone (the investor) eventually has to. Since two portfolios close to each other on the efficient frontier might have completely different compositions, it is of interest to see how run time is affected when the amount of generated portfolios increase.

It is also important to note that we limited the number of available shares to OMXS30. Markowitz’s initial model implicitly assumes that the investor can choose from a lot more shares than 29. In section 9.1 we mentioned why we limited the number of available shares, but for investors willing to pay for data, the number of available shares will probably contain a lot more shares. This makes it interesting to note how run-time increases when we increase the number of available shares.

All our tests showed that the weighted sum approach had shorter run time than the \( \varepsilon \)-constraint method. This might not come as a surprise since the \( \varepsilon \)-constraint method amounts to solving a quadratic programming problem with equality and inequality constraints. Whereas the weighted sum approach amounts to solving a quadratic programming problem that only has equality constraints\(^1\).

One should be wary towards run time as a vital measurement for how good these types of methods are. Markowitz’s model and ours as well are usually used for longer holding periods. In this thesis our assumed holding period was one year. Relative to that a few minutes or even hours is not critical. This strengthens the argument for not using preference-based models. Another weakness with run time measured in these ways, is that it does not tell you with what rate the quality of gathered information increases. If we increase the amount of generated nondominated solutions we would preferably want these to fill the “blank” spaces on the nondominated set and in an ideal world the most unexplored parts of the nondominated set. Instead we might see that increasing the amount of nondominated solutions only leads to clustering on certain parts of the nondominated set. This is the main reason for introducing dispersion as a crucial measurement of how good the model works.

9.2.2 Dispersion

If generating the “whole” nondominated set is a rationale for not choosing a preference-based method, dispersion should be important. In this thesis we first chose to measure this by decreasing the minimum allowed distance between nondominated solutions. The problem with this approach can be seen in figure 9.1. If we begin with comparing nondominated solutions A, C, D, E and F with B, only A and D will be removed. If we instead begin with comparing nondominated solutions B, C, D, E and F with A, all nondominated solutions except A and F will be removed. The problem will occur during the second iteration when if E is already removed, F will not be, even though these two are the nondominated solutions closest to each other.

\(^1\)Disregarding the constraint \( x_i \geq 0 \) which MATLAB treats in a different way
We tried to suppress this effect by scrambling the order in which the comparison was done ten times for each distance. This resulted in some damping but as can be seen in figures 8.7a and 8.7b there is still some oscillation.

Figure 9.1: How order affects elimination

In figure 8.7a we can see that when the allowed minimum allowed distance is relatively small the weighted sum approach is better, whereas for larger distances shown in figure 8.7b the difference seems negligible. The latter result is important since it shows that the ranges for the different methods seem to be about the same. One of the weaknesses of the $\varepsilon$-constraint method is that we have to choose ranges for the $\varepsilon_i$'s. This involves approximating the nadir-values, figure 8.7b therefore shows that this approximation seems to work rather well for our problem.

Due to the weakness mentioned above we used two more measurements: (1) A weighted mean of all distances, and (2) the mean of the smallest distance each nondominated solution has to all other nondominated solutions. The results of the proposed measurements might at first be a bit confusing but become more clear if we look at the nondominated solutions generated by each method. These were shown in figures 8.9. In figure 6.1 we showed how one can interpret the weighted sum approach in the two-objective case. With three objectives the only difference is that instead of a line moving outward from the origin, we now have a plane. If at least one objective function is rapidly changing, a small change of the weighting parameters will yield a big jump on the nondominated set. If this is not the case, a little change of the weighting parameters will yield a smaller jump. In figure 8.9b this can be seen by noting the clustering on the "upper" parts of the nondominated set and the hole in the middle. By this logic, decreasing the step size of the weighting parameters will generate more points on the flatter parts than it will on the steeper parts of the nondominated set. Decreasing the step size of the $\varepsilon_i$'s will not lead to this type of clustering. This explains the weighted mean illustrated in figure 8.8b. As the amount of generated portfolios increase, smaller distances will have a larger share of the weighted mean for the weighted sum approach than for the $\varepsilon$-constraint method.

Figures 8.9 also shows where the $\varepsilon$-constraint method performs badly. When solving the problem at hand we wanted to be able to use MATLAB’s quadprog function, this meant that we optimized the covariance function and placed ESG and expected return as $\varepsilon$-constraints. In figure 8.9c we see that in upper parts of the nondominated set, the
solutions generated by the $\varepsilon$-constraint method is rather sparse when compared to the weighted sum approach. This happens because when one $\varepsilon_i$ (or two) are close to its maximum value\textsuperscript{1} on the nondominated set, a small increase in $\varepsilon_i$ might mean a big trade-off in standard deviation. Both the ESG and expected return criteria are linear functions, as we increase one of them diversification will decrease, in general this also means a higher standard deviation. This also explains why there seems to be some clustering as we move ”higher” up on the nondominated set in figure 8.9a. This can partly explain the rather big decrease seen in figure 8.7a.

Measurements evaluating the dispersion are hard to find in previous research, this is probably due to it being a rather abstract measurement. To evaluate dispersion one often needs to inspect the nondominated solutions generated by one method with another method. For larger models this might not be an option, one is then faced with a complex issue. As we see it, possible improvements that can be made for both methods are adaptive step lengths. By in some way measuring the steepness of the nondominated set, adaptive changes of the weighting parameters could be made. For the $\varepsilon$-constraint method it might be of interest to use adaptive changes for $\varepsilon_i$'s as we reach the boundaries of the nondominated set, and especially where at least one of the linear functions have large values.

From the measurements we proposed and computed, it is hard to say which method one should prefer for our type of problem. We only tested the models dispersion on the data set we had, and the results can therefore not be assumed to hold for other data sets. It seems as if the two methods are better at different parts of the nondominated set, to say that one part is more important than the other is hard and we will not try to answer this question.

\section{Portfolio selection}

According to us, the fundamental concept of portfolio theory is to provide an investor with a decision basis for the decision that the portfolio selection constitutes. Markowitz’s mean-variance framework gives the investor a trade-off between expected return and volatility at a two-dimensional efficient frontier corresponding to efficient portfolios. The main result is that an investor should choose a portfolio on the efficient frontier since all other portfolios are worse in some aspect. Furthermore, the mean-variance framework emphasizes the benefits of a diversified portfolio, which in some sense is one of the few ”free lunches” in the financial markets.

Following the recent years increased interest in ESG concerns it seems suitable to include these concerns in the portfolio selection process. Our approach\textsuperscript{2} is to include these as a third criterion, in addition to expected return and volatility. By doing this, the efficient frontier is extended from a line in the two-dimensional plane to a surface in the three-dimensional space. A comparison between figure 4.2 and 8.1 makes this perfectly clear.

In the same way as an investor using the traditional Markowitz framework should choose a portfolio on the red line in figure 4.2, an investor which includes ESG concerns in

\textsuperscript{1}Since both ESG and expected return are to be maximized
\textsuperscript{2}As described thoroughly in the preceding sections of this thesis, the mathematics of this approach follows the work of other, e.g. Steuer, Qi, et al. (2008). The way our work differs is in the focus on ESG concerns, where others seem to be interested in additional criteria in general.
the portfolio choice should choose one of the portfolios represented as red circles in figure 8.1. However, it is quite complicated to in a comprehensive way evaluate to which extent this affects the individual investor. That is, there is no obvious way of comparing a two-dimensional line with a three-dimensional surface. Moreover, it was not the purpose of our work to compare the traditional Markowitz framework with the extended framework.

9.3.1 Level curves and the ESG-to-variability ratio

One way of relating the extended to the traditional framework is to consider the three-dimensional front as a set of level curves. This is shown in figure 8.3.

Panel (a) of 8.3 shows the traditional trade-off (i.e. between expected return and volatility) for fixed values of the ESG-index. An advantage of this representation is that it transforms the nondominated set to efficient frontier recognized by investors. One can see that the level curves lie close together. This implies that increased ESG is cheap, in terms of decreased $E[R_P]$ and increased $StDev(R_P)$. The conclusion is that an investor should (or would) choose a portfolio with relatively high ESG-rating. This is also implied by table 8.3, in which we see that portfolios with quite different characteristics have very similar ESG-values.

Panel (b) of 8.3 shows the trade-off between ESG-rating and volatility (for fixed $E[R_P]$). Striking is that each level curve has a sharp bend, reminiscent of a knee. The reason for this is evident if one studies the number of stocks included in the portfolios that lies along a particular contour line, as shown in figure B.1. At first, the decreasing number of stocks has a large impact on ESG but only a small impact on volatility. However, as one approaches the “knee” the number of stocks in the portfolio is rapidly decreasing. This is because only a few stocks posses high ESG-ratings. That is, to achieve a higher ESG-rating in the portfolio, a decreasing number of stocks can be included in it. As the number of stocks in the portfolio decreases, so does the benefit of diversification, i.e. the volatility increases rapidly. Thus after the knee, only a small increase in ESG costs relatively much in increased volatility of the portfolios. For the traditional Markowitz frontier, a common way to determine a single portfolio is to use (and maximize) the Sharpe ratio\(^1\). Sharpe (1966) defines it as:

$$S = \frac{E[R_P - R_f]}{\sqrt{Var(R_P)}}$$

(9.3.1)

where $R_f$ denotes the risk-free interest rate. Thus by maximizing $S$, one maximizes the expected excess return per unit of risk. In analogy with this, we propose the following: assume instead that an investor is interested in the trade-off between ESG and risk. The investor should then (as with the Sharpe ratio) try to maximize the excess ESG per unit of risk. Let $\overline{ESG}$ denote some benchmark value for ESG. Then we define the “ESG-to-variability ratio” as:

$$E = \frac{ESG_p - \overline{ESG}}{\sqrt{Var(R_P)}}$$

(9.3.2)

which the investor should maximize.

As benchmark, it seems suitable to pick the mean value of all\(^2\) listed stocks at the

---

\(^1\)Sharpe called it the “reward-to-variability” ratio in his original work, but with time the name has become the Sharpe ratio.

\(^2\)Stocks listed on Small, Medium and Large cap. In total 245 stocks.
Using the Folksam index described above gives $ESG = 0.3014$. Now by maximizing $E$ for each contour line in figure 8.3b gives a “optimal”\(^1\) portfolio for each fixed value of $E[R_P]$. This is shown in figure B.2. We see that these portfolios are approximately where the curvature is sharpest, i.e. where the marginal utility of taking on additional risk decreases sharply. The main benefit of doing this additional optimization is that it reduces the number of portfolios the investor should consider, thus simplifying the portfolio selection further.

One can of course compute the Sharpe ratio-optimal portfolios for the level curves in 8.3a, using a risk-free interest rate of 2%\(^3\) and equation 9.3.1 gives the result displayed in B.3. The argument to why this is beneficial is the same as before.

### 9.3.2 Some remarks on the investor specific portfolios

One of our main results is the efficient frontier as presented in 8.1. With the corresponding portfolios (i.e. efficient solutions) this represents the decision basis for the investor. One may argue that this finishes our work, since we make no attempt to guess the preferences of individual investors. However, since we are interested of how the inclusion of ESG concerns affects the portfolio choice we need to extract specific portfolios from the set of efficient solutions. Our way of doing this is outlined in section 8.1.2 that resulted in some hypothetical investors presented in table 8.2. Computing the optimal portfolios for these investors gave the portfolio weights displayed in figure 8.3.

Since we now have specific portfolios, some analysis may be performed. We see that independent of the preferences of the investors, some stocks are excluded from the portfolios. For example, INVE-B, LUPE and MTG-B are not part of any of the portfolios presented in table 8.3. These stocks are also associated with a low ESG ranking (with rankings 29, 24 and 26, respectively as presented in table C.4). We can expect that their absence from the portfolios in part be due to their low ESG-ratings. We need however to nuance the picture slightly. If a stock is to be included in a portfolio is also dependent on its expected return, its variance and its covariance with other stocks. If we limit ourselves to just look at the former two, we see that the mentioned stocks ranks in the middle for expected return, with ranks 15, 19 and 9. Furthermore, the ranks for variance are 4, 26 and 23. This leads us to the conclusion that the exclusion of INVE-B is mostly part to the bad ESG-rating (since medium exp. return and good variance). For LUPE and MTG-B both ESG and variance are bad, and expected return is in the middle ground. Thus it is quite clear that they should not be included in a portfolio, however, if it is the bad ESG-rating or the high (bad) variance that causes the exclusion may depend on investor specific preferences. That is, a high ESG-seeking investor may accept a high variance and a high return-seeking investor may accept a low ESG-rating. Even if the specific mechanisms are invisible\(^4\) to us, the main conclusion is that the inclusion of ESG concerns as a third objective in the portfolio optimization affects the portfolios of investors.

---

\(^1\)An argument for this could be that holding all stocks represents a very-well diversified portfolio, thus only subject to systematic risk, and as “risk-free” as a stock portfolio can get.

\(^2\)Optimal according to our alternation of the Sharpe ratio

\(^3\)Ten-year Swedish government bond, for April 2014

\(^4\)We are unable to in detail study which criterion that causes the inclusion or exclusion of a stock in a portfolio, since it is the aggregate values of the objective functions that matter
Chapter 10

Conclusion

In recent years investors in the financial markets around the globe have begun to focus on non-financial factors in their portfolio selection processes (UNPRI, 2014a). ESG is a catch-all term for these additional factors. In this thesis we conclude that it seems reasonable for investors to include ESG concerns in a mathematical framework to enable portfolio optimization. This stems from two reasons: investors have ESG concerns and companies working with ESG issues perform better than their counterparts. Having a mathematical framework as an aid in the portfolio selection process will hopefully lead to better decisions being made by decision makers.

However, extending the Markowitz model to a three-dimensional trade-off increases the complexity faced by investors. To alleviate this drawback, we propose to represent the efficient frontier as a set of level curves and thus relate the extended model to the original one. As an extension of this, we propose a variant of the Sharpe ratio, namely the "ESG-to-variability ratio". This reduces the number of portfolios the investor should consider, and thus simplifying the process of selecting a portfolio.

When studying the effects of the inclusion of ESG as a third objective function we saw that some stocks were in fact penalized and thus excluded from the (hypothetical) investor specific portfolios determined.

A model is only as good as the data used in it. In this thesis we use the Single-index model to estimate expected returns and covariances. In a real-world setting a more sophisticated model would be appropriate. One should also keep in mind the obvious difficulties encountered when trying to predict and model the future.

Two multi-criteria optimization methods were evaluated, the ε-constraint method and the weighted sum approach. Both of these methods generate a nondominated set which is presented to the investor. The original methods are fairly primitive and two improvements were implemented. We avoided unnecessary iterations using early exits with the ε-constraint method and used research originally designed to compute the boundaries of the ε_i’s to normalize the weighted sum approach.

Run time was evaluated to see how the methods would fair when input data and generated nondominated solutions were increased. These tests showed that the weighted sum approach was the computationally faster of the two. The main weakness with these measurements are that they do not take into account if and how fast the methods converge towards a true nondominated set. We therefore proposed three measurements of dispersion. The results showed that each method has its own strengths and weaknesses.

We suggest that future research should aim at developing adaptive step lengths for
both methods. The main conclusion is that no method is preferable over the other, for small data sets we recommend that the methods are used together.

We would like to end this thesis by stressing that financial mathematical frameworks are always simplifications of a complex reality. One should always be aware of the simplifications and assumptions behind the model. Integrating the ESG criterion does not necessarily mean that we are any better off in trying to describe reality, but it gives investors the possibility of evaluating measurements not usually included in the portfolio selection process.
Bibliography


Mavrovats, George (2009). “Effective implementation of the i ε j i ε i ε i i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i ε i ε i ε i i


Appendix A

C can be assumed to be positive definite

A matrix $H$ is positive semidefinite if $x^\top H x \geq 0$, for all $x \in \mathbb{R}^n$. A matrix $H$ is positive definite if $x^\top H x > 0$, for all nonzero $x \in \mathbb{R}^n$.

We have: $C = \beta \beta^\top \sigma_m^2 + D(\sigma_e^2)$ (The hats used to denote estimates in 4.3.12 are excluded here). We see that (and using that $\sigma_m^2$ is a scalar):

$$x^\top C x = x^\top (\beta \beta^\top \sigma_m^2 + D(\sigma_e^2)) x = x^\top \beta \beta^\top x \sigma_m^2 + x^\top D(\sigma_e^2) x$$

Beginning with the first term:

$$x^\top \beta \beta^\top x = (x^\top \beta)(\beta^\top x) = (x^\top \beta)(x^\top \beta)^\top = \|x^\top \beta\|^2 \geq 0$$

for all $x \in \mathbb{R}^n, \beta \in \mathbb{R}^n$. By the definition of variance, $\sigma_m^2 \geq 0$. Therefore:

$$x^\top \beta \beta^\top x \sigma_m^2 \geq 0, \text{ for all } x \in \mathbb{R}^n \text{ and } \beta \in \mathbb{R}^n$$

Now recall the second term: $x^\top D(\sigma_e^2) x$. Note that:

$$D(\sigma_e^2) = \begin{bmatrix} \sigma_{1e}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{2e}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{ne}^2 \end{bmatrix}$$

Thus:

$$x^\top D(\sigma_e^2) x = x_1^2 \sigma_{1e}^2 + x_2^2 \sigma_{2e}^2 + \cdots + x_n^2 \sigma_{ne}^2$$

By definition, $\sigma_{ie}^2 \geq 0$. Consequently we have:

$$x^\top D(\sigma_e^2) x \geq 0$$

Now we can conclude that:

$$x^\top C x = x^\top \beta \beta^\top x \sigma_m^2 + x^\top D(\sigma_e^2) x \geq 0$$

which holds for all $x \in \mathbb{R}^n$ and $\beta \in \mathbb{R}^n$. Hence $C$ is a positive semidefinite matrix.

However, in practice $\sigma_{ie}^2 > 0$, since the residual terms from the regressions performed in the single-index model have a variance (otherwise the returns of an individual stock must be perfectly correlated with the market, i.e. the linear regression gives no residual terms). Thus the second term in the expression above is strictly greater than zero (for nonzero $x$). Thus, we generally believe that the covariance matrix is $C$ positive definite.

To summarize: $C$ is always positive semidefinite and can in this thesis also be assumed to be positive definite.
Appendix B

Figures

Figure B.1: Number of stocks in portfolio along a contour curve

Figure B.2: Contour lines with ESG-to-variability ratio maximizing portfolios
Figure B.3: Contour lines with Sharpe ratio maximizing portfolios
Appendix C

Tables

UN Global Impact

<table>
<thead>
<tr>
<th>Principle</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Businesses should support and respect the protection of internationally proclaimed human rights; and make sure that they are not complicit in human rights abuses.</td>
</tr>
<tr>
<td>2</td>
<td>Businesses should uphold the freedom of association and the effective recognition of the right to collective bargaining; the elimination of all forms of forced and compulsory labour; the effective abolition of child labour; and the elimination of discrimination in respect of employment and occupation.</td>
</tr>
<tr>
<td>3</td>
<td>Businesses should support a precautionary approach to environmental challenges; undertake initiatives to promote greater environmental responsibility; and encourage the development and diffusion of environmentally friendly technologies.</td>
</tr>
<tr>
<td>4</td>
<td>Businesses should work against corruption in all its forms, including extortion and bribery.</td>
</tr>
</tbody>
</table>

Table C.1: Ten principles in the areas of human rights, labour, the environment and anti-corruptions
<table>
<thead>
<tr>
<th>Name</th>
<th>Symbol</th>
<th>Folksam</th>
<th>Miljö</th>
<th>MR</th>
<th>E</th>
<th>S</th>
<th>G</th>
<th>STOXX</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABB Ltd</td>
<td>ABB</td>
<td>4.40</td>
<td>4.55</td>
<td>90.50</td>
<td>96.90</td>
<td>80.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alfa Laval</td>
<td>ALFA</td>
<td>4.31</td>
<td>3.64</td>
<td>91.30</td>
<td>63.30</td>
<td>90.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASSA ABLOY B</td>
<td>ASSA B</td>
<td>4.08</td>
<td>3.68</td>
<td>77.00</td>
<td>80.60</td>
<td>39.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AstraZeneca</td>
<td>AZN</td>
<td>4.85</td>
<td>4.62</td>
<td>81.70</td>
<td>68.10</td>
<td>82.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atlas Copco A</td>
<td>ATCO A</td>
<td>4.19</td>
<td>4.03</td>
<td>75.60</td>
<td>99.30</td>
<td>82.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boliden</td>
<td>BOL</td>
<td>5.43</td>
<td>4.20</td>
<td>90.10</td>
<td>93.80</td>
<td>64.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrolux B</td>
<td>ELUX B</td>
<td>4.89</td>
<td>4.45</td>
<td>96.30</td>
<td>95.20</td>
<td>72.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ericsson B</td>
<td>ERIC B</td>
<td>5.92</td>
<td>4.90</td>
<td>91.90</td>
<td>90.10</td>
<td>81.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Getinge B</td>
<td>GETI B</td>
<td>3.59</td>
<td>2.77</td>
<td>98.90</td>
<td>84.90</td>
<td>58.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hennes &amp; Mauritz B</td>
<td>HM B</td>
<td>4.03</td>
<td>4.87</td>
<td>95.80</td>
<td>95.40</td>
<td>45.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investor B</td>
<td>INVE B</td>
<td>1.47</td>
<td>2.31</td>
<td>8.90</td>
<td>63.00</td>
<td>19.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lundin Petroleum</td>
<td>LUPE</td>
<td>1.51</td>
<td>2.35</td>
<td>57.60</td>
<td>61.40</td>
<td>87.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modern Times Group B</td>
<td>MTG B</td>
<td>2.52</td>
<td>2.84</td>
<td>28.70</td>
<td>53.10</td>
<td>87.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nokia Oyj</td>
<td>NOKI SEK</td>
<td>#N/A #N/A</td>
<td>98.70</td>
<td>47.20</td>
<td>49.60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nordea Bank</td>
<td>NDA SEK</td>
<td>4.12</td>
<td>4.48</td>
<td>96.50</td>
<td>89.20</td>
<td>82.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sandvik</td>
<td>SAND</td>
<td>4.31</td>
<td>4.03</td>
<td>47.50</td>
<td>98.90</td>
<td>71.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SKF B</td>
<td>SKF B</td>
<td>5.52</td>
<td>4.66</td>
<td>95.80</td>
<td>87.00</td>
<td>96.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swedbank A</td>
<td>SWED A</td>
<td>4.80</td>
<td>4.59</td>
<td>92.40</td>
<td>65.70</td>
<td>90.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swedish Match</td>
<td>SWMA</td>
<td>4.03</td>
<td>3.22</td>
<td>78.70</td>
<td>72.80</td>
<td>53.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tele2 B</td>
<td>TEL2 B</td>
<td>2.52</td>
<td>2.03</td>
<td>19.90</td>
<td>12.80</td>
<td>55.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TeliaSonera</td>
<td>TLSN</td>
<td>4.26</td>
<td>4.03</td>
<td>77.70</td>
<td>41.30</td>
<td>16.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Volvo B</td>
<td>VOLV B</td>
<td>6.03</td>
<td>4.62</td>
<td>58.50</td>
<td>83.90</td>
<td>54.20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table C.2: ESG data used
<table>
<thead>
<tr>
<th>Stock</th>
<th>Miljö</th>
<th>MR</th>
<th>E</th>
<th>S</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABB</td>
<td>11</td>
<td>9</td>
<td>10</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>ALFA</td>
<td>12</td>
<td>19</td>
<td>9</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>ASSA B</td>
<td>18</td>
<td>18</td>
<td>17</td>
<td>14</td>
<td>26</td>
</tr>
<tr>
<td>AZN</td>
<td>8</td>
<td>5</td>
<td>13</td>
<td>18</td>
<td>10</td>
</tr>
<tr>
<td>ATCO A</td>
<td>15</td>
<td>15</td>
<td>18</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>BOL</td>
<td>4</td>
<td>12</td>
<td>11</td>
<td>6</td>
<td>17</td>
</tr>
<tr>
<td>ELUX B</td>
<td>7</td>
<td>11</td>
<td>4</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>ERIC B</td>
<td>2</td>
<td>1</td>
<td>8</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>GETI B</td>
<td>23</td>
<td>25</td>
<td>1</td>
<td>11</td>
<td>18</td>
</tr>
<tr>
<td>HM B</td>
<td>20</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>24</td>
</tr>
<tr>
<td>INVE B</td>
<td>28</td>
<td>27</td>
<td>28</td>
<td>21</td>
<td>27</td>
</tr>
<tr>
<td>LUPE</td>
<td>27</td>
<td>26</td>
<td>22</td>
<td>22</td>
<td>8</td>
</tr>
<tr>
<td>MTG B</td>
<td>25</td>
<td>24</td>
<td>25</td>
<td>25</td>
<td>7</td>
</tr>
<tr>
<td>NOKI SEK</td>
<td>#N/A</td>
<td>#N/A</td>
<td>2</td>
<td>26</td>
<td>23</td>
</tr>
<tr>
<td>NDA SEK</td>
<td>17</td>
<td>10</td>
<td>3</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>SAND</td>
<td>12</td>
<td>15</td>
<td>24</td>
<td>2</td>
<td>16</td>
</tr>
<tr>
<td>SCA B</td>
<td>6</td>
<td>12</td>
<td>15</td>
<td>17</td>
<td>2</td>
</tr>
<tr>
<td>SCV B</td>
<td>10</td>
<td>23</td>
<td>20</td>
<td>23</td>
<td>3</td>
</tr>
<tr>
<td>SEB A</td>
<td>22</td>
<td>3</td>
<td>5</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>SECU B</td>
<td>24</td>
<td>22</td>
<td>26</td>
<td>16</td>
<td>25</td>
</tr>
<tr>
<td>SKA B</td>
<td>5</td>
<td>12</td>
<td>19</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td>SKF B</td>
<td>3</td>
<td>3</td>
<td>12</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>SSAB A</td>
<td>16</td>
<td>20</td>
<td>#N/A</td>
<td>#N/A</td>
<td>#N/A</td>
</tr>
<tr>
<td>SHB A</td>
<td>18</td>
<td>5</td>
<td>23</td>
<td>24</td>
<td>22</td>
</tr>
<tr>
<td>SWED A</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>19</td>
<td>5</td>
</tr>
<tr>
<td>SWMA</td>
<td>20</td>
<td>21</td>
<td>14</td>
<td>15</td>
<td>21</td>
</tr>
<tr>
<td>TEL2 B</td>
<td>25</td>
<td>28</td>
<td>27</td>
<td>28</td>
<td>19</td>
</tr>
<tr>
<td>TLSN</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>27</td>
<td>28</td>
</tr>
<tr>
<td>VOLV B</td>
<td>1</td>
<td>5</td>
<td>21</td>
<td>12</td>
<td>20</td>
</tr>
</tbody>
</table>

Table C.3: ESG data as ranks
<table>
<thead>
<tr>
<th>Stock</th>
<th>ESG</th>
<th>Rank</th>
<th>Exp. Return</th>
<th>Rank</th>
<th>Variance</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABB.ST</td>
<td>0.79</td>
<td>5</td>
<td>0.0947</td>
<td>20</td>
<td>0.0556</td>
<td>9</td>
</tr>
<tr>
<td>ALFA.ST</td>
<td>0.72</td>
<td>14</td>
<td>0.1162</td>
<td>13</td>
<td>0.0702</td>
<td>13</td>
</tr>
<tr>
<td>ASSA-B.ST</td>
<td>0.62</td>
<td>21</td>
<td>0.1027</td>
<td>16</td>
<td>0.0474</td>
<td>6.5</td>
</tr>
<tr>
<td>AZN.ST</td>
<td>0.74</td>
<td>12</td>
<td>0.0333</td>
<td>27</td>
<td>0.0322</td>
<td>1</td>
</tr>
<tr>
<td>ATCO-A.ST</td>
<td>0.75</td>
<td>9</td>
<td>0.1231</td>
<td>10</td>
<td>0.0726</td>
<td>14.5</td>
</tr>
<tr>
<td>BOL.ST</td>
<td>0.77</td>
<td>7</td>
<td>0.1673</td>
<td>2</td>
<td>0.1273</td>
<td>25</td>
</tr>
<tr>
<td>ELUX-B.ST</td>
<td>0.80</td>
<td>4</td>
<td>0.1189</td>
<td>12</td>
<td>0.1053</td>
<td>20</td>
</tr>
<tr>
<td>ERIC-B.ST</td>
<td>0.84</td>
<td>1</td>
<td>0.0751</td>
<td>23</td>
<td>0.0816</td>
<td>17</td>
</tr>
<tr>
<td>GETI-B.ST</td>
<td>0.67</td>
<td>17</td>
<td>0.0831</td>
<td>22</td>
<td>0.0673</td>
<td>12</td>
</tr>
<tr>
<td>HM-B.ST</td>
<td>0.73</td>
<td>13</td>
<td>0.0911</td>
<td>21</td>
<td>0.0474</td>
<td>6.5</td>
</tr>
<tr>
<td>INVE-B.ST</td>
<td>0.29</td>
<td>29</td>
<td>0.1076</td>
<td>15</td>
<td>0.0449</td>
<td>4</td>
</tr>
<tr>
<td>LUPE.ST</td>
<td>0.52</td>
<td>24</td>
<td>0.0948</td>
<td>19</td>
<td>0.1633</td>
<td>26</td>
</tr>
<tr>
<td>MTG-B.ST</td>
<td>0.49</td>
<td>26</td>
<td>0.1234</td>
<td>9</td>
<td>0.1159</td>
<td>23</td>
</tr>
<tr>
<td>NOKI-SEK.ST</td>
<td>0.65</td>
<td>19</td>
<td>0.1221</td>
<td>11</td>
<td>0.3316</td>
<td>29</td>
</tr>
<tr>
<td>NDA-SEK.ST</td>
<td>0.78</td>
<td>6</td>
<td>0.1331</td>
<td>7</td>
<td>0.0726</td>
<td>14.5</td>
</tr>
<tr>
<td>SAND.ST</td>
<td>0.67</td>
<td>16</td>
<td>0.1578</td>
<td>3</td>
<td>0.1116</td>
<td>22</td>
</tr>
<tr>
<td>SCV-B.ST</td>
<td>0.66</td>
<td>18</td>
<td>0.0264</td>
<td>29</td>
<td>0.1272</td>
<td>24</td>
</tr>
<tr>
<td>SEB-A.ST</td>
<td>0.80</td>
<td>3</td>
<td>0.1385</td>
<td>5</td>
<td>0.0854</td>
<td>18</td>
</tr>
<tr>
<td>SECU-B.ST</td>
<td>0.44</td>
<td>27</td>
<td>0.0970</td>
<td>18</td>
<td>0.0661</td>
<td>11</td>
</tr>
<tr>
<td>SKA-B.ST</td>
<td>0.74</td>
<td>11</td>
<td>0.1094</td>
<td>14</td>
<td>0.0548</td>
<td>8</td>
</tr>
<tr>
<td>SKF-B.ST</td>
<td>0.82</td>
<td>2</td>
<td>0.1280</td>
<td>8</td>
<td>0.0780</td>
<td>16</td>
</tr>
<tr>
<td>SCA-B.ST</td>
<td>0.75</td>
<td>10</td>
<td>0.0739</td>
<td>25</td>
<td>0.0428</td>
<td>3</td>
</tr>
<tr>
<td>SSAB-A.ST</td>
<td>0.55</td>
<td>23</td>
<td>0.1797</td>
<td>1</td>
<td>0.1701</td>
<td>27</td>
</tr>
<tr>
<td>SHB-A.ST</td>
<td>0.57</td>
<td>22</td>
<td>0.1021</td>
<td>17</td>
<td>0.0657</td>
<td>10</td>
</tr>
<tr>
<td>SWED-A.ST</td>
<td>0.76</td>
<td>8</td>
<td>0.1354</td>
<td>6</td>
<td>0.1061</td>
<td>21</td>
</tr>
<tr>
<td>SWMA.ST</td>
<td>0.62</td>
<td>20</td>
<td>0.0284</td>
<td>28</td>
<td>0.0452</td>
<td>5</td>
</tr>
<tr>
<td>TEL2-B.ST</td>
<td>0.31</td>
<td>28</td>
<td>0.0602</td>
<td>26</td>
<td>0.1716</td>
<td>28</td>
</tr>
<tr>
<td>TLSN.ST</td>
<td>0.51</td>
<td>25</td>
<td>0.0746</td>
<td>24</td>
<td>0.0355</td>
<td>2</td>
</tr>
<tr>
<td>VOLV-B.ST</td>
<td>0.70</td>
<td>15</td>
<td>0.1485</td>
<td>4</td>
<td>0.0981</td>
<td>19</td>
</tr>
</tbody>
</table>

Table C.4: Parameters and ranks