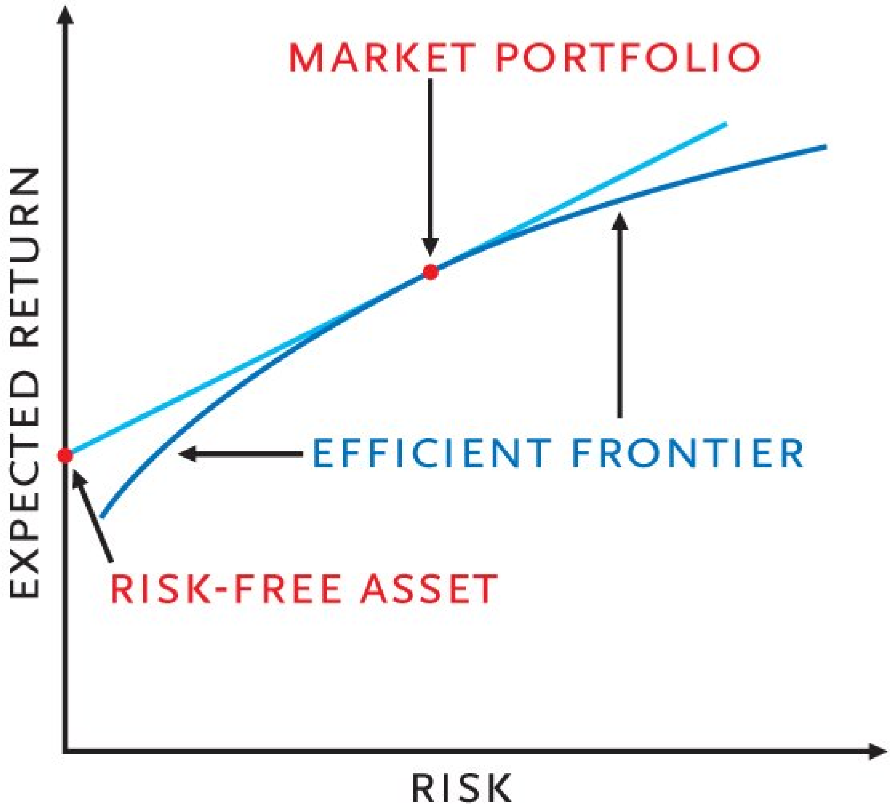
The Black-Litterman Model
Hype or Improvement?



Anisa Salomons
a.j.salomons@gmail.com

Abstract

This thesis explores a popular asset allocation model: the Black-Litterman model. First, an overview is given of the foundations of modern portfolio theory with the mean-variance model and the CAPM. Next, we discuss some improvements that could be made over the mean-variance model. The Black-Litterman model addresses some of these flaws and tries to improve them. The model has been described mathematically, and various definitions of the parameters are compared.
Finally, an empirical study has been performed to compare the performance of the Black-Litterman model to mean-variance optimization. The models have been compared in a three asset universe that consists of a momentum portfolio, a HML portfolio and a size portfolio. The views of the investor have been forecasted by a regression analysis on factors that describe the economic climate. The regression analysis also provides a consistent manner to specify the uncertainty on the views of the investor.
The conclusion can be drawn that BL-model improves on the mean-variance model, in our sample period, however the result is dependent on a well chosen benchmark.

Acknowledgment

First and foremost I would like to thank my supervisors Bas Bosma of Deloitte and Evert Vrugt from ABP Investments. They have meticulously corrected the texts I produced, even though the writing must sometimes have been awful to read. Their insistence on details have greatly improved my writing, the writing process and finally my thesis. It has been a pleasure to work together.
Of course, I would like to thank my professor Harry Trentelman. He made it possible for me to explore a subject that is not directly related to my graduation subject. His mathematical viewpoint provided a nice exchange with the finance, his views fine tuned the mathematics of my thesis.
Finally, it would like to thank my boyfriend Mathijs for his support, coping with my moodiness and my constant occupation of the computer during the final phases of the thesis.
Of course, all faults remain my own.

Contents

Chapter 1

Introduction

Portfolio selection is concerned with selecting a portfolio of investments that will fulfill the investment objectives over the investment horizon. What these objectives are differs per investor, but a positive and stable payoff on the investments is always desirable.
The portfolio selection problem is complex for at least two reasons, the large number of investment opportunities available and the difficulty to forecast the future. Aside from the many different investment opportunities that are available, it is nowadays also relatively easy to invest in nearly every country around the world. It is possible to not only invest in the Netherlands, but for a more risky investment one could choose, for example, investments in Russia or China. The possibility to invest globally expands the investment universe to nearly infinite size and makes it difficult for a person to examine all possibilities.
Investing is always a risky enterprise. An initial investment is made for a certain amount of money, but it is never certain that the value of the investment will increase. Even though there are numerous models to assist an investor in her investment decision, it is never possible to forecast the future with certainty. These two problems make it difficult to select a portfolio.
There are mainly two approaches to portfolio selection, a heuristic approach and a quantitative approach. In the heuristic approach the portfolio is selected with limited help of a model. The investor forms views about future performance of investments from news in the media. These views are used to select investments that are believed to have some favorable characteristics that the investor looks for in her investment portfolio.
The quantitative approach uses a mathematical model to make the final allocation of investments. The model evaluates the characteristics of the investments and determines which ones should be added to the portfolio. Harry Markowitz is the founder of quantitatively making investment decisions with his 1952 paper 'Portfolio selection'. He proposed that when determining an investment one should not only look at the possible pay-

off of the investment, but also take into account how certain one is that this payoff will actually be acquired. By formulating a mathematical model and making this trade-off explicit it became possible to allocate investments quantitatively. Although the model inspired a rich field of science and is used by many, it does have some important flaws. The model often results in counterintuitive portfolios, which poorly reflect the views of the investor. Investors worked around this problem by introducing extra constraints that would limit the range of the possible outcomes.

The two approaches to investing are depicted here more disjoint than they are in practice. Almost every investor nowadays uses the assistance of some sort of model in assembling an investment portfolio, at least to examine the characteristics of the portfolio.

This thesis examines a model that combines the two approaches in one model. The model was developed by Fisher Black and Robert Litterman of Goldman Sachs, their model is actually used at Goldman Sachs to determine investments. The first publication on the model was in 1990, and subsequently in 1991, 1992. Despite the multiple publications they never described the model very thoroughly. This thesis sets out to explain the model mathematically as well as conceptually.

Outline of the thesis The first chapter gives a background in the financial terminology and notation that will be used in the thesis. We next move to the basis of portfolio selection with the model of Markowitz and the subsequently developed capital asset pricing model.

Subsequently we arrive at the focus of this thesis: the Black-Litterman model.

The last chapter covers an empirical study of the Black-Litterman model in combination with some investment strategies.

Chapter 2

Financial terminology

The discussion of portfolio selection has its own vocabulary. The most important vocabulary and concepts will be discussed in this chapter, as well as the accompanying notation. These are basic finance concepts like asset, return, risk, portfolio and diversification.

2.1 Assets and its characteristics

2.1.1 Asset classes[1]

An investor can choose from thousands of different assets. Not only are there many ordinary shares to choose from, but there are also other investment opportunities available. These opportunities can be divided in classes of assets with the same characteristics.
The most well known asset class is *equity*, also known as ordinary shares or in the United States as stock. Equity is the ownership of a part of a company. Equity of public companies can be traded on a stock exchange.
Another asset class is bonds. A *bond* is a debt certificate issued by a borrower to a lender. The debt certificate says that the borrower owes the lender a debt and is obliged to repay the principal and interest at a later date. The later date is set at the issuance of the bond and the interest rate can be fixed or variable. Bonds, as every investment, vary in the degree of risk attached to them. The length of the borrowing period and the entity that issues the bonds are important risk factors. Short term government bonds are generally regarded as very safe investments.
The final class under consideration is currency. Investing in foreign currency can be useful either to bet on a change in the exchange rate or to insure, or hedge, investments in that currency against changes in the exchange rate.
There are also other asset classes, but they are not of interest to the present

[1] The definitions are derived from Smullen and Hand (2005) and Moles and Terry (1997).

research. In general the word asset will be used to describe an investment, when it matters which class of assets is under consideration, we will be more specific.

2.1.2 Return and risk

The investor invests her money in a portfolio of assets and is interested in making a profit. More specifically she is interested in making a profit relative to the invested money. The rate of return (*return* for short) is the difference between the amount of invested money at the start and the value of the investment at the end of the period plus some additional (net) cash flows as for example dividend, divided by the starting value, i.e.

$$\text{return} = \frac{\text{profit}}{\text{invested money}} = \frac{\text{end value} + \text{cash flows} - \text{starting value}}{\text{starting value}}.$$

There are other definitions of return, but this one suffices for our purposes. Return is defined for a period in the past, but in asset allocation one is interested in the future behavior of an asset, the future return. Markowitz (1959) expresses the future or forecasted return as the expected value of the return.

What is actually meant by the term expected return $\mathrm{E}(r)$, is a forecast of the return, as we want to forecast the future return of the asset. If r_t denotes the return up to time t, then $\mathrm{E}(r)$ is shorthand for $\mathrm{E}(r_{t+1}|I_t)$ which means the forecast of the return at time $t+1$ given all the information up to and including time t. The expected return is one of two important characteristics of an asset relevant to mean-variance optimization.

The other important characteristic is the *risk*. Intuitively risk should mean something like the chance that one looses on a investment. However, mathematically it is quit challenging to define it and several different definitions exist. Markowitz (1952) defined risk as the variance of the return. Variance measures the deviation around a point, negative deviations from this point as well as positive deviations. In the case of an investment, the variance of return measures the deviation of the return around the expected return. An investor would only consider less expected return, i.e. a negative deviation as a risk, while positive deviations also add to the the variance of return, which makes variance a counterintuitive measure of risk. In reaction to this measure, others have been proposed that measure risk differently, or only measure negative deviations.

The square root of the variance is called the standard deviation in mathematics, in finance this measure is called the *volatility*.

Markowitz (1952) postulates that the only two measures of interest of an asset are its expected return and its risk, measured in variance.

2.1.3 Expected return and variance of return

Portfolio theory relies heavily on the probabilistic measures expected value, covariance and variance. We will recall their definitions and some important properties of these measures. For a more in-depth treatment one could read for example the book of Johnson and Wichern (1998).

Definition

Expected value The sample expected value also known as sample mean of a sample $x_i,\ i = 1 \ldots n$ is defined as the average value of the sample: $\mathrm{E}(X) = \frac{1}{n}\sum_{i=1}^{n} x_i$. The mean of a variable is often denoted by the Greek letter μ.

Variance The variance is a measure of how much a variable varies around the expected value (μ). One definition of the sample variance is: $\mathrm{var}(X) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \mu)^2 = \sigma^2$, where μ is the mean value of the sample.

Standard deviation The standard deviation (σ) is the square root of the variance, in finance this is often called the volatility. The standard deviation is a more intuitive measure of variability than variance, as it is more directly related to the normal distribution. The normal distribution and its relation with mean and volatility is explained in the next paragraph.

Correlation coefficient The correlation coefficient measures how two random variables are co related when a linear relationship is understood between the two variables. If two random variables X and Y have a linear relationship of the form $Y = aX + b$, then the factor a is defined as the correlation coefficient ρ.
The correlation coefficient varies between -1 and 1. Two variables are perfectly positive correlated if the correlation coefficient $\rho = 1$, this means that if X increases so does Y, and by the same amount. A perfect negative correlation is found when $\rho = -1$, this means that if X increases, Y decreases and Y decreases by the same amount. Two random variables are called uncorrelated if $\rho = 0$.

Covariance The covariance provides a measure of the strength of the correlation between two random variables. The covariance of X_1 and X_2 can be related to the correlation coefficient ρ of the two random variables and the respective standard deviations (σ_1 and σ_2). They are related as follows: $\mathrm{cov}(X_1, X_2) = \rho\sigma_1\sigma_2 = \sigma_{12}$, if X_1 and X_2 are independent then the correlation coefficient $\rho = 0$ as is the covariance $\mathrm{cov}(X_1, X_2) = 0$.

Properties The expected value, variance and covariance have a few useful properties. The proof of these properties will be omitted, the reader is referred to a good book on statistics like Johnson and Wichern (1998). Let a and b be real valued scalars, A a real valued matrix and let X, Y and X_i be random variables then the following properties hold,

$$\mathrm{E}(aX+b) = a\mathrm{E}(X) \tag{2.1}$$
$$\mathrm{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{E}(X_i) \tag{2.2}$$
$$\mathrm{cov}(X,Y) = \mathrm{cov}(Y,X) \tag{2.3}$$
$$\mathrm{cov}(X,X) = \mathrm{var}(X) \tag{2.4}$$
$$\mathrm{cov}(aX,bY) = ab\,\mathrm{cov}(X,Y) \tag{2.5}$$
$$\mathrm{var}(aX+bY) = a^2\mathrm{var}(X) + b^2\mathrm{var}(Y) + 2ab\,\mathrm{cov}(X,Y) \tag{2.6}$$
$$\mathrm{var}\left(\sum_{i=1}^{n} X_i\right) = \mathrm{cov}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j\right) \tag{2.7}$$
$$\mathrm{var}(AX) = A\mathrm{var}(X)A'. \tag{2.8}$$

Normal distribution It is often assumed that the distribution of the asset returns have a normal probability distribution, for example by Black and Litterman (1991a). When an investors estimates that an asset has an expected return of 4% with a variance of $(1\%)^2$, then the probability distribution of the return could be drawn in a graph, see figure 2.1.

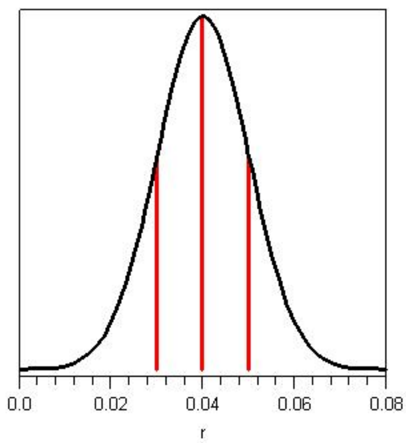


Figure 2.1: The normal probability distribution.

The values that the return can assume are plotted on the horizontal axis, the mean can be found in the middle at 4%, the variance of $(1\%)^2$, or equivalently

the standard deviation of 1%, can be seen in the figure at $r = \mu - \sigma = 3\%$ and $r = \mu + \sigma = 5\%$.
The investor has implicitly professed the opinion that with a probability of 68% the return will be one standard deviation away from the mean, thus there is a 68% probability that return is in the interval 3% to 5%. This is a well known property of the normal distribution, another often used rule of thumb is that there is 95% probability that the returns are in the two standard deviation interval around the mean, i.e. from 2% to 6%.

2.2 The portfolio of assets

2.2.1 Portfolio consists of assets

A portfolio consists of various assets, where the proportion of an asset in the total value of the portfolio is called its *weight*. A portfolio could for example consist for one third of equity ASML, half could be equity of KPN and one sixth equity of Aegon.
A portfolio consisting of n assets, is represented mathematically by a vector $\mathbf{w} \in \mathbb{R}^n$. This would make the vector of weights in this example $\mathbf{w} = \left(\frac{1}{3}, \frac{1}{2}, \frac{1}{6}\right)'$.
The weights in the vector are proportional to the total portfolio and therefore have to sum to one, $\sum_{i=1}^{n} w_i = 1$.

Definition 1 (portfolio)**.** *A portfolio consisting of n assets is represented by a vector $\mathbf{w} \in \mathbb{R}^n$ such that $\sum_{i=1}^{n} w_i = 1$.*

2.2.2 Borrowing and lending assets

The weights in the portfolio need not be positive, it is possible to borrow or lend an asset. Borrowing an asset would lead to a negative weight in the portfolio.
Borrowing makes sense if the investor anticipates a price decrease of an asset. For example, supposes an investor borrows an asset that is worth 20 euro at time $t = 0$, and she is obliged to return the asset at time $t = 3$. The moment she receives the asset, she sells the asset on the stock market. At time $t = 3$ she has to return the asset. If at this time the price of the asset has decreased to for example 15 euro, she makes a profit. She buys the asset on the stock market for 15 euro and returns the asset. She has now made of profit of $20 - 15 = 5$ and a return of $\frac{-15+20}{20} = 25\%$.
Borrowing an asset, taking a negative position is called *shorting* an asset. A positive position in an asset is called going *long*. Not every investor is allowed to short assets, pension funds are often prohibited from this practice.
Borrowing of assets makes it possible to invest in a portfolio without investing any own funds, such a portfolio is called a *zero-investment portfolio*. In

such a portfolio, the long positions are financed by short positions in other asset, the end result is a portfolio whose weights sum to zero.

2.2.3 Notation

The concepts risk and return can be given a mathematical representation. The return of asset i is denoted by r_i. The expected return of asset i becomes $\mathrm{E}(r_i)$, its variance σ_i^2 and the covariance of asset i and j is σ_{ij}.
For a portfolio that consists of n assets, the return of each asset in the portfolio is captured by the vector of returns $\mathbf{r} \in \mathbb{R}^n$. The vector of returns also has an expected value, $\mathrm{E}(\mathbf{r}) \in \mathbb{R}^n$. The covariance and the variance of the assets in the portfolio are represented in the covariance matrix $\Sigma \in \mathbb{R}^{n \times n}$, the diagonal entries of which are formed by the variance of the assets ($\sigma_{ii} = \sigma_i^2$) as this is the covariance of an asset with itself.
The covariance matrix is symmetric, this is due to the symmetry of the covariance, see equation (2.3).
The return of the portfolio (r_p) is determined by the return of the assets in the portfolio: $r_p = \sum_{i=1}^n w_i r_i$. The expected value and the variance of the return of the portfolio follow after straightforward computation from the properties of the expected value, covariance and variance.

Proposition 1. *The expected return* $\mathrm{E}(r_p)$ *of a portfolio is* $\mathbf{w}'\mathrm{E}(\mathbf{r})$.
The variance of return of a portfolio is $\mathrm{var}(r_p) = \mathbf{w}'\Sigma\mathbf{w}$.

Proof. The proof follows from the properties of the expected value and the variance as described in equations (2.1) to (2.7).

$$\begin{aligned}
\mathrm{E}(r_p) = \mathrm{E}\left(\sum_{i=1}^n w_i r_i\right) &=^1 \sum_{i=1}^n \mathrm{E}\left(w_i r_i\right) \\
&=^2 \sum_{i=1}^n w_i \mathrm{E}(r_i) = \mathbf{w}'\mathrm{E}(\mathbf{r})
\end{aligned}$$

$=^1$ *Property* (2.2) *is used to interchange the summation and the expected value operator.*
$=^2$ *Propery* (2.1) *is used, constants are invariant to the expected value operator.*

The variance of return for the portfolio can be computed in the following way.

$$\begin{aligned}
\mathrm{var}(r_p) = \mathrm{var}\left(\sum_{i=1}^n w_i r_i\right) &=^1 \sum_{i=1}^n \sum_{j=1}^n \mathrm{cov}(w_i r_i, w_j r_j) \\
&=^2 \sum_{i=1}^n \sum_{j=1}^n w_i w_j \mathrm{cov}(r_i, r_j) =^3 \sum_{i=1}^n \sum_{j=1}^n w_i w_j \sigma_{ij} \\
&= \mathbf{w}'\Sigma\mathbf{w}
\end{aligned}$$

$=^1$ *Property (2.7) is used to change the variance of the summation into the summations of the covariances.*
$=^2$ *Property (2.5) is used, constants are invariant to the covariance operator.*
$=^3$ *The covariance of asset i and asset j is alternatively written as σ_{ij}.*

□

2.3 Rationale for diversification[2]

One of the important notions in asset allocation is that diversification reduces risk. There is even an English proverb that supports diversification: "Don't put all your eggs in one basket". Although it seems a reasonable investment rule it is good to investigate mathematically under which circumstances diversification reduces risk.
The effect of diversification can be quantified by using the formulas for the sum of variances. We will not give a general proof of the conditions under which diversification diminishes the variance of the portfolio, as there are many parameters that can be varied. Instead we will consider two instances of diversification. In the first instance all n assets have mutually uncorrelated returns, that is $\sigma_{ij} = 0$ for all assets i unequal to j. All the assets have equal expected return $\mathrm{E}(r)$ and variance σ^2. The portfolio will be constructed from an equal weighting scheme, i.e. taking equal proportions of each asset: $w_i = \frac{1}{n}$ for each asset i. This makes the expected return of the portfolio $\mathrm{E}(r_p)$ equal to the expected return of an individual asset in the portfolio.

$$\mathrm{E}(r_p) = \sum_{i=1}^{n} w_i \mathrm{E}(r_i) = \frac{1}{n} \sum_{i=1}^{n} \mathrm{E}(r) = \mathrm{E}(r).$$

The expected return of the portfolio is in this example independent of the number of assets in the portfolio. The variance of this portfolio $\mathrm{var}(r_p)$, however, does depend on the number of assets:

$$\begin{aligned}\mathrm{var}(r_p) &= \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij} =^1 \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_{ij} \\ &=^2 \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{\sigma^2}{n}\end{aligned}$$

$=^1$ *$\frac{1}{n}$ is substituted for the weights w_i.*
$=^2$ *It is used that $\sigma_{ij} = 0$ for all $i \neq j$ and $\sigma_{ii} = \sigma^2$.*

[2]The text in this section is derived from Luenberger (1998).

It can be seen that the variance does depend on the number of assets in the portfolio and it decreases as the number of assets increases. The variance of the portfolio even approaches zero as the number of assets grows to infinity.

$$\lim_{n\to\infty} \text{var}(r_p) = \lim_{n\to\infty} \left(\frac{1}{n^2} \sum_{i=1}^{n} \sigma^2 \right) = \lim_{n\to\infty} \left(\frac{\sigma^2}{n} \right) = 0$$

Hence, for uncorrelated assets, with equal expected return, diversification reduces the variance of the portfolio and can eliminate it altogether, while the expected return of the portfolio remains the same. This implies that it would be best to compile a portfolio of as many as possible mutually uncorrelated assets with equal weight, as this allows the variance to be reduced to zero in the limit.

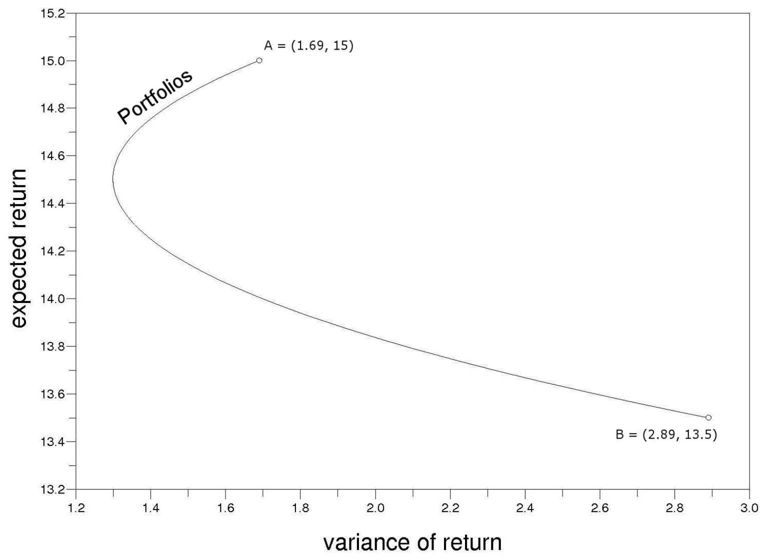


Figure 2.2: Diversification diminishes the variance of the portfolio.

This can best be illustrated by an example.

Example 1. *Assume we have two assets A and B. Asset A has expected return of* 13.5%*, asset B has an expected return of* 15%*. In vector notation:* $\text{E}(\mathbf{r}) = \begin{pmatrix}13.5\% & 15\%\end{pmatrix}'$.
The first asset has a volatility of 17%*, the second asset has volatility* 13% *and the correlation between the assets is* 0.23*. We would like to construct the covariance matrix of these two assets, therefore the volatility needs to be squared to obtain the variance, correlation coefficient and the asset volatilities needs to be multiplied to obtain the covariance of the assets. This gives*

$\sigma_{11} = \sigma_1^2 = (17\%)^2 = 289(\%)^2$, $\sigma_{22} = \sigma_2^2 = (13\%)^2 = 169(\%)^2$, and covariance $\sigma_{12} = \rho\sigma_1\sigma_2 = 0.23 \cdot 17\% \cdot \mathbf{13}\% = 51(\%)^2$. In matrix notation $\Sigma = \begin{pmatrix} 289(\%)^2 & 50.8(\%)^2 \\ 50.8(\%)^2 & 169(\%)^2 \end{pmatrix}$.
We will construct portfolios that consists of these two assets. Starting with the portfolio that only consists of asset A, slowing adding asset B until the portfolio consists only of this asset. This can be accomplished by assigning weight λ and $1-\lambda$ to asset A and B respectively, where λ will vary between zero and one. When $\lambda = 1$ then the portfolio consists of asset A, when $\lambda = 0$ then the portfolio consists of asset B. The weight vector of the portfolio thus is $\mathbf{w} = \begin{pmatrix}\lambda & 1-\lambda\end{pmatrix}'$.
The expected return of the portfolios is then computed by $\mathrm{E}(r_p) = \mathbf{w}'\mathrm{E}(\mathbf{r}) = \lambda\mathrm{E}(r_1) + (1-\lambda)\mathrm{E}(r_2) = \lambda\, 13.5\% + (1-\lambda)15\%$. The variance of the portfolio equals $\mathrm{var}(r_p) = \mathbf{w}'\Sigma\mathbf{w} = \sum_{i=1}^n \sum_{j=1}^n w_i w_j \sigma_{ij} = \lambda^2\sigma_1^2 + 2\lambda(1-\lambda)\sigma_{12} + (1-\lambda)^2\sigma_2^2 = \lambda^2 289(\%)^2 + 2\lambda(1-\lambda)50.8(\%)^2 + (1-\lambda)^2 169(\%)^2$.
The resulting portfolios are plotted in figure 2.2. It can be seen that the variance of the portfolio can be diminished by combining the two assets. This is due to the correlation coefficient of the assets, which is not very large and thus makes the covariance of the assets much smaller than the individual variances. Therefore portfolios which are a combination of these two assets will have a lower variance, than portfolios than consists of the single assets.

The next situation under consideration, is the situation where the returns of the assets are correlated. Suppose, as before, that each asset has an expected return of $\mathrm{E}(r)$ and variance σ^2, but now each return pair has a covariance of $\sigma_{ij} = a\sigma^2$ for $i \neq j$ and $a \in \mathbb{R}$. The portfolio consists again of n equally weighted assets. The result is that the expected return of the portfolio is the same as in the previous instance, but the variance is different:

$$\begin{aligned}
\mathrm{var}(r_p) &= \sum_{i=1}^n \sum_{j=1}^n w_i w_j \sigma_{ij} = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} \\
&=^1 \frac{1}{n^2}\left(\sum_{i=j} \sigma_{ij} + \sum_{i\neq j} \sigma_{ij}\right) =^2 \frac{1}{n^2}\left(\sum_{i=j} \sigma^2 + \sum_{i\neq j} a\sigma^2\right) \\
&= \frac{1}{n^2}\left(n\sigma^2 + (n^2-n)a\sigma^2\right) \\
&= \frac{\sigma^2}{n} + \left(1 - \frac{1}{n}\right)a\sigma^2 = \frac{(1-a)\sigma^2}{n} + a\sigma^2
\end{aligned}$$

$=^1$ *The single summation is divided in two separate summations, one for $i = j$ and one for $i \neq j$.*
$=^2$ *The variance is the same for all $i = j$: $\sigma_{ii} = \sigma^2$, the covariance is $\sigma_{ij} = a\sigma^2$ for all $i \neq j$.*

If the assets are mutually correlated with $a\sigma^2$, then the variance cannot be eliminated altogether by increasing the number of assets.

$$\lim_{n\to\infty} \mathrm{var}(r_p) = \lim_{n\to\infty} \left(\frac{(1-a)\sigma^2}{n} + a\sigma^2 \right) = a\sigma^2$$

This analysis of diversification is somewhat crude, for it is assumed that all assets have the same expected return and the covariances have a simple structure. In general, diversification may reduce the overall expected return, while reducing the variance. How much expected return needs to be traded for a lowering of risk, is an important question.

2.4 Summary

This chapter can be summarized in a few sentences. The portfolio of assets is represented by a vector of weights $\mathbf{w} \in \mathbb{R}^n$, such that $\sum_{i=1}^{n} w_i = 1$.
Every asset has an expected return ($\mathrm{E}(r_i)$) and risk measured by the variance of return (σ_i^2). The portfolio of assets also has an return: $\mathrm{E}(r_p) = \mathbf{w}'\mathrm{E}(\mathbf{r})$ and a variance $\mathrm{var}(r_p) = \mathbf{w}'\Sigma\mathbf{w}$.
Adding not perfectly correlated assets to the portfolio reduces the variance of the portfolio.

Chapter 3

Mean-variance optimization

Quantitative asset allocation originates from the work of Markowitz in 1952. Nowadays, Markowitz' mean-variance optimization is still the basis for quantitative asset allocation. This also holds for the main subject of this this report the model of Black and Litterman. Therefore, it remains relevant to discuss the model of Markowitz prior to the discussion of our main subject. The model will first be derived conceptually to form a general idea of the intricacies of the subject. Subsequently, the mathematical formulation of the model will be discussed as well as some strong points and some drawbacks of the model. Finally, the capital asset pricing model (CAPM) will be discussed. The CAPM can be used to determine the expected return of an asset under certain assumptions. The CAPM follows from Markowitz's model and together they form the basis of modern portfolio theory.
The main text in this chapter is derived from Luenberger (1998), Markowitz (1987) and Sharpe (1964).

3.1 Model development

To develop an investment model it is good to have an idea of the way in which investors select a portfolio. An investor follows the economic news in order to form views on which markets, sectors or specific companies are going to perform better and which are going to perform worse. However, these views alone are not enough to select a portfolio. Such views have to be translated in a tractable form. How does one, for example, translate a view that the American economy will outperform the European economy in an asset allocation? Quantitative models can give guidance in asset allocation, after the initial translation from idea to input has been done. A model not only needs input that can be optimized, it also needs an objective function. This is a function that describes the allocation process and that the investor wants to optimize.

The objective function An investor aims to make a positive return on her investments. Therefore, the expected return of the portfolio of assets would seem to be to a logical objective function.
To maximize the expected return of the portfolio, one has to simply invest in the single asset with the highest expected return. Addition of other assets with a lower expected return to the portfolio would lower the expected return of the portfolio.
Investing in a single however, asset goes against the notion of diversification and the result will be that the performance of the portfolio is based on the performance of the one asset. If the asset performs well, so does the investment. It also means the converse, if the asset performs badly, so does the portfolio. The performance of the investments solely depends on the one asset. This makes the return very erratic and therefore very risky. It would be better if the return could be more steady.
Therefore, it is also important to investigate the risk involved in the portfolio. One should not only maximize the expected return but also take into account the risk of the portfolio; one should balance the risk and the expected return of the portfolio. A less risky result could, for example, be obtained by diversifying the portfolio with investments in companies in different sectors. If the equity in one sector perform poorly, it could be that equity in another sector do well. This diversification leads to a more steady expected return of the portfolio. However, probably some expected return has to be traded to obtain this less risky portfolio. It would be desirable if the trade-off between risk and expected return would become explicit.
An investor that only takes on additional risk, in exchange for additional expected return is called a *risk averse investor*. Risk aversion is thought to best describe human investment behavior.
There is a specific field concerned with defining functions that can categorize preferences and formalizes the principle of risk aversion, this field is called *utility theory*.
The reasoning in this paragraph leads to the following conclusion. The objective should be to maximize expected return for a certain level of risk, or equivalently minimize risk for a certain level of expected return.

Markowitz Harry Markowitz (1952), wrote the seminal paper on quantitative portfolio selection. He identified the forecasted return with expected return and risk with variance of return. He went on to postulate that the above objective is the one to strive for and developed a mathematical model for portfolio selection.
The objective, in terms of variance becomes to minimize the variance of return for a certain level of expected return. The expected value is often called the mean value. Therefore this kind of optimization is called *mean-variance (MV)* optimization.

Markowitz (1987) defines a portfolio that minimizes variance for a certain level of expected return, or equivalently maximizes expected return for a certain level of variance as an *efficient portfolio*.

Definition 2 (Efficient portfolio)**.** *A portfolio is mean-variance efficient if there does not exist another portfolio with a higher mean and no higher variance or less variance and no less mean.*

The efficient frontier can be depicted in a risk-return diagram, see figure 3.1. In this example we have three assets, these assets can be combined to form different portfolios. After weights are selected for the portfolio, the variance of the portfolio $\mathbf{w}'\Sigma\mathbf{w}$ and the expected return $\mathbf{w}'\mathbf{r}$ can be computed. The risk and the expected return of the portfolio is then depicted in the risk-return diagram.
The curve of efficient portfolios is called the efficient frontier, it has the form of a tipped parabola. There do not exist portfolios beyond the efficient frontier which have less variance for the level of expected return.
The top of the parabola is the minimum variance portfolio, the portfolio that has the minimal variance over all possible combinations in our investment universe. The portfolios that correspond to the preferences of a risk-averse investor are located on the top half of the parabola. These portfolios all have the characteristic that they have the maximal expected return for a certain level of variance.

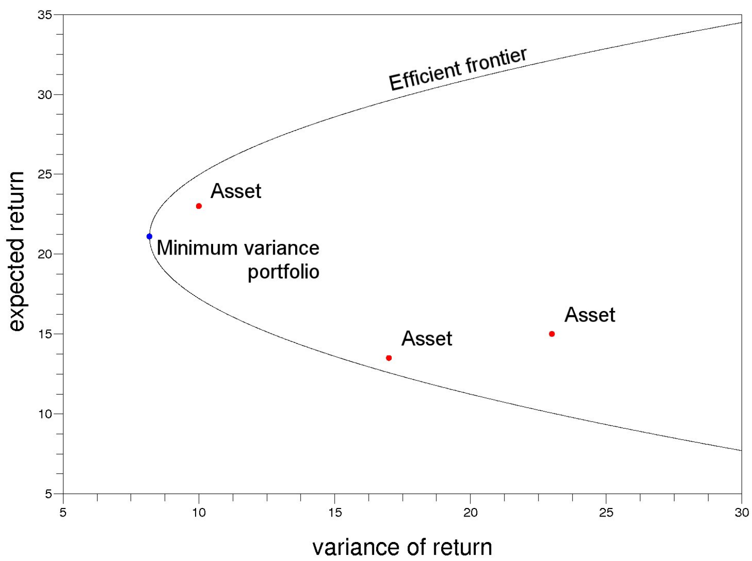


Figure 3.1: Efficient frontier of a three asset portfolio.

3.2 The mathematics of MV-optimization

The general idea of mean-variance analysis has been explained in the previous paragraph. In this paragraph we will concentrate on the mathematical formulation of the optimization problem and its solutions. In general we distinguish between three cases for the optimization problem, classified according to the constraints on the problem. One of the main advantages of mean-variance optimization is the flexibility to cope with the extra restrictions an investor faces in practice. These restrictions leads to mathematically different problems.
The section starts with standard mean-variance optimization without constraints. Subsequently we will move to mean-variance optimization with equality constraints, and finally we will discuss mean-variance optimization with inequality constraints.

3.2.1 Unconstrained mean-variance optimization

The objective of mean-variance optimization is to maximize the expected return of a portfolio of assets for a given level of risk. If each of the n assets in the portfolio has a weight $w_i \in \mathbb{R}$, then mean-variance optimization needs to determine the optimal allocation of weights. This can be accomplished by determining the solution to $\max_{\mathbf{w}\in\mathbb{R}^n} \mathrm{E}(r_p)$ subject to $\mathrm{var}(r_p) = c$ where c is the desired level of risk. In subsection 2.2.3 it has been shown that the expected return and variance of return of a portfolio are $\mathbf{w}'\mathrm{E}(\mathbf{r})$ and $\mathbf{w}'\Sigma\mathbf{w}$ respectively. The standard mean-variance optimization problem can therefore be formulated as follows:

Problem 1 (Standard MV-optimization)**.**

$$\begin{aligned} &\max_{\mathbf{w}\in\mathbb{R}^n} \mathbf{w}'\mathrm{E}(\mathbf{r}) \\ &\text{subject to } \frac{1}{2}\mathbf{w}'\Sigma\mathbf{w} = c \end{aligned}$$

The factor $\frac{1}{2}$ is only introduced for convenience, it does not alter the problem. The solution can be found via the method of Lagrange. The Lagrangian becomes $L(\mathbf{w}, \lambda) = \mathbf{w}'\mathrm{E}(\mathbf{r}) + \lambda(\frac{1}{2}\mathbf{w}'\Sigma\mathbf{w} - c)$. The solution $\mathbf{w}^*$ has to fulfill simultaneously $\dfrac{\partial L}{\partial \mathbf{w}} = \mathbf{0}$ and $\dfrac{\partial L}{\partial \lambda} = 0$.

$$\begin{aligned} \frac{\partial L}{\partial \mathbf{w}} &= \mathrm{E}(\mathbf{r}) + \lambda\frac{1}{2}[\Sigma\mathbf{w} + (\mathbf{w}'\Sigma)'] = \mathbf{0} \\ &=^1 \mathrm{E}(\mathbf{r}) + \lambda\Sigma\mathbf{w} = \mathbf{0} \qquad (3.1) \\ \frac{\partial L}{\partial \lambda} &= \frac{1}{2}\mathbf{w}'\Sigma\mathbf{w} - c = 0 \qquad (3.2) \end{aligned}$$

$=^1$ *It is used that the covariance matrix is symmetric,* $\Sigma' = \Sigma$, *therefore* $(\mathbf{w}'\Sigma)' = \Sigma'\mathbf{w} = \Sigma\mathbf{w}$.

From equation (3.1) a formula for $\mathbf{w}^*$ is obtained: $\mathbf{w}^* = -(\lambda\Sigma)^{-1}\mathrm{E}(\mathbf{r})$. This formula is substituted in equation (3.2) to obtain the value of the parameter $\lambda = -\sqrt{\frac{1}{2c}\mathrm{E}(\mathbf{r})'\Sigma^{-1}\mathrm{E}(\mathbf{r})}$:

$$\begin{aligned}
c &= \frac{1}{2}\mathbf{w}'\Sigma\mathbf{w} =^1 \frac{1}{2}[(\lambda\Sigma)^{-1}\mathrm{E}(\mathbf{r})]'\Sigma(\lambda\Sigma)^{-1}\mathrm{E}(\mathbf{r}) \\
c &=^2 \frac{1}{2}\mathrm{E}(\mathbf{r})'(\lambda\Sigma)^{-1}\Sigma(\lambda\Sigma)^{-1}\mathrm{E}(\mathbf{r}) =^3 \frac{1}{2}\lambda^{-2}\mathrm{E}(\mathbf{r})'\Sigma^{-1}\mathrm{E}(\mathbf{r}) \Rightarrow \\
\lambda &= -\sqrt{\frac{1}{2c}\mathrm{E}(\mathbf{r})'\Sigma^{-1}\mathrm{E}(\mathbf{r})}. \qquad (3.3)
\end{aligned}$$

$=^1$ *The optimal weight* $\mathbf{w}^* = -(\lambda\Sigma)^{-1}\mathrm{E}(\mathbf{r})$ *is substituted.*
$=^2$ *The square brackets are expanded:* $[(\lambda\Sigma)^{-1}\mathrm{E}(\mathbf{r})]'$ $= \mathrm{E}(\mathbf{r})'((\lambda\Sigma)^{-1})'$ *and it is used that* Σ *is symmetric hence* $\mathrm{E}(\mathbf{r})'((\lambda\Sigma)^{-1})' = \mathrm{E}(\mathbf{r})'((\lambda\Sigma)')^{-1} = \mathrm{E}(\mathbf{r})'(\lambda\Sigma)^{-1}$.
$=^3$ *The lambdas are grouped together by expanding the inverse of* $(\lambda\Sigma)^{-1} = \lambda^{-1}\Sigma^{-1}$.

The parameter λ is often called the risk-aversion parameter, as it represents a risk-averse investor for values of $\lambda \leq 0$.
In the usual formulation of the mean-variance problem, the risk level is not set beforehand, the problem is simultaneously solved for all risk levels. The level of risk is then not varied by the parameter c but is varied by the risk-aversion parameter λ. This allows the formulation of the mean-variance problem in one single formula without constraints.

Problem 2 (Unconstrained MV-optimization)**.**

$$\max_{\mathbf{w}\in\mathbb{R}^n} \mathbf{w}'\mathrm{E}(\mathbf{r}) - \frac{1}{2}\lambda\mathbf{w}'\Sigma\mathbf{w} \quad \forall\lambda \geq 0.$$

This is an ordinary maximization problem without any constraints, hence the solution can be found by differentiation to the vector of weights $\mathbf{w}$. This objective function is equal to the Lagrangian of Problem 1 except for the sign of λ, in this case it has a negative sign in the previous problem it had a positive sign. The sign of λ does not alter the problem, it only alters the set for which λ describes risk-aversion, previously this was for $\lambda \leq 0$ now it is for $\lambda \geq 0$. The vector that maximizes this problem, the optimal vector of weights equal $\mathbf{w}^* = (\lambda\Sigma)^{-1}\mathrm{E}(\mathbf{r})$.

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\mathbf{w}}(\mathbf{w}'\mathrm{E}(\mathbf{r}) - \frac{1}{2}\lambda\mathbf{w}'\Sigma\mathbf{w}) = \mathrm{E}(\mathbf{r}) - \lambda\Sigma\mathbf{w} = \mathbf{0} \\
&\Rightarrow \mathbf{w}^* = (\lambda\Sigma)^{-1}\mathrm{E}(\mathbf{r})
\end{aligned}$$

By varying the risk aversion parameter λ the optimal solutions $\mathbf{w}^*$ can be found to the differing problems. These solutions in turn can be plotted in a risk-return graph, by computing the expected return and the variance of the portfolio for the different solutions. We have already seen the line in figure 3.1, in this case λ needs to be positive for risk-averseness. The positive value corresponds to the more general definition of risk-aversion that will be discussed in paragraph 3.3.1.

The covariance matrix The covariance matrix plays an important role in mean-variance optimization and has to meet two requirements. To determine the optimal weights of the assets in the portfolio, the covariance matrix needs to be inverted $\mathbf{w}^* = (\lambda\Sigma)^{-1}\mathrm{E}(\mathbf{r})$. A basic requirement for matrix inversion is that the determinant of $\Sigma \neq 0$, this is known as the matrix Σ is nonsingular.
The other important property for the covariance matrix can be deduced from equation (3.3), where the square root of $\mathrm{E}(\mathbf{r})'\Sigma^{-1}\mathrm{E}(\mathbf{r})$ is taken. In order for the square root to have real values it is needed that $\mathrm{E}(\mathbf{r})'\Sigma^{-1}\mathrm{E}(\mathbf{r}) > 0$ for all values of $\mathrm{E}(\mathbf{r})$. The property is called positive definiteness: the matrix Σ^{-1} needs to be positive definite. It can be seen that Σ^{-1} is positive definite if and only if Σ is positive definite, see a good book on linear algebra like Graham (1987).
It is well known that if Σ is positive definite, then it is nonsingular. Therefore, we make the following assumption.

Assumption. *The covariance matrix is positive definite ($\Sigma > 0$).*

The covariance matrix, under this assumption, should give no problems during inversion. However care has to be taken, when the covariance matrix is estimated. Since, if the estimation procedure is flawed, this could lead to singular or nearly singular matrices that are therefore not invertible.

3.2.2 Equality constraints

The simple instances of mean-variance optimization, outlined in the previous subsection, are not very realistic, since normally an investor faces constraints. At least there should be a constraint that requires to invest all the available resources. This is called a full investment constraint. This constraint can be translated into a formula that requires that the weights of the assets in the portfolio have to sum to one, $\mathbf{1}'\mathbf{w} = 1$.
Other equality constraints could be imposed, these k restrictions can be represented by a matrix $A \in \mathbb{R}^{k\times n}$ and a vector $\mathbf{b} \in \mathbb{R}^k$ such that $A\mathbf{w} = \mathbf{b}$. Adding the equality constraints transforms the unconstrained optimization problem of Problem 2 to a Lagrange problem, which is relatively easy to solve.

Problem 3 (MV with equality constraints)**.**

$$\max_{\mathbf{w}\in\mathbb{R}^n} \mathbf{w}'\mathrm{E}(\mathbf{r}) - \frac{1}{2}\lambda\mathbf{w}'\Sigma\mathbf{w} \qquad \forall\lambda \geq 0.$$
$$\text{subject to } A\mathbf{w} = \mathbf{b}$$

The Lagrangian of this problem is

$$L(\mathbf{w},\boldsymbol{\gamma}) = \mathbf{w}'\mathrm{E}(\mathbf{r}) - \frac{1}{2}\lambda\mathbf{w}'\Sigma\mathbf{w} + \boldsymbol{\gamma}'(A\mathbf{w} - \mathbf{b}),$$
$$\text{where } \boldsymbol{\gamma} \in \mathbb{R}^k \text{ is a Lagrange multiplier}$$

A necessary and sufficient condition for the existence of a solution $\mathbf{w}^*$ is the existence of a vector of weights that simultaneously fulfills the equations: $\frac{\partial L}{\partial \mathbf{w}} = \mathbf{0}$ and $\frac{\partial L}{\partial \boldsymbol{\gamma}} = \mathbf{0}$.

$$\begin{aligned} \frac{\partial L}{\partial \mathbf{w}} &= \mathrm{E}(\mathbf{r}) - \lambda\Sigma\mathbf{w} + (\boldsymbol{\gamma}'A)' = \mathbf{0} \\ &= \mathrm{E}(\mathbf{r}) - \lambda\Sigma\mathbf{w} + A'\boldsymbol{\gamma} = \mathbf{0} \qquad (3.4) \\ \frac{\partial L}{\partial \boldsymbol{\gamma}} &= A\mathbf{w} - \mathbf{b} = \mathbf{0} \qquad (3.5) \end{aligned}$$

From equation (3.4) it follows that $\mathbf{w}^* = (\lambda\Sigma)^{-1}[A'\boldsymbol{\gamma} + \mathrm{E}(\mathbf{r})]$, which can be substituted in equation (3.5) to give an expression for $\boldsymbol{\gamma}$.

$$\begin{aligned} \mathbf{b} &= A\mathbf{w} \\ \mathbf{b} &=^1 A(\lambda\Sigma)^{-1}(A'\boldsymbol{\gamma} + \mathrm{E}(\mathbf{r})) \\ \mathbf{b} &= A(\lambda\Sigma)^{-1}A'\boldsymbol{\gamma} + A(\lambda\Sigma)^{-1}\mathrm{E}(\mathbf{r}) \\ \boldsymbol{\gamma} &= [A(\lambda\Sigma)^{-1}A']^{-1}[\mathbf{b} - A(\lambda\Sigma)^{-1}\mathrm{E}(\mathbf{r})] \end{aligned}$$

$=^1$ *The optimal weight* $\mathbf{w} = (\lambda\Sigma)^{-1}(A'\boldsymbol{\gamma} + \mathrm{E}(\mathbf{r}))$ *is substituted.*

Therefore the solution to the problem including equality constraints is: $\mathbf{w}^* = (\lambda\Sigma)^{-1}[A'\boldsymbol{\gamma} + \mathrm{E}(\mathbf{r})]$, with $\boldsymbol{\gamma} = [A(\lambda\Sigma)^{-1}A']^{-1}[\mathbf{b} - A(\lambda\Sigma)^{-1}\mathrm{E}(\mathbf{r})]$. The solution to the unconstrained problem can be seen in this solution: $\mathbf{w}^* = (\lambda\Sigma)^{-1}[A'\boldsymbol{\gamma} + \mathrm{E}(\mathbf{r})] = (\lambda\Sigma)^{-1}\mathrm{E}(\mathbf{r}) + (\lambda\Sigma)^{-1}A'\boldsymbol{\gamma}$. If $A = 0$ and $b = 0$ the equality constraint solution reduces to the unconstrained solution $\mathbf{w}^* = (\lambda\Sigma)^{-1}\mathrm{E}(\mathbf{r})$.

3.2.3 Inequality constraints

In addition to equality constraints, there could also be other constraints to the problem. Restrictions on borrowing, lending or the amount of borrowing and lending are very common. Such restrictions are generally of the form $\mathbf{l} \leq A\mathbf{w} \leq \mathbf{u}$, where $\mathbf{l} \in \mathbb{R}^k$ is a lower bound and $\mathbf{u} \in \mathbb{R}^k$ is an upper bound.

Alternatively, borrowing of equities can be prohibited: $w_i \geq 0$, which is called a short selling constraint. The mean-variance optimization problem with the addition of inequality constraints can be written as:

Problem 4 (MV with inequality constraints)**.**

$$\max_{\mathbf{w}\in\mathbb{R}^n} \mathbf{w}'\mathrm{E}(\mathbf{r}) - \frac{1}{2}\lambda\mathbf{w}'\Sigma\mathbf{w} \qquad \forall\lambda \geq 0$$
$$\text{subject to } \mathbf{l} \leq A\mathbf{w} \leq \mathbf{u}$$

The inequality constraints greatly complicate the problem. It is no longer possible to find an analytical solution, it has become a parametric quadratic programming problem (parametric due to the parameter λ and quadratic due to the quadratic term in the variance). Several algorithms are available to solve parametric quadratic programming problems. Markowitz (1959) has developed a so called critical line algorithm to solve this problem.

3.2.4 Separation theorem

Previously, the portfolio could only consist of risky assets, assets with an expected return and variance unequal to zero. When there is a risk-free asset available, an asset that has zero variance, it changes the portfolio selection problem, Tobin (1958) investigated this problem and has developed the separation theorem.

The risk-free asset (r_f) is depicted in figure 3.2, as it has zero variance it is placed on the expected return axis. The risk-free asset makes it possible to draw a new efficient frontier that has a better risk-return balance. This is accomplished by forming a portfolio that consists of a combination of the risk-free asset and the tangency portfolio. This line is represented in the figure as the 'new efficient frontier'.

The new portfolio (p_n) is constructed in the following way: $p_n = \lambda r_f + (1 - \lambda)p_t$, where p_t is the tangency portfolio and r_f is the risk-free asset. The parameter λ can be varied to obtain a series of portfolio and eventually draw out the new efficient frontier. If $\lambda = 1$ the new portfolio consist only of the risk-free asset, if $\lambda = 0$ the portfolio consist only of the tangency portfolio. If $\lambda < 0$ the risk-free asset is borrowed to finance a larger position in the tangency portfolio.

The expected return and variance of this new portfolio, can be computed from the mean and variance of the risk-free asset and those of the tangency portfolio. The tangency portfolio has expected return $\mathrm{E}(p_t) = \mu_t$ and variance of return $\mathrm{var}(p_t) = \sigma_t^2$. The risk-free asset has no variance, thus $\mathrm{var}(r_f) = 0$ and expected return $\mathrm{E}(r_f) = r_f$.

The expected return of the new portfolio (r_n) thus equals $\mathrm{E}(r_n) = \mathrm{E}(\lambda r_f + (1 - \lambda)r_t) = \lambda\mathrm{E}(r_f) + (1 - \lambda)\mathrm{E}(r_t) = \lambda r_f + (1 - \lambda)\mu_t$ (properties (2.1) and (2.2) are used). The variance of the new portfolio equals $\mathrm{var}(p_n) = \mathrm{var}(\lambda r_f +$

$(1-\lambda)r_t) = \lambda^2 \mathrm{var}(r_f) + (1-\lambda)^2 \mathrm{var}(r_t) = 0 + (1-\lambda)^2 \sigma_t^2$ (property (2.6) is used). The new portfolio in the variance-expected return diagram can thus be parametrized by $(\mathrm{var}(r_n), \mathrm{E}(r_n)) = ((1-\lambda)^2\sigma_t^2, \lambda r_f + (1-\lambda)\mu_t)$. The coefficients form a tipped parabola, due to the quadratic term in the variance. However, in the volatility-expected return diagram it would become a straight line. Then, the volatility or standard deviation of the new portfolio is given by $\sigma_n = \sqrt{\mathrm{var}(r_n)} = (1-\lambda)\sigma_t$. Therefore, the new portfolio in the volatility-expected return diagram is parametrized by $((1-\lambda)\sigma_t, \lambda r_f + (1-\lambda)\mu_t)$. The series of portfolios is portrayed in figure 3.2 by the new efficient frontier.

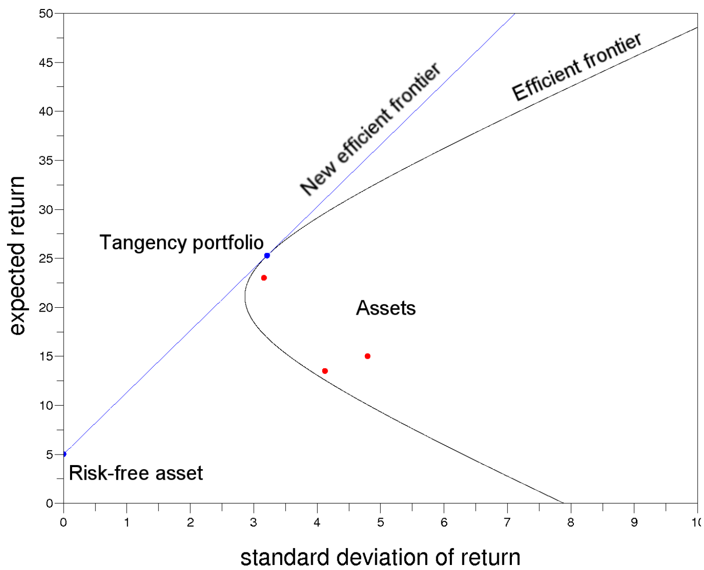


Figure 3.2: The separation theorem: the risk-free asset is used to set the risk level

In the presence of a risk-free asset, the portfolio selection problem becomes a two-part problem. First determine the tangency portfolio and next adjust the tangency portfolio to the desired risk-level by going long or short in the risk-free asset.

The selection of the risky assets in the portfolio can now be separated from the attitude towards risk. The separation theorem is important in the next development in modern portfolio theory, the development of the capital asset pricing model by Sharpe, Lintner, Mossin and Treynor.

3.2.5 Summary

The mean-variance optimization problem can be divided in three versions, each with its own degrees of difficulty. The unconstrained problem is an

ordinary maximization problem. Adding equality constraints makes the problem more realistic, but also increases the difficulty. It is still possible to find an analytical solution, with the help of Lagrange multipliers. The last version of the mean-variance optimization is the most realistic one, and also the most difficult to solve. There no longer exists a closed form solution, but a solution can be found via a quadratic programming algorithm. A strong point of mean-variance optimization is its flexibility, it is very easy to add additional constraints to the problem.
Tobin's separation theorem allows the problem to be divided into two parts if there is a risk-free rate of borrowing and lending available. The fist step is to select the tangency portfolio and the next step is to adjust the portfolio to the desired risk level by borrowing or lending of the risk-free asset.

3.3 Weak points of mean-variance analysis

3.3.1 Utility theory and mean-variance analysis

The mean-variance criterion makes the exchange between risk and expected return explicit. The criterion states a preferences for portfolios with a higher expected return relative to portfolios with a lower level of expected return (for the same level of risk). This seems a reasonable criterion for portfolio selection. However, care has to be taken in applying the criterion, since in some cases the criterion results in unlikely preferences.
This failure has been analyzed by various authors among which Hanoch and Levy (1969). They analyze preferences with the help of utility theory and subsequently compare these preferences with those obtained from mean-variance optimization. They conclude that in certain cases the preferences resulting from mean-variance optimization differ from those obtained by utility theory. Before we move to the explanation of these results we will first discuss an example.

The MV-criterion is not always valid Hanoch and Levy (1969) analyze when the mean-variance criterion captures the preferences of a risk-averse investor correctly. Example 2, adapted from Hanoch and Levy (1969) is an instance in which the MV-criterion results in unlikely preferences.

Example 2. *An investor can choose between two assets X and Y, whose returns are random variables (the notation convention, to depict only matrices in capital letters, is violated for this example). The first asset (X) has a return of 1 with probability 0.8 and a return of 100 with probability 0.2. The second asset (Y) has a return of 10 with probability 0.99 and a return of 1000 with probability 0.01.*

x	$\mathbb{P}(X = x)$	y	$\mathbb{P}(Y = y)$
1	0.8	10	0.99
100	0.2	1000	0.01

$\mathrm{E}(X) = 1 \cdot 0.8 + 100 \cdot 0.2 = 20.8$, $\mathrm{E}(Y) = 10 \cdot 0.99 + 1000 \cdot 0.01 = 19.9$
$\mathrm{E}(X) > \mathrm{E}(Y)$, $\mathrm{var}(X) = 1568 < \mathrm{var}(Y) = 9703$

Table 3.1: The mean and variance of asset X and Y.

Computation of the expected return learns that the expected return of asset X is 20.8, and the expected return of asset Y is equal to 19.9. Hence the expected return of asset X is larger than the expected return of asset Y. The variance of asset X (=1568) is smaller than the variance of asset Y (= 9703). The mean-variance criterion says to prefer asset X in this case, as it has a greater expected return and a smaller variance than asset Y. However, it seems more natural to prefer asset Y: in that case one has almost certainly a return of 10 as opposed to a return of 1.

This example shows that the mean-variance criterion sometimes gives counterintuitive answers. The example will be continued later, when the rudimentaries of utility theory have been developed.

Utility theory[1] Utility theory can be used to rank preferences. Formally, a utility function is a function u defined from a space Z representing the various possible portfolios to the real line ($\mathbb{R}$).

Definition 3 (Utility function)**.** *An utility function, is a function $u\colon Z \to \mathbb{R}$. It is a non-decreasing, continuous function that captures the investors preferences.*

An investor will prefer portfolio P_1 to P_2 if the expected utility of portfolio P_1 is greater than the expected utility of portfolio P_2. The specific utility function used varies among individuals, depending on their individual risk tolerance and their individual financial environment. The simplest utility function is a linear one $u(x) = x$. An investor using this utility function ranks portfolios by their expected values, risk does not play a role. The linear utility function is said to be risk neutral since there is no trade off between risk and expected return in the order of preferences.
A wide range of utility function are allowed, however in practice certain standard types are popular. The most commonly used utility functions are

[1]The description of utility theory has been derived from Luenberger (1998).

exponential $u(x) = -\exp(-ax)$ with $a > 0$, logarithmic $u(x) = \log(x)$, power $u(x) = bx^b$ for $b < 1$ and $b \neq 0$ and the quadratic function $u(x) = x - bx^2$ for $b > 0$.
A logarithmic utility function could be used to rank the portfolios in the previous example.

Example 3 (continued from example 2)**.** *The logarithmic function $u(x) = \log_{10} x$, would rank the assets of example 2 as depicted in table 3.2.*

x	$\mathbb{P}(X = x)$	y	$\mathbb{P}(Y = y)$
1	0.8	10	0.99
100	0.2	1000	0.01

$\mathrm{E}[u(x)] = \mathbb{P}(X = x_1)u(x_1) + \mathbb{P}(X = x_2)u(x_2) = 0.8 \cdot 0 + 0.2 \cdot 2 = 0.4$
$\mathrm{E}[u(y)] = \mathbb{P}(Y = y_1)u(y_1) + \mathbb{P}(Y = y_2)u(y_2) = 0.99 \cdot 1 + 0.01 \cdot 3 = 1.02$
$\mathrm{E}[u(y)] > \mathrm{E}[u(x)]$

Table 3.2: The expected utility of asset X and Y.

The expected utility of X is $\mathrm{E}[u(x)] = 0.4$, this is bigger than the expected utility of Y: $\mathrm{E}[u(y)] = 1.02$. The expected utility of Y is greater than that of X, hence asset Y should be preferred to X, as seems in accordance with intuition.

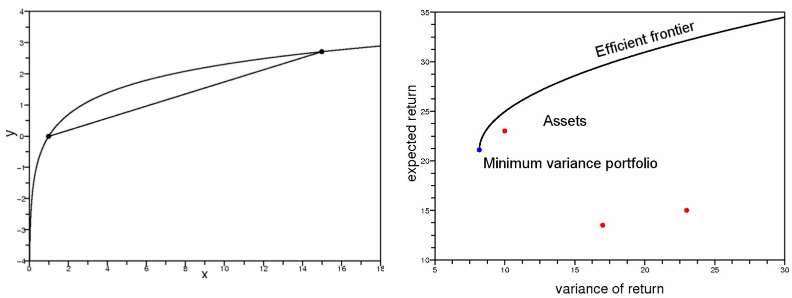
Equivalent utility functions An utility function is used to provide a ranking among alternatives; its actual numerical value (called its cardinal value) has no real meaning. What is important, is how the function ranks alternatives when the expected utility is computed (called its ordinal value). An expected utility function is not unique, there are several functions that provide the same ranking. This non-uniqueness is due to the linear nature of the expected return. The utility function $v(x) = au(x) + b$, provides the same ranking as the utility function $u(x)$. This can be seen by taking the expected value of the utility function. The expected value of $v(x)$ is $\mathrm{E}[v(x)] = \mathrm{E}[au(x) + b] = a\mathrm{E}[u(x)] + b$. Adding a constant to each value does not change the ranking of the values nor does multiplication by a factor. Hence, the rankings of the expected utility function of $u(x)$ and $v(x)$ are the same. When utility functions produce the same ranking, they are called equivalent, $u(x)$ and $v(x)$ are in this case equivalent. The equivalence of utility functions can be used to scale utility functions conveniently.

Risk aversion The main purpose of an utility function is to provide the investor with a systematic way of ranking alternatives. The utility function and hence the ranking of alternatives can be used to capture the principle of risk aversion. Risk aversion is represented in utility terms by a concave utility function.

Definition 4 (Concave utility and risk aversion). *A function u defined on an interval $[a, b]$ of the real numbers is said to be concave if for any α with $0 \leq \alpha \leq 1$ and any x and y in $[a, b]$ it holds that*

$$u(\alpha x + (1-\alpha)y) \geq \alpha u(x) + (1-\alpha)u(y). \tag{3.6}$$

An utility function $u(x)$ is said to be risk averse on $[a, b]$ if it is concave on $[a, b]$. If $u(x)$ is concave everywhere, it is said to be risk averse.



(a) The function $\log(x)$ is concave. (b) The efficient frontier is concave.

Figure 3.3: Concave functions and risk-averseness.

The property that a function is concave can be formulated in several ways. A general condition for concavity is that the straight line drawn between any two points on the graph of the function must lie below (or on) the graph itself. In simple terms, an increasing concave function has a slope that flattens for increasing values. In mathematical terms, this is equivalent to the second derivative of the function being negative on the whole domain, $u''(x) < 0$ for all x.
The efficient frontier is an example of a concave function, in figure 3.3b it can be seen that any line drawn between two points on the efficient frontier lies below the graph. Furthermore, the efficient frontier describes the preferences of a risk-averse investor: a portfolio on the efficient frontier that has a higher risk, also has a higher expected return.
A special case is the risk-neutral utility function $u(x) = x$. This function is concave according to the preceding definition, but it is a degenerate case. Strictly speaking, this function represents no risk aversion. Normally the term risk averse is reserved for the case where $u(x)$ is strictly concave, which means that there is strict inequality in equation (3.6) whenever $x \neq y$.

Example 4 (continued from example 2). *The logarithmic function is an example of an concave utility function. This can be seen by computing the second derivative:*

$$u(x) = \log_{10} x \;\Rightarrow\; u'(x) = \frac{1}{\log(10)x}$$
$$u''(x) = \frac{-1}{\log(10)x^2} < 0 \textit{ for all } x \in \mathbb{R}$$

The $\log(x)$ *function is a strictly concave utility function values of* x *on* $[0, \infty)$*, hence it can represent the preferences of a risk-averse investor. It has already been shown that the expected utility of asset* X*, with this utility function, is smaller than the expected utility of asset* Y*. Therefore a risk-averse investor should prefer asset* Y *to asset* X*. The mean-variance criterion, which should represent the preferences of a risk-averse investor, says to prefer asset* X*. Where does the mean-variance criterion go wrong?*

Risk aversion coefficients The degree of risk aversion exhibited by a utility function is related to the magnitude of the curvature in the function. The stronger the curvature, the greater the risk aversion. This notion can be quantified in terms of the second derivative of the utility function.
The degree of risk aversion is formally defined by the Arrow-Pratt absolute risk aversion coefficient, which is

$$a(x) = -\frac{u''(x)}{u'(x)}$$

The term $u'(x)$ is used in the denominator to normalize the coefficient. This normalization causes $a(x)$ to be comparable for all equivalent utility functions. The coefficient function $a(x)$ expresses how risk aversion changes with the wealth level. For many investors, risk aversion decreases as their wealth increases, reflecting that they are willing to take more risk when they are financially secure.

3.3.2 The MV-criterion implies normally distributed returns

Hanoch and Levy (1969) have studied the question when the mean-variance criterion is a valid efficiency criterion for a risk averse investor. An efficiency criterion is said to be valid if it produces the same efficient set for all concave utility functions. The ranking of the elements in the efficient set still depends on the specific utility function. They found, as Tobin (1958) already suspected, that the mean-variance criterion is valid if and only if the distribution of the returns is of a two parameter family. The proof is omitted as it would carry to far for this thesis.
They concluded that the mean-variance "criterion is optimal, when the distributions considered are all Gaussian normal. But the symmetric nature

of this distribution seems to deny its usefulness as a good approximation to reality, for at least some types of risky portfolios.
Even for symmetric distributions, the mean-variance criterion is not valid, when the distribution has more than two parameters."

3.3.3 Shortcomings of mean-variance optimization

Thus far, the theoretical background of mean-variance optimization has been described, in which setting it makes good sense. However, when applying it to real live problems some flaws do arise. The main problems are that the optimization procedure often results in concentrated portfolios, that the model requires much input data and, finally, that the model lacks robustness.

Concentrated portfolios In general diversification is thought of as a reasonable approach to spreading risk. Adding assets to a portfolio that are less than perfectly correlated to the assets already in the portfolio reduces the variance of the portfolio. Mean-variance optimization however, can result in portfolios with large long and short positions in only a few assets (Black and Litterman, 1992), which opposes the notion of diversification. If the parameters that are used in the optimization, like the vector of expected return and the covariance matrix, would be known with certainty, it would be reasonable to invest in such concentrated portfolios, but as the expected returns are just forecasts this seems a very risky investment choice. Concentrated portfolios are very counterintuitive, which is one of the reasons for the lack of popularity of using unconstrained mean-variance optimizers in making investment decisions (Michaud, 1989).

The model requires much input data A more practical problem is that the model requires input of expected return, variance and covariance of every asset under consideration.
If an investor has 1000 assets in her portfolio, it becomes a very cumbersome task to give estimates for all the input parameters. There are solutions to this problem: historical data could be used to give an estimate of the expected return. But historical estimates are often bad predictors of future behavior (Black and Litterman, 1992).
Another problem is how the investor should formulate her believes about future performance. Often an investor holds relative views on asset performance, for example that asset a will outperform asset b. Mean-variance analysis needs a specific estimate of the expected return of a single asset and cannot handle relative views.

The model is not robust The main deficiency is that the model is not robust. This means that a small change in the values of the input parameters can cause a large change in the composition of the portfolio. The

mean-variance model assumes that the input data is correct, without any estimation error. The model does not address this uncertainty and sets out to optimize the parameters as if they were certain. Michaud (1989) describes mean-variance optimizers even as "estimation-error maximizers".
Best and Grauer (1991) have analyzed the behavior of the mean-variance optimizer under changes in the asset mean. They show that a small increase in an asset mean can cause a very different portfolio composition. Half of the assets can be eliminated from the original portfolio, by a small increase in the mean of one asset. They state that for a mean-variance problem with a budget constraint, the rate of change for the vector of weights depends on the change in the mean. More specifically it depends on λ, the risk-aversion parameter, Σ^{-1}, the inverted variance-covariance matrix and q, which specifies the change in the asset mean.
Michaud (1989) considers the covariance matrix as the main culprit for the non-robust behavior. The covariance matrix is often estimated from data, but this estimation procedure can produce matrices that are nearly singular or singular. This causes problems when the matrix is inverted during the optimization procedure.

Summary The idea of mean-variance optimization has an intuitive appeal and is very useful for educational purposes. When the optimization procedure is used for practical purposes then the resulting portfolios are counter-intuitive and the optimization procedure should be constrained. Furthermore, the model is non-robust and the assumption of normally distributed expected returns is not always a good assumption.
Mean-variance analysis has been a good starting point for the development of portfolio selection, but it could be improved on.

3.4 Capital asset pricing model[2]

The work of Markowitz on portfolio selection became relevant with the publication of the *capital asset pricing model (CAPM)*. This theory, that has been separately developed by William Sharpe (1964), John Lintner (1965), Jan Mossin (1966) and Jack Treynor (1961), builds on the mean-variance analysis of Markowitz to develop a model that can compute the expected return of an asset if an equilibrium would exist in the market.

3.4.1 Assumptions

The model is based on the following assumptions. Suppose that every investor bases his investing decisions on mean-variance theory. Suppose also that the investors agree on the future performance of every asset in the

[2]The text has been derived from Luenberger (1998) and Sharpe (1970).

investment universe, hence everyone assigns the same mean, variance and covariance to the assets. Assume there is a unique risk-free rate of borrowing or lending available to all investors and that there are no transaction costs. Under these assumptions an equilibrium can be established in the market. From this equilibrium the pricing formula can be derived.

Assumption. *A1 All investors use mean-variance analysis to select a portfolio.*

A2 All investors have homogeneous believes about the future return, variance and covariance of assets.

A3 There is a unique risk-free rate of borrowing and lending available for all investors.

3.4.2 Equilibrium

From Tobin's separation theory (see section 3.2.4) it is known that everyone will invest in a single portfolio of risky assets. In addition, investors can borrow or lend at the risk-free rate, to adjust the portfolio to the desired risk level. Furthermore, since everyone uses the same means, variances, and covariances, to determine the optimal portfolio, everyone will compile the same risky portfolio.
Some investors will seek to avoid risk and will have a high percentage of the risk-free asset in their portfolios. Other investors who are more aggressive, will have a high percentage of the risky portfolio. However, every individual will form a portfolio that is a mix of the risk-free asset and the same risky portfolio.
If everyone purchases the same portfolio of risky assets, what must that portfolio be? The answer to this question is the key insight underlying the CAPM.
As all investors share the same view, and at that moment there is only one optimal portfolio, it will result in a rising price of the assets in the optimal portfolio and hence a downward adjustment of the expected return. The opposite happens with the assets not in the portfolio. These price changes lead to a revision of the portfolios. And this goes on and on, until an equilibrium is reached. In this equilibrium the optimal portfolio is the one that contains all assets proportional to their capitalization weights, that is the market portfolio. This means that the market portfolio is mean-variance efficient in equilibrium.
This theory of equilibrium is usually applied to assets that are traded repeatedly over time, such as equities traded on a stock exchange. The conclusion of the mean-variance approach and Tobin's separation theorem is that the optimal portfolio, in which everyone invests, must be the market portfolio($\mathbf{w}_m$).

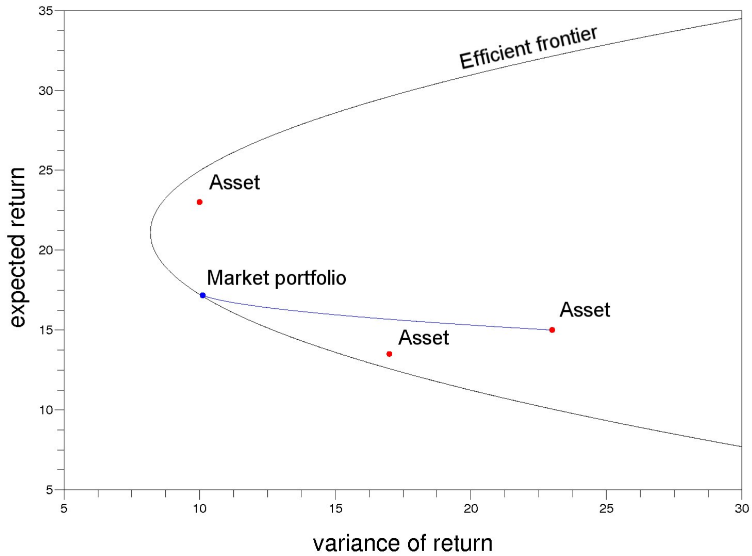


Figure 3.4: The curve that connects the market portfolio and a single asset.

3.4.3 The pricing formula

Theorem 3.4.1 (The capital asset pricing model)**.** *If the market portfolio M is mean-variance efficient, the expected return of an asset i satisfies*

$$\mathrm{E}(r_i) - r_f = \frac{\sigma_{iM}}{\sigma_M^2}\left(\mathrm{E}(r_M) - r_f\right) = \beta_i \left(\mathrm{E}(r_M) - r_f\right) \tag{3.7}$$

where $\beta_i = \dfrac{\sigma_{iM}}{\sigma_M^2}$

σ_{iM} *represents the covariance between asset* i

and the market portfolio M

σ_M^2 *represents the variance of the market portfolio*

Proof. The proof rests on connecting the market portfolio, with a single asset. This forms a curve that should be tangent to the capital market line, i.e. the line connecting the market portfolio with the risk-free asset, in the point M. From this equality it is possible to find the pricing formula.
We start by forming the curve that connects the market portfolio and a single asset, see figure 3.4. For any $\alpha \in \mathbb{R}$ consider the portfolio consisting of a portion α invested in asset i and a portion $1-\alpha$ invested in the market portfolio M. The return on this portfolio would be $r_\alpha = \alpha r_i + (1-\alpha) r_M$. The portion α can be negative, $\alpha < 0$, which corresponds to borrowing at the risk-free rate. The expected rate of return of this portfolio is formed

by the expected return of the asset and the expected return of the market portfolio

$$\mathrm{E}(r_\alpha) = \alpha \mathrm{E}(r_i) + (1-\alpha)\mathrm{E}(r_M). \tag{3.8}$$

The variance of the return is

$$\sigma_\alpha^2 = \alpha^2 \sigma_i^2 + 2\alpha(1-\alpha)\sigma_{iM} + (1-\alpha)^2 \sigma_M^2. \tag{3.9}$$

As α varies, the values of $(\mathrm{E}(r_\alpha), \sigma_\alpha^2)$ trace out a curve in the risk-return diagram, as shown in figure 3.4. In particular, $\alpha = 0$ corresponds to the market portfolio M. The curve cannot cross the capital market line, as this is the efficient frontier. Hence as α passes through zero, the curve must be tangent to the capital market line at M. The tangency condition will be used to derive the pricing formula.
When the two lines are tangent at M, the two lines must have the same slope in that point. To get at this equality, we need to calculate the derivative of the connecting curve. The curve is parametrized by the function for expected return (3.8) and the function for the standard deviation (3.9). To compute a derivative for $(\mathrm{E}(r), \sigma)$, first the derivatives are computed separately and later they are combined to form the desired derivative.

$$\frac{\mathrm{dE}(r_\alpha)}{\mathrm{d}\alpha} = \mathrm{E}(r_i) - \mathrm{E}(r_M)$$
$$\frac{\mathrm{d}\sigma_\alpha^2}{\mathrm{d}\alpha} = 2\alpha\sigma_i^2 + 2(1-\alpha)\sigma_{iM} - 2\alpha\sigma_{iM} - 2(1-\alpha)\sigma_M^2)$$

Thus at the market portfolio $\alpha = 0$, this gives

$$\left.\frac{\mathrm{d}\sigma_\alpha^2}{\mathrm{d}\alpha}\right|_{\alpha=0} = \sigma_{iM} - \sigma_M^2$$

We then use the following relation to compute the derivative of the curve,

$$\frac{\mathrm{dE}(r_\alpha)}{\mathrm{d}\sigma_\alpha^2} = \frac{\mathrm{dE}(r_\alpha)}{\mathrm{d}\alpha}\frac{\mathrm{d}\alpha}{\mathrm{d}\sigma_\alpha^2} = \frac{\mathrm{dE}(r_\alpha)}{\mathrm{d}\alpha}\left(\frac{\mathrm{d}\sigma_\alpha^2}{\mathrm{d}\alpha}\right)^{-1} \quad \text{if} \quad \frac{\mathrm{d}\sigma_\alpha^2}{\mathrm{d}\alpha} \neq 0$$

to obtain

$$\left.\frac{\mathrm{dE}(r_\alpha)}{\mathrm{d}\sigma_\alpha^2}\right|_{\alpha=0} = \frac{\mathrm{E}(r_i) - \mathrm{E}(r_M)}{\sigma_{iM} - \sigma_M^2}.$$

This slope must be equal to the slope of the capital market line. Hence,

$$\frac{\mathrm{E}(r_i) - \mathrm{E}(r_M)}{\sigma_{iM} - \sigma_M^2} = \frac{\mathrm{E}(r_M) - r_f}{\sigma_M^2}$$

Solve for $\mathrm{E}(r_i)$, to obtain the final result

$$\begin{aligned}\mathrm{E}(r_i) &= \frac{\mathrm{E}(r_M) - r_f}{\sigma_M^2}(\sigma_{iM} - \sigma_M^2) + \mathrm{E}(r_M) \\ &= (\mathrm{E}(r_M) - r_f)\frac{\sigma_{iM}}{\sigma_M^2} + r_f \\ &= (\mathrm{E}(r_M) - r_f)\beta_i + r_f\end{aligned}$$

This is equivalent to formula (3.7) in the capital asset pricing model. □

Two kinds of risk An interesting result of the CAPM is that it allows risk to be divided in two parts. To develop this result the return (r_i) of asset i is written as

$$r_i = r_f + \beta_i(\mathrm{E}(r_M) - r_f) + \epsilon_i. \tag{3.10}$$

Where ϵ_i is a random variable to indicate the uncertainty in the return. The CAPM formula can be used to derive two results about ϵ_i.
From the formula for the expected value of r_i (3.8) the first result follows: the expected value of ϵ_i must be zero. The second result follows by taking the correlation of the return of an asset (3.10) with the return of the market portfolio r_M: from this it follows that the covariance of ϵ_i with the market portfolio is zero, $\mathrm{cov}(\epsilon_i, \sigma_M) = 0$. Therefore the variance of an asset is:

$$\sigma_i^2 = \beta_i^2\sigma_M^2 + \mathrm{var}(\epsilon_i) \tag{3.11}$$

The first part $\beta_i^2\sigma_i^2$ is called *systematic risk*. This is the risk associated with the market as a whole. This risk cannot be reduced by diversification because every asset with nonzero beta contains this risk. The second part, $\mathrm{var}(\epsilon_i)$, is termed the *non-systematic*, idiosyncratic, or specific risk. This risk is uncorrelated with the market and can be reduced by diversification. It is the systematic (or non-diversifiable) risk, measured by beta, that is most important, since it directly combines with the systematic risk of other assets. A result of CAPM is that expected return depends on this beta.

3.5 Summary

In this chapter, the mean-variance approach has been described in detail. The most complex instance is the one with inequality constraints. This instance can be solved as a parametric quadratic programming problem.
There are several limitations to the mean-variance criterion. The main limitations are that it can result in counterintuitive portfolios, the method is non-robust and it assumes normally distributed returns.
Finally the capital asset pricing model has been derived. The result of this model is that in an equilibrium situation every investor should hold the

market portfolio combined with the risk-free asset. Furthermore it concludes that the expected return of an asset is proportional to the market beta. The CAPM will play a role in the next chapter, where the Black-Litterman model will be discussed.

Chapter 4

The Black-Litterman model

The mean-variance model was a groundbreaking model in portfolio selection. It allows investors to quantitatively select a portfolio on the basis of their views. But, as explained in the previous chapter, the model has its shortcomings. Practitioners do use this model, but worked around the problems mentioned by adding many constraints.
Fischer Black and Robert Litterman, then working at Goldman Sachs, set out to improve the original mean-variance model. They choose a practitioners perspective and wanted to develop a model that could be used at Goldman Sachs for portfolio selection. Therefore the mathematics of the model should be tractable, the inputs should be intuitive to the investment manager and the optimized portfolio should reflect the investors views. They have published a few articles on their model (1991a, 1991b, 1992), but none of them are very precise about the mathematics of the model.
Various authors have tried to shed light on the model. Satchell and Scowcroft tried to demystify the mathematics of the model in their 2000 article titled "A demystification of the Black-Litterman model: Managing quantitative and traditional portfolio construction". Also there is a lucid description of the model in chapter seven of Lee's book (2000) on tactical asset allocation. The bank UBS uses a model very similar to the Black-Litterman model, which is described by Scowcroft and Sefton in chapter 4 of the book by Satchell and Scowcroft (2003). Also Idzorek (2004), from Ibbotson associates, wrote an interesting guide about the Black-Litterman model: "A step-by-step guide to the Black-Litterman model". His paper gives an overview of the articles about the model and describes a new method to set the level of uncertainty in the model.
In this chapter the Black-Litterman model (BL-model) is explained, with emphasize on the mathematics of the model. In consecutive order, the objectives of the model, the mathematical derivation and a further explanation of the parameters in the model will be discussed.

4.1 The Black-Litterman model in general

Black and Litterman set out to accomplish more intuitive portfolios by computing a better estimate for the expected return vector. This expected return vector could then directly be used to compute the portfolio weights, or the expected return vector could be fed to a mean-variance optimizer to provide a solution to a constrained optimization problem.
Black and Litterman identified two sources of information about the expected returns and they combined these two sources of information in one expected return formula. The first source of information is obtained quantitatively, these are the expected returns that follow from the CAPM and thus should hold if the market is in an equilibrium. The CAPM returns form a backbone to the process, and are used to dampen the possibly extreme views of the second source of information.
The second source of information are the views held by the investment manager. The investment manager has access to different information and could therefore have different opinions about the expected return of the asset, than those that would hold in an equilibrium. The views of the investor are used to tilt the equilibrium views, they provide information to invest less or more in a certain asset, then would follow from the equilibrium views.
Combining these two sources of information results in a new vector of expected returns. This improved vector of expected returns can then be used in the portfolio optimization process.

4.1.1 Two sources of information

Quantitative views can provide a stable reference point. This reference point is often called a *benchmark portfolio*, neutral view or *equilibrium view*.

Quantitative views on expected return Black and Litterman look at investments in a global context and invest in three asset classes: equity, bonds and currency. Therefore, the quantitative views should reflect this global framework and they choose an international CAPM to compute the equilibrium returns. The standard CAPM of Sharpe (see section 3.4) is used to compute the expected return of equity and bonds, which is supplemented with the universal hedge ratio of Black (1989) to obtain the ratio of currency that should be hedged.
The CAPM computes the expected return of an asset when there is an equilibrium in the market. The major drawback of the formula, $\mathrm{E}(r_i) - r_f = \frac{\sigma_{iM}}{\sigma_M^2}(\mathrm{E}(r_M) - r_f)$, is that it is difficult to specify the expected return of the market portfolio $\mathrm{E}(r_M)$, as the world market is practically unbounded. Therefore, a benchmark portfolio is often used as a proxy for the market. For such a finite benchmark it is possible to compute the expected return and therefore to estimate the expected return of the single asset.

A benchmark portfolio could be an index, for example the Standard & Poor's Global 1200 index (S&P Global 1200). The S&P Global 1200 covers equity in 29 countries and approximately 70 percent of global market capitalization[1]. From the returns of the assets in the benchmark it is possible to estimate the expected return of the market portfolio. The topic of equilibrium returns will be discussed more thoroughly in section 4.2.2.

The investor's view on expected returns An investor also has views on the expected return of assets. The BL-model allows investors to express their views in an absolute sense, asset A will have an expected return of x, as well as in a relative sense, e.g. asset A will outperform asset B. The Black-Litterman model allows the investor to represent views in such a relative way, which is much closer to how investors think (Scowcroft and Sefton, 2003).
Additionally, the investor might not be equally certain about every view. This should also be reflected in the model, so that not all views are treated on an equal footing. Therefore, Black and Litterman made it possible to express a level of certainty to each view separately.
The number of views that the investor wants to take into account is flexible. It can range from no views at all to as much views as there are assets under consideration. This makes the model much better to use. Investors often focus only on a small part of the potential investment universe, choosing assets that they feel are undervalued, finding assets with positive momentum, or identifying relative value trades. In the Black-Litterman model, it is only necessary to specify a view if the investor holds one.

Integrate quantitative and traditional approach An additional advantage of this approach is that is combines two approaches that were formerly considered separately. Scowcroft and Sefton, in chapter 4 of "Advances in portfolio construction and implementation (2003)" address this broader issue in investment management: the dichotomy between the traditional investment manager and the more quantitative approach to investments.
The traditional manager forms views from news about performance of companies, markets, interest rates, etc. A view could be that pharmaceutical companies will outperform food companies. A view could also be that the French companies will outperform German companies. Or that Unilever will outperform Kraft Foods. From these views she compiles a portfolio, without the help or with limited help of a quantitative model. Typically, the traditional investment manager has little background in mathematics, and is hesitant to use quantitative models, as she feels that techniques of mean-variance analysis and related procedures do not capture effectively

[1] http://www2.standardandpoors.com/spf/pdf/index/factsheet_global1200.pdf

their value added.
The quantitative manager, on the other hand, is surprised by the lack of rigor that the traditional manager uses to select a portfolio. The Black-Litterman model can integrate these diverse approaches. This framework allows the traditional managers to give their views/forecasts, and these views are combined with the quantitative model to give final forecasts that reflect a blend of both viewpoints.

4.2 Mathematics of the Black-Litterman model

Mathematically, the main challenge is to combine the two separate sources of information into one vector of expected returns. This has to be done in such a way that the mathematics remains tractable and the parameters are intuitive to the user.
One could combine the expected returns of the neutral reference points with the views of the investor heuristically. If one is positive about an asset, then simply increase the weight of the asset, and vice versa for assets about which one has a negative outlook. The question then becomes how much to increase it. Furthermore, assets are correlated: if one asset is expected to do well and therefore the weight is increased, then the weights of other positively correlated asset should also be increased. It would be very cumbersome to do this all by hand. A more constructive approach is needed. Black and Litterman have combined these two separate sources of information in a constructive manner and suggest two methods to accomplish this. First, the mixed estimation method of Theil (1971), which is related to the generalized least square method to estimate dependent parameters. Secondly, they suggest that the new vector of expected returns should be 'assumed to have a probability distribution that is a product of two normal distributions' (Black and Litterman, 1991a). Satchell and Scowcroft (2000) propose a Bayesian approach to accomplish this blending of probability distributions.
Not only the method to compute the combined vector of expected returns is poorly described mathematically, but also the characteristics of the variables are unclear. It is unclear what the parameters represent and how they should be specified. This makes it very difficult to use to model.
Of the two approaches suggested by Black and Litterman the most widely used approach is the Bayesian one. Especially after the publication of an article by Stephen Satchell and Alan Scowcroft (2000) about the derivation of the BL-formula. In this article they use a Bayesian approach to combine equilibrium views with the views of the investor. The Bayesian approach updates currently held opinions with the neutral reference point to form new opinions. The reasoning of Satchell and Scowcroft (2000) will be largely followed in the explanation of the Bayesian approach.

After the Bayesian approach, the mixed estimation method of Theil (1971) will be described. The derivation of the main formula is very concise when using this method.

4.2.1 Preliminaries

Satchell and Scowcroft (2000) went a long way in describing the Black-Litterman model mathematically. Their notation and description of basic terms will be used in the thesis.

As before, there are n assets in the universe. The portfolio of assets is represented by a vector of weights $\mathbf{w} \in \mathbb{R}^n$. The return of the assets is a random variable that can be represented by the vector $\mathbf{r} \in \mathbb{R}^n$, and it has an expected value $\mathrm{E}(\mathbf{r}) \in \mathbb{R}^n$.

Black and Litterman adapt the definition of expected return to represent *expected excess returns* by which they mean the expected returns in the domestic currency minus the domestic risk-free rate $\mathrm{E}(\mathbf{r}) - r_f$.

Definition 5 (Expected excess return). *The expected excess return of an asset is the expected return in the domestic currency minus the domestic risk free rate,* $\mathrm{E}(\mathbf{r}) - r_f$.

For conciseness we will use the term expected return to represent expected excess returns in this chapter. In some cases the term expected excess return will still be used for emphasize.

Furthermore, we assume that the variance of return exists and is well defined. This means that the vector of returns $\mathbf{r}$ is assumed to have a positive definite covariance matrix $\Sigma \in \mathbb{R}^{n \times n}$.

Finally, Satchell and Scowcroft (2000) assume that returns are normally distributed random variable. This is a common assumption in finance. However, this assumption is also often shown to be flawed, see for example Embrechts et al. (2003).

To sum up there are two assumptions:

Assumption. *A1 Returns have a normal probability distribution.*

A2 The covariance matrix of returns, Σ, is positive definite.

4.2.2 Equilibrium

The market equilibrium returns form the backbone of the portfolio. There are roughly three approaches in the literature for computing these equilibrium returns. One could use historical means or average returns, the equilibrium returns as defined by the CAPM, or reversed optimized returns from some benchmark portfolio.

In the case of historical means, one estimates the future returns by making

an average of the returns over a certain time interval. For example to estimate the future expected returns an average is taken over the returns of the past six months.
Many authors have pointed out, for example Frankfurter et al. (1971), that past returns are a bad estimator of future expected returns. Frankfurter et al. performed an experiment in which it was assumed that the returns were normally distributed. An estimate of future returns was made by sampling these normally distributed returns over four different run lengths. The obtained estimates are used in a mean-variance optimizer to compile portfolios. The experiment shows that "the make up of the efficient portfolio varies substantially among sample trials, even though generation of all sample data was based upon the assumed parameter values."
Black and Litterman state that the equilibrium returns ($\boldsymbol{\pi}$) can be computed from $\boldsymbol{\pi} = \delta\Sigma\mathbf{w}_m$. The parameter δ is called the (world) risk aversion coefficient and $\mathbf{w}_m$ is the world market portfolio. They claim that the equilibrium returns are derived from the CAPM. However their notation differs from that of the standard representation of the CAPM formula. A transformation of the original CAPM shows however, that their depiction is correct. Satchell and Scowcroft (2000) transform the CAPM formula into the formula for $\boldsymbol{\pi}$. This transformation also results in new insights in the parameter δ.

Proposition 2. *The CAPM formula* $\mathrm{E}(r_i) - r_f = \beta(\mathrm{E}(r_m) - r_f)$ *can be written as* $\boldsymbol{\pi} = \delta\Sigma\mathbf{w}_m$*, where* $\beta = \frac{\mathrm{cov}(r_i, \mathbf{r}'\mathbf{w}_m)}{\sigma_m^2}$*,* $\boldsymbol{\pi} = \mathrm{E}(\mathbf{r}) - r_f$ *and the parameter* δ *is* $\delta = \frac{\mathrm{E}(r_m) - r_f}{\sigma_m^2}$*.*

Proof. The CAPM formula will first be transformed to vector notation in order to have both formulas in the same dimensions. Next we will expand the CAPM vector formula until we are back to basic variables, rearrange the variables and finally combine them into the new variables δ, Σ and $\mathbf{w}_m$.

$$\begin{aligned}
\mathrm{E}(r_i) - r_f &= \beta(\mathrm{E}(r_m) - r_f) \\
\mathrm{E}(\mathbf{r}) - r_f &=^1 \boldsymbol{\beta}(\mathrm{E}(r_m) - r_f) \\
\boldsymbol{\pi} &=^2 \boldsymbol{\beta}\mu_m =^3 \frac{\mathrm{cov}(\mathbf{r}, \mathbf{r}'\mathbf{w}_m)}{\sigma_m^2}\,\mu_m \\
&=^4 \frac{\mathrm{cov}(\mathbf{r}, \mathbf{r}')\mathbf{w}_m}{\sigma_m^2}\,\mu_m =^5 \frac{\mathrm{E}(r_m) - r_f}{\sigma_m^2}\,\mathrm{cov}(\mathbf{r}, \mathbf{r}')\mathbf{w}_m \\
&=^6 \delta\Sigma\mathbf{w}_m
\end{aligned}$$

$=^1$ *The equation is transformed from a one dimensional formula to a vector formula.*
$=^2$ $\boldsymbol{\pi} = \mathrm{E}(\mathbf{r}) - r_f$ *and* $\mu_m = \mathrm{E}(r_m) - r_f$.
$=^3$ $\boldsymbol{\beta} = \frac{\mathrm{cov}(\mathbf{r},\mathbf{r}'\mathbf{w}_m)}{\sigma_m^2}$ *see equation* (3.7).
$=^4$ $\mathrm{cov}(aX, Y) = a\,\mathrm{cov}(X, Y)$ *if* a *is not a random variable, see rule* (2.5).
$=^5$ *Some rearranging of the terms takes place .*
$=^6$ $\delta = \frac{\mathrm{E}(r_m)-r_f}{\sigma_m^2} = \frac{\mu_m}{\sigma_m^2}$ *and* $\Sigma = \mathrm{cov}(\mathbf{r}, \mathbf{r}')$.

Thus the CAPM formula can be written as $\boldsymbol{\pi} = \delta\Sigma\mathbf{w}_m$, where $\delta = \frac{\mathrm{E}(r_m)-r_f}{\sigma_m^2}$ the excess return on the market portfolio divided by its variance. □

From the CAPM it follows that if the market is in an equilibrium, then every investor should hold the market portfolio ($\mathbf{w}_m$). It is possible to compute an estimate for the returns if the weights of this market portfolio are known, however herein lies the difficulty of the CAPM. The weights of the world market portfolio are very difficult to obtain. Therefore the third approach of reverse optimization, advocated by Scowcroft and Sefton (2003), is the most practical one.
A benchmark or index portfolio is used as a proxy for the market weights, taking away the problem of estimating these weights. The equilibrium returns, as in the previous case, follow from the CAPM formula $\boldsymbol{\pi} = \delta\Sigma\mathbf{w}_m$. The parameter δ is in this case the the expected return of the benchmark portfolio divided by its variance. The relation between the market portfolio and the expected returns can also be seen from the unconstrained mean-variance optimization problem. The optimization problem with mean $\boldsymbol{\pi}$ and variance Σ is: $\max_{\mathbf{w}\in\mathbb{R}^n} \mathbf{w}_m'\boldsymbol{\pi} - \frac{\delta}{2}\mathbf{w}_m'\Sigma\mathbf{w}_m$. The solution to this problem is $\mathbf{w}_m = (\delta\Sigma)^{-1}\boldsymbol{\pi}$. In this case the weights are already known and we are interested in the vector of expected returns ($\boldsymbol{\pi}$), thus for a mean-variance efficient portfolio $\mathbf{w}_m$ reverse optimization entails

$$\boldsymbol{\pi} = \delta\Sigma\mathbf{w}_m. \tag{4.1}$$

In this context δ could be chosen equal to $\sqrt{\frac{1}{2c}\mathrm{E}(\mathbf{r})'\Sigma^{-1}\mathrm{E}(\mathbf{r})}$, where c is the desired level of risk as measured in variance, see problem 1 on page 16.
In practice, the values of the risk aversion coefficient vary around 3.
The rationale for taking a benchmark portfolio instead of a world market portfolio also stems from more practical issues. The performance of a investment manager is often measured against a benchmark. The goal could, for example, be to earn a better return than the S&P Global 1200. In that case it would be wise to invest, in the absence of any views, in the benchmark portfolio. When the investment manager does have views, she could deviate from the benchmark portfolio in other to tilt the portfolio in the direction she seems fit. Therefore, it is a practical and safe choice to take the benchmark portfolio as an equilibrium portfolio.

4.2.3 Expressing Views

In many articles on the Black-Litterman model, for example Idzorek (2004), it is emphasized that the investor can express relative views, for example that asset A will outperform asset B by 2%. This manner of expressing is an important improvement of the BL-model over traditional mean-variance optimization, as this manner of expressing views is more intuitive than expressing absolute views.
Let us move to the mathematical description of the manner of expressing views and the view matrix P. An investor often holds views about performance of assets, asset classes or markets. The mathematical representation of these views needs to meet a few characteristics. The views have to be specified relative to the vector of expected return $\mathrm{E}(\mathbf{r})$, the views have to be specified relative to each other and it has to be possible to express a level certainty in the view. These prerequisites lead to the following specification.

$$
\begin{aligned}
P\mathrm{E}(\mathbf{r}) = \mathbf{q} + \boldsymbol{\epsilon},\ &\text{where } \boldsymbol{\epsilon} \sim \mathrm{N}(0, \Omega) \\
&P \in \mathbb{R}^{k \times n} \text{ is known, } \mathbf{q} \in \mathbb{R}^{k} \text{ is known} \\
&\boldsymbol{\epsilon} \in \mathbb{R}^{k} \text{ is an error vector with known variance } \Omega \in \mathbb{R}^{k \times k} \\
&\mathrm{E}(\mathbf{r}) \in \mathbb{R}^{n} \text{ is unknown and needs to be estimated}
\end{aligned}
\tag{4.2}
$$

Assets that are under consideration can be specified in the matrix P, the vector $\mathbf{q}$ expresses the relative change in performance and the vector of random variables $\boldsymbol{\epsilon}$ expresses the uncertainty of the view. The vector $\boldsymbol{\epsilon}$ is normally distributed with mean zero and variance Ω. That the mean is zero means that the investor does not have a standard bias against a certain set of assets. It is assumed that the views are mutually uncorrelated and therefore the covariance matrix Ω is diagonal. A variance of zero represents absolute certainty about the view. The vector $\mathrm{E}(\mathbf{r})$ is the unknown expected return vector that needs to be estimated.
What is often not noted is that Black and Litterman let the manner of formulating views in the matrix P completely free, they not did give any characteristics. A more general idea therefore, that can be found in some literature (see Scowcroft and Sefton (2003)) is to express views on a portfolio of assets. Then the matrix P is considered as a series of portfolios and the vector $\mathbf{q}$ holds the expected return of the corresponding portfolio. It is difficult for a person to estimate the expected return of a portfolio of assets. However, this more general definition does capture all manners of expressing views.
A portfolio could exist of one asset, which would correspond to expressing an absolute view on an asset; a portfolio could be zero-investment, this would correspond to expressing a relative view, and finally one has the possibility to express views on more than two assets.
It is important to note that the vector $\mathbf{q}$ denotes the forecasted relative

performance of the assets. This can be illustrated by an example. An investor holds one view about the relative performance of the assets A and B. She specifies the view in the formula $P\mathrm{E}(\mathbf{r}) = \mathbf{q}$, with $P = \begin{pmatrix} 1 & -1 & 0 \end{pmatrix}$. The vector $\mathrm{E}(\mathbf{r})$ is unknown and equals $\mathrm{E}(\mathbf{r}) = \begin{pmatrix} \mathrm{E}(r_A) & \mathrm{E}(r_B) & \mathrm{E}(r_C) \end{pmatrix}'$, the variable q represents the forecast of the assets return, she forecasts this to be: $q = 2\%$. Performing the actual multiplication highlights the meaning of the formula: $P\mathrm{E}(\mathbf{r}) = \mathrm{E}(r_A) - \mathrm{E}(r_B) = 2\%$, the investor has expressed a view that she expects the difference between the expected return of asset A and B to be 2%.
When the matrix P represents a collection of portfolios, every row of P represents one portfolio, and the corresponding element of $\mathbf{q}$ is its expected return. The example would then translate to a zero-investment portfolio whose expected return is 2%.
An example can make this manner of expressing views more clear.

Example 5 (Expressing views)**.** *Consider an investor that has a benchmark of eights asset classes, see table 4.1.*[2]

Asset Class	**Weight**
US Bonds	19.34 %
International Bonds	26.13 %
US Large Growth	12.09 %
US Large Value	12.09 %
US Small Growth	1.34 %
US Small Value	1.34 %
International developed equity	24.18%
International emerging equity	3.49%
Sum	**100.0%**

Table 4.1: The benchmark of assets.

She holds three views about the performance of assets:
View 1: The asset class international developed equity will have an absolute excess return of 5.25 %
View 2: International bonds will outperform US bonds by 25 basis points, or equivalently 0.25 %.
View 3: US large growth and US small growth asset will outperform US large value and US small value by 2%.

[2]The example is derived from Idzorek (2004).

These views can be expressed in matrix form in the following way:

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & -0.5 & 0.5 & -0.5 & 0 & 0 \end{pmatrix} \mathrm{E}(\mathbf{r}) = \begin{pmatrix} 5.25 \\ 0.25 \\ 2 \end{pmatrix} + \boldsymbol{\epsilon}.$$

There are some variations possible on the specification of the view matrix P and the confidence level $\boldsymbol{\epsilon}$. Various authors have different views, which will be discussed in paragraph 4.4.2 and 4.4.1.
The subsection can be summarized in the following assumption:

Assumption. *A3 The investor has $k < n$ views, expressed as a linear relationship $P\mathrm{E}(\mathbf{r}) = \mathbf{q} + \boldsymbol{\epsilon}$. Where $P \in \mathbb{R}^{k \times n}$, $\mathbf{q} \in \mathbb{R}^k$, $\boldsymbol{\epsilon} \in \mathbb{R}^k \sim \mathrm{N}(0, \Omega)$ and $\Omega \in \mathbb{R}^{k \times k}$ is a diagonal covariance matrix.*

4.2.4 The Bayesian approach: Combining views with equilibrium[3]

Black and Litterman do not explicitly state how they arrive at the BL-formula, they suggest the mixed estimation method of Theil (1971) that will be discussed in the next subsection, and they suggest that $\mathrm{E}(\mathbf{r})$ is the product of two normal distributions. Satchell and Scowcroft (2000) use a Bayesian approach to obtain an estimate for the vector of expected returns as a product of two normal distributions. This approach will be discussed in this subsection. First we will discuss Bayes' formula in general and then move to the application of the formula to the problem at hand.
In classical statistical analysis, unknown parameters are estimated by a set of observed data. The Bayesian approach, however proposes that views about the state of the world are subjective.[4] Instead of estimating parameters as if they were fixed, as if only the one set of observation should be used, the Bayesian approach proposes to continuously update the set of observed data with recently observed data in order to sharpen the subjective prior beliefs about the current state. The centerpiece driving all Bayesian methods is Bayes' Theorem.

Theorem 4.2.1 (Bayes' Theorem)**.**

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\,\mathbb{P}(A)}{\mathbb{P}(B)} \tag{4.3}$$

This is often restated as

$$\mathbb{P}(A|B) \propto \mathbb{P}(A)\ell(A|B)$$

[3]The text is derived from the works by Satchell and Scowcroft (2000) and Lee (2000).
[4]A good introductory book on Bayesian statistics is that of Lee (1997).

for $\mathbb{P}(B) \neq 0$. Here $\propto$ means 'is proportional to' and ℓ denotes the likelihood function. $\mathbb{P}$ denotes a discrete or a continuous probability distribution.
The probability distribution prior to any observations is $\mathbb{P}(A)$, this is called the prior distribution. After the observations of new data represented by B a new distribution can be derived for $\mathbb{P}(A|B)$, this is called the posterior distribution.

Proof. This is a reworking of the traditional formula for conditional probabilities:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A)\,\mathbb{P}(A)}{\mathbb{P}(B)}.$$

□

Bayes' Theorem is mathematically not very challenging. To apply the theorem to the problem at hand is less straightforward. What should be the prior distribution and the posterior distribution in this case and how should they be specified?
It is assumed that the investor forms her views using knowledge of the equilibrium expected returns. Therefore, the equilibrium expected returns are considered the prior returns and these will be updated with the views of the investor. The posterior distribution combines both sources of information. Using Bayes' formula this yields:

$$\mathbb{P}\left(P\mathrm{E}(\mathbf{r})|\mathrm{E}(\mathbf{r})\right) = \frac{\mathbb{P}\left(\mathrm{E}(\mathbf{r})|P\mathrm{E}(\mathbf{r})\right)\ \mathbb{P}\left(P\mathrm{E}(\mathbf{r})\right)}{\mathbb{P}\left(\mathrm{E}(\mathbf{r})\right)}. \tag{4.4}$$

We would like to make one note of comment on the formula used by Satchell and Scowcroft (2000), they use:

$$\mathbb{P}(\mathrm{E}(\mathbf{r})|\boldsymbol{\pi}) = \frac{\mathbb{P}(\boldsymbol{\pi}|\mathrm{E}(\mathbf{r}))\mathbb{P}(\mathrm{E}(\mathbf{r}))}{\mathbb{P}(\boldsymbol{\pi})} \tag{4.5}$$

to compute the posterior distribution. However, in their proof they substitute the probability distribution of $P\mathrm{E}(\mathbf{r})$ instead of the one for $\mathrm{E}(\mathbf{r})$. To be mathematically correct one should use formula (4.4). The formula they use looks intuitive, as it seems to compute the probability distribution of the expected return vector given the equilibrium returns, however the formula is used incorrectly.

Probability distributions The probability distribution of the views can be deduced from assumption A3: $P\mathrm{E}(\mathbf{r})|\mathrm{E}(\mathbf{r}) \sim N(\mathbf{q}, \Omega)$.

Assumption. *A4 $P\mathrm{E}(\mathbf{r})|\mathrm{E}(\mathbf{r})$ is normally distributed with mean $\mathbf{q}$ and diagonal covariance matrix Ω: $P\mathrm{E}(\mathbf{r})|\mathrm{E}(\mathbf{r}) \sim N(\mathbf{q}, \Omega)$.*

Equilibrium returns The reasoning on the distribution of the expected returns is more complicated, than the reasoning on the distribution of the investors views. Black and Litterman (1992) assume first that the return of the assets are normally distributed with mean $\mathrm{E}(\mathbf{r})$ and variance Σ. Furthermore, they assume that this mean itself is a random variable, is unobservable and stochastic. In Bayesian statistics, $\mathrm{E}(\mathbf{r})$ would be called a hyperparameter, and could be approximated by another application of Bayes' Theorem.
They however, do not follow this path. Instead, they choose a distribution for $\mathrm{E}(\mathbf{r})$ heuristically. They assume that the market is always moving to equilibrium, and is not necessarily in equilibrium. Therefore, the mean of the expected return should be equal to the expected returns that would hold if the market is in equilibrium, i.e. equal to the CAPM returns. The variance of $\mathrm{E}(\mathbf{r})$ is assumed to be proportional to the variance of the returns r, proportional to Σ with proportionality constant τ. The constant τ will be close to zero, because the uncertainty in the mean of the return is much smaller than the uncertainty in the return itself. "The equilibrium risk premiums together with $\tau\Sigma$ determine the equilibrium distribution for expected excess returns."

Assumption. *A5 The expected return* $\mathrm{E}(\mathbf{r})$ *is a random variable, which is normally distributed with mean* $\boldsymbol{\pi}$ *and variance* $\tau\Sigma$*:* $\mathrm{E}(\mathbf{r}) \sim N(\boldsymbol{\pi}, \tau\Sigma)$.

The contribution of Black and Litterman has been to put this problem into a tractable form.
This brief introduction into Bayesian statistics will be sufficient to derive the formula for the posterior distribution of the expected return.

Theorem 4.2.2 (Black-Litterman formula)**.** *Under the assumptions A1 to A4 the posterior distribution* $E(\mathbf{r})|P\mathrm{E}(\mathbf{r})$ *is normally distributed with mean*

$$[(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1}[(\tau\Sigma)^{-1}\boldsymbol{\pi} + P'\Omega^{-1}\mathbf{q}] \tag{4.6}$$

and variance

$$[(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1}. \tag{4.7}$$

Proof. The proof is a straightforward application of Bayes' Theorem.[5] The probability distributions for the prior distribution and the likelihood function will be substituted, resulting in a complicated exponent that will be transformed by some matrix manipulation to resemble a normal distribution with the to be proved mean and variance.
We will briefly restate the assumptions that will be used in the proof.

Assumption. *A4* $P\mathrm{E}(\mathbf{r})|\mathrm{E}(\mathbf{r})$ *is normally distributed with mean* $\mathbf{q}$ *and diagonal covariance matrix* Ω*:* $P\mathrm{E}(\mathbf{r})|\mathrm{E}(\mathbf{r}) \sim N(\mathbf{q}, \Omega)$.

[5] The proof is adapted from a proof in Satchell and Scowcroft (2000).

A5 *The expected return* $\mathrm{E}(\mathbf{r})$ *is a random variable, which is normally distributed with mean* $\boldsymbol{\pi}$ *and variance* $\tau\Sigma$*:* $\mathrm{E}(\mathbf{r}) \sim N(\boldsymbol{\pi}, \tau\Sigma)$.

Bayes' Theorem in this context can be expressed as:

$$\mathbb{P}(\mathrm{E}(\mathbf{r})|P\mathrm{E}(\mathbf{r})) = \frac{\mathbb{P}(P\mathrm{E}(\mathbf{r})|\mathrm{E}(\mathbf{r}))\mathbb{P}(\mathrm{E}(\mathbf{r}))}{\mathbb{P}(P\mathrm{E}(\mathbf{r}))}. \tag{4.8}$$

Assumption A4 states that the probability density function of $P\mathrm{E}(\mathbf{r})|\mathrm{E}(\mathbf{r})$ is multivariate normally distributed with $P\mathrm{E}(\mathbf{r})|\mathrm{E}(\mathbf{r}) \sim N(\mathbf{q}, \Omega)$:

$$\mathbb{P}(P\mathrm{E}(\mathbf{r})|\mathrm{E}(\mathbf{r})) = \frac{1}{\sqrt{(2\pi)^k \det(\Omega)}} \exp[-\frac{1}{2}(P\mathrm{E}(\mathbf{r}) - \mathbf{q})'\Omega^{-1}(P\mathrm{E}(\mathbf{r}) - \mathbf{q})].$$

Assumption A5 states that the probability density function of $\mathrm{E}(\mathbf{r})$ is multivariate normally distributed with $\mathrm{E}(\mathbf{r}) \sim N(\boldsymbol{\pi}, \tau\Sigma)$:

$$\mathbb{P}(\mathrm{E}(\mathbf{r})) = \frac{1}{\sqrt{(2\pi)^n \det(\tau\Sigma)}} \exp[-\frac{1}{2}(\mathrm{E}(\mathbf{r}) - \boldsymbol{\pi})'(\tau\Sigma)^{-1}(\mathrm{E}(\mathbf{r}) - \boldsymbol{\pi})].$$

These distributions can be substituted in equation (4.8) to obtain the posterior distribution. We will first concentrate on the numerator of the equation (4.8):

$$\mathbb{P}(\mathrm{E}(\mathbf{r})|P\mathrm{E}(\mathbf{r})) \propto \mathbb{P}(P\mathrm{E}(\mathbf{r})|\mathrm{E}(\mathbf{r}))\ \mathbb{P}(\mathrm{E}(\mathbf{r})).$$

Substituting the distributions into the formula gives

$$\begin{aligned}\mathbb{P}(\mathrm{E}(\mathbf{r})|P\mathrm{E}(\mathbf{r})) \propto\ & \frac{1}{\sqrt{(2\pi)^k \det(\Omega)}} \exp[-\frac{1}{2}(P\mathrm{E}(\mathbf{r}) - \mathbf{q})'\Omega^{-1}(P\mathrm{E}(\mathbf{r}) - \mathbf{q})] \\ & \cdot \frac{1}{\sqrt{(2\pi)^n \det(\tau\Sigma)}} \exp[-\frac{1}{2}(\mathrm{E}(\mathbf{r}) - \boldsymbol{\pi})'(\tau\Sigma)^{-1}(\mathrm{E}(\mathbf{r}) - \boldsymbol{\pi})].\end{aligned}$$

We leave out all the constants and are left with:

$$\begin{aligned}\mathbb{P}(\mathrm{E}(\mathbf{r})|P\mathrm{E}(\mathbf{r})) \propto \exp[&-\frac{1}{2}(P\mathrm{E}(\mathbf{r}) - \mathbf{q})'\Omega^{-1}(P\mathrm{E}(\mathbf{r}) - \mathbf{q}) \\ &- \frac{1}{2}(\mathrm{E}(\mathbf{r} - \boldsymbol{\pi}))'(\tau\Sigma)^{-1}(\mathrm{E}(\mathbf{r}) - \boldsymbol{\pi})].\end{aligned}$$

This formula will be transformed in a such way that the general formula for a normal distribution becomes apparent, i.e. the probability density function of a multivariate normally distributed variable $X \in \mathbb{R}^n$ with mean $\boldsymbol{\mu}$ and variance Σ can be represented as $\mathbb{P}(X) \propto \exp[-\frac{1}{2}(X - \boldsymbol{\mu})'\Sigma^{-1}(X - \boldsymbol{\mu})]$. In our case it should become apparent that $\boldsymbol{\mu}$ equals the mean of the BL-formula and Σ the variance of the formula. In order to get to this result we will concentrate on the $(X - \boldsymbol{\mu})'\Sigma^{-1}(X - \boldsymbol{\mu})$ part of the exponent and leave out the rest for the moment.

Expanding the brackets in the exponent and dropping the exponent and the factor a half in the exponent gives:

$$\begin{aligned}&\mathrm{E}(\mathbf{r})'P'\Omega^{-1}P\mathrm{E}(\mathbf{r}) - \mathrm{E}(\mathbf{r})'P'\Omega^{-1}\mathbf{q} - \mathbf{q}'\Omega^{-1}P\mathrm{E}(\mathbf{r}) + \mathbf{q}'\Omega^{-1}\mathbf{q}\\ &\quad + \mathrm{E}(\mathbf{r})'(\tau\Sigma)^{-1}\mathrm{E}(\mathbf{r}) - \mathrm{E}(\mathbf{r})'(\tau\Sigma)^{-1}\boldsymbol{\pi} - \boldsymbol{\pi}'(\tau\Sigma)^{-1}\mathrm{E}(\mathbf{r}) + \boldsymbol{\pi}'(\tau\Sigma)^{-1}\boldsymbol{\pi}.\end{aligned}$$

The term $\mathbf{q}'\Omega^{-1}P\mathrm{E}(\mathbf{r})$ is equal to $\mathrm{E}(\mathbf{r})'P'\Omega^{-1}\mathbf{q}$ due to the symmetric property of Ω, the same holds for $\mathrm{E}(\mathbf{r})'(\tau\Sigma)^{-1}\boldsymbol{\pi}$ and $\boldsymbol{\pi}'(\tau\Sigma)^{-1}\mathrm{E}(\mathbf{r})$, now due to the symmetry of Σ. The previous equation can thus be shortened to:

$$\begin{aligned}&\mathrm{E}(\mathbf{r})'P'\Omega^{-1}P\mathrm{E}(\mathbf{r}) - 2\mathbf{q}'\Omega^{-1}P\mathrm{E}(\mathbf{r}) + \mathbf{q}'\Omega^{-1}\mathbf{q} + \mathrm{E}(\mathbf{r})'(\tau\Sigma)^{-1}\mathrm{E}(\mathbf{r})\\ &\qquad - 2\boldsymbol{\pi}'(\tau\Sigma)^{-1}\mathrm{E}(\mathbf{r}) + \boldsymbol{\pi}'(\tau\Sigma)^{-1}\boldsymbol{\pi}.\end{aligned}$$

This can be expanded even further to:

$$\begin{aligned}&\mathrm{E}(\mathbf{r})'[P'\Omega^{-1}P + (\tau\Sigma)^{-1}]\mathrm{E}(\mathbf{r}) - 2[\mathbf{q}'\Omega^{-1}P + \boldsymbol{\pi}'(\tau\Sigma)^{-1}]\mathrm{E}(\mathbf{r})\\ &\qquad + \mathbf{q}'\Omega^{-1}\mathbf{q} + \boldsymbol{\pi}'(\tau\Sigma)^{-1}\boldsymbol{\pi}.\end{aligned}$$

The formula will be simplified by introducing three symbols C, H, A.

$$\begin{aligned}&C = (\tau\Sigma)^{-1}\boldsymbol{\pi} + P'\Omega^{-1}\mathbf{q},\\ &H = (\tau\Sigma)^{-1} + P'\Omega^{-1}P, \text{ where } H \text{ is symmetrical } H = H',\\ &A = \mathbf{q}'\Omega^{-1}\mathbf{q} + \boldsymbol{\pi}'(\tau\Sigma)^{-1}\boldsymbol{\pi}.\end{aligned}$$

With this shortened notation we can rewrite the exponent to:

$$\mathrm{E}(\mathbf{r})'H\mathrm{E}(\mathbf{r}) - 2C'\mathrm{E}(\mathbf{r}) + A = \mathrm{E}(\mathbf{r})'H'\mathrm{E}(\mathbf{r}) - 2C'\mathrm{E}(\mathbf{r}) + A.$$

We introduce the identity matrix $I = H^{-1}H$ into the equation which will prove useful.

$$\begin{aligned}&\mathrm{E}(\mathbf{r})'H'\mathrm{E}(\mathbf{r}) - 2C'\mathrm{E}(\mathbf{r}) + A\\ &= (H\mathrm{E}(\mathbf{r}))'H^{-1}H\mathrm{E}(\mathbf{r}) - 2C'H^{-1}H\mathrm{E}(\mathbf{r}) + A\\ &= (H\mathrm{E}(\mathbf{r}) - C)'H^{-1}(H\mathrm{E}(\mathbf{r}) - C) + A - C'H^{-1}C\\ &= (\mathrm{E}(\mathbf{r}) - H^{-1}C)'H(\mathrm{E}(\mathbf{r}) - H^{-1}C) + A - C'H^{-1}C.\end{aligned}$$

The terms in $A - C'HC$ do not depend on $P\mathrm{E}(\mathbf{r})$ and therefore disappear into the constant of integration. Thus, reintroducing the exponent and the factor half in the exponent leaves us with:

$$\mathbb{P}(\mathrm{E}(\mathbf{r})|P\mathrm{E}(\mathbf{r})) \propto \exp[-\frac{1}{2}(\mathrm{E}(\mathbf{r}) - H^{-1}C)'H(\mathrm{E}(\mathbf{r}) - H^{-1}C)].$$

Hence, $\mathrm{E}(\mathbf{r})|P\mathrm{E}(\mathbf{r})$ has mean

$$H^{-1}C = [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1}[(\tau\Sigma)^{-1}\boldsymbol{\pi} + P'\Omega^{-1}\mathbf{q}]$$

and variance

$$H^{-1} = [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1}.$$

□

The Bayesian approach after a somewhat lengthy calculation thus, leads to a formula for the posterior distribution of the expected return that combines the investors views and the market views. The concept of updating the views of the investor with the views of the market is fairly intuitive, and makes sense. A major drawback of this method is the assumption about the equilibrium returns. The parameter τ has to be set, but there is no real indication about how to do this.
Before we investigate the other way to derive the Black-Litterman formula, we investigate the meaning of the formula.

Interpretations The solution in equation (4.6) looks quite complicated. Nevertheless, the formula does comply with the specifications that if the investor does not have any views on the assets, the equilibrium distribution of returns should be used. This will lead to holding the market portfolio as is also implied by the CAPM.

$$\widehat{\mathrm{E}(\mathbf{r})} = [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1}[(\tau\Sigma)^{-1}\boldsymbol{\pi} + P'\Omega^{-1}\mathbf{q}]$$

The BL-formula has the same result as can be seen by taking $P = 0$. In that case we are only left with $[(\tau\Sigma)^{-1}]^{-1}[(\tau\Sigma)^{-1}\boldsymbol{\pi}] = \boldsymbol{\pi}$, the equilibrium returns. Therefore the BL-formula fulfills this basic characteristic.
It is also interesting to investigate the other limiting case: if the investor is totally confident in his views what will the vector of expected returns be? A total confidence in views would imply that the variance on this view is zero, it is not possible to substitute this in the BL-formula because this would require the inversion of a zero matrix. It can be shown the the BL-formula then converges to the views of the investor ($\mathbf{q}$).

4.2.5 Mixed estimation method of Theil

Black and Litterman also propose a different method to derive the combined formula of expected returns: the mixed estimation method of Theil and Goldberger (1961), which is more thoroughly described in the book by Theil (1971). The model is an adaptation of the generalized least square estimation method. Theil (1971) shows that it yields the same results as a Bayesian estimation method. The main advantage of this method is that it is much easier to derive the estimator with the mixed estimation technique. The mixed estimation method of Theil is a generalized least square estimation method. Where traditional least square estimation only allows data from one source, the mixed estimation method allows two separate sources of information.
In order to explain the mixed estimation method of Theil, we will first recall generalized least square estimation and subsequently develop the connection with mixed estimation.

Generalized least square estimation[6] Generalized least square estimation gives an estimator for a dependent variable. In general, one wants to estimate the vector $\boldsymbol{\beta}$, from the observed data in the vector $\mathbf{y}$. The observed data has some estimation error, represented by the vector $\mathbf{e}$, it is assumed that this estimation error has a normal probability distribution. Additionally, the relation between the observed data $\mathbf{y}$ and the to be estimated data is linear in some matrix X. More specifically a relationship $\mathbf{y} = X\boldsymbol{\beta} + \mathbf{e}$ is assumed to hold.
The main theorem, also known as Aitken's theorem is the following.

Theorem 4.2.3 (Aitken)**.** *Given is the specification $\mathbf{y} = X\boldsymbol{\beta} + \mathbf{e}$, where $\mathbf{y} \in \mathbb{R}^k$ is a known vector, where $X \in \mathbb{R}^{(k\times n)}$ is a full column rank matrix, $\boldsymbol{\beta} \in \mathbb{R}^n$ is a unknown vector and $\mathbf{e} \in \mathbb{R}^k$ is an error term. The error term has zero mean $\mathrm{E}(\boldsymbol{\epsilon}) = \mathbf{0}$ and variance $\mathrm{var}(\boldsymbol{\epsilon}) = \Sigma$. Suppose that X is a non-stochastic matrix and $\mathrm{E}(\mathbf{y}|X) = X\boldsymbol{\beta}$. Also suppose that $\mathrm{var}(\mathbf{y}|X) = \Sigma \in \mathbb{R}^{k\times k}$, where Σ is a positive definite matrix.*
Then,

$$\hat{\boldsymbol{\beta}} = (X'\Sigma^{-1}X)^{-1}(X'\Sigma^{-1}\mathbf{y}) \tag{4.9}$$

is the best linear unbiased estimator for $\boldsymbol{\beta}$ and the covariance matrix of the estimator is $(X'\Sigma^{-1}X)^{-1}$. It is the best estimator in the sense that any other estimator of $\boldsymbol{\beta}$ which is also linear in the vector $\mathbf{y}$ and unbiased has a covariance matrix which exceeds that of $\hat{\boldsymbol{\beta}}$ by a positive definite matrix.

Given a set of observed data, the generalized linear estimation technique obtains an estimator for $\boldsymbol{\beta}$. In this case there are two separate sets of data to make an estimate of $\mathrm{E}(\mathbf{r})$ from.

Mixed Estimation The information from the investment manager and information from the market can be used to estimate the vector of expected returns $\mathrm{E}(\mathbf{r})$. Mixed estimation says now to incorporate both sources of information in the model and not dismiss one of the sources as false. It is probable that one has a priori information about the unknown parameter, and these ideas could then be updated with sample information. The concept of mixed estimation is very close to Bayesian inference, which also uses a prior distribution which is then updated with sample information. A derivation of the Black-Litterman formula with this method can be found in Koch (2004).

Model specification The two sources of information have previously been specified in terms of probability distributions. For generalized least

[6] An explanation of generalized least square estimation can be found in any good econometrics book, for example Theil (1971).

square estimation they have to be specified as a linear model. The to be estimated parameter is the expected return $\mathrm{E}(\mathbf{r}) = \boldsymbol{\beta}$.
The prior information consists of the views of the investment manager about the performance of certain assets. The views of the investment manager are already represented by Black and Litterman in the desired form.

$$\begin{aligned}
&\mathbf{q} = P\mathrm{E}(\mathbf{r}) + \boldsymbol{\epsilon},\\
&\quad \mathbf{q} \in \mathbb{R}^k \text{ a known vector}\\
&\quad P \in \mathbb{R}^{k\times n},\ k < n \text{ is a known matrix of rank } k\\
&\quad \boldsymbol{\epsilon} \in \mathbb{R}^k \text{ the vector of errors}\\
&\quad \text{such that } \mathrm{E}(\boldsymbol{\epsilon}) = 0,\ \mathrm{var}(\boldsymbol{\epsilon}) = \Omega \text{ non-singular}\\
&\quad \text{or equivalently } \mathrm{E}(\mathbf{q}) = P\mathrm{E}(\mathbf{r}),\ \mathrm{var}(\mathbf{q}) = \Omega
\end{aligned}$$

The sample observation are the equilibrium returns. They have previously been expressed as $\mathrm{E}(\mathbf{r}) \sim \mathrm{N}(\boldsymbol{\pi}, \tau\Sigma)$. This assumption will now be slightly modified, the equilibrium return vector $\boldsymbol{\pi}$ is our observed variable and $\mathrm{E}(\mathbf{r})$ is the value we would like to estimate.

$$\begin{aligned}
&\boldsymbol{\pi} = \mathrm{E}(\mathbf{r}) + \mathbf{u},\\
&\quad \boldsymbol{\pi} \in \mathbb{R}^n \text{ the observed vector of equilibrium returns,}\\
&\quad \mathrm{E}(\mathbf{r}) \in \mathbb{R}^n \text{ is the to be estimated vector of expected returns,}\\
&\quad \mathbf{u} \in \mathbb{R}^n \text{ a vector of errors,}\\
&\quad \text{such that } \mathrm{E}(\mathbf{u}) = 0,\ \mathrm{var}(\mathbf{u}) = \tau\Sigma \ \text{ non-singular,}\\
&\quad \text{or equivalently } \mathrm{E}(\boldsymbol{\pi}) = \mathrm{E}(\mathbf{r}) \text{ and } \mathrm{var}(\boldsymbol{\pi}) = \tau\Sigma.
\end{aligned}$$

Aitken's theorem expresses the linear relationship as $\boldsymbol{\pi} = X\boldsymbol{\beta} + \mathbf{u}$. Our linear relationship corresponds to this when the matrix X is identified with the identity matrix I, which clearly has full column rank.
To estimate $\mathrm{E}(\mathbf{r})$, both sources of information are used by stacking the respective vectors and matrices. The linear equation then becomes:

$$\begin{pmatrix}\boldsymbol{\pi}\\ \mathbf{q}\end{pmatrix} = \begin{pmatrix} I \\ P\end{pmatrix}\mathrm{E}(\mathbf{r}) + \begin{pmatrix}\mathbf{u}\\ \boldsymbol{\epsilon}\end{pmatrix} \tag{4.10}$$
$$\text{where } \mathrm{var}(\begin{pmatrix}\mathbf{u}\\ \boldsymbol{\epsilon}\end{pmatrix}) = \mathrm{var}(\begin{pmatrix}\boldsymbol{\pi}\\ \mathbf{q}\end{pmatrix}) = \begin{pmatrix}\tau\Sigma & 0\\ 0 & \Omega\end{pmatrix} = W.$$

Note that this can be written in the form: $\mathbf{y} = X\boldsymbol{\beta} + \mathbf{e}$. It can be seen that the linear system (4.10) complies with Aitken's Theorem. The matrix $X = \begin{pmatrix} I \\ P\end{pmatrix}$ has full column rank due to the identity matrix and the matrix W is positive definite due to the positive definiteness of Σ and the diagonality of Ω. Furthermore, the two sources of information are independent.

The estimator of $\mathrm{E}(\mathbf{r})$ can be found by substituting the appropriate parameters in equation (4.9) (this short derivation can be found in Koch (2004)):

$$\begin{aligned}\widehat{\mathrm{E}(\mathbf{r})} &= (X'W^{-1}X)^{-1}(X'W^{-1}\mathbf{y}) \\ &= [\begin{pmatrix} I & P' \end{pmatrix} \begin{pmatrix} \tau\Sigma & 0 \\ 0 & \Omega \end{pmatrix}^{-1} \begin{pmatrix} I \\ P \end{pmatrix}]^{-1} [\begin{pmatrix} I & P' \end{pmatrix} \begin{pmatrix} \tau\Sigma & 0 \\ 0 & \Omega \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{\pi} \\ \mathbf{q} \end{pmatrix}] \\ &= [\begin{pmatrix} (\tau\Sigma)^{-1} & P'\Omega^{-1} \end{pmatrix} \begin{pmatrix} I \\ P \end{pmatrix}]^{-1} [\begin{pmatrix} (\tau\Sigma)^{-1} & P'\Omega^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\pi} \\ \mathbf{q} \end{pmatrix}] \\ &= [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1}[(\tau\Sigma)^{-1}\boldsymbol{\pi} + P'\Omega^{-1}\mathbf{q}]\end{aligned}$$

The derivation of the formula is in this case remarkably shorter than in the Bayesian approach.

Least square estimate of $\boldsymbol{\tau}$ The mixed estimation method of Theil (1971) additionally provides an estimate for τ, however our specification of the model does not fulfill the requirements for the estimator. The estimator not only requires a full column rank matrix X, in our case the identity matrix I, but it also requires that there are more observations than to be estimated data, i.e. row dimension of X is larger than the column dimension. The identity matrix cannot fulfill this requirement, as column dimension equals the row dimension, and the estimator for τ cannot be used.

$$\hat{\tau} \neq \frac{1}{n-k}(\boldsymbol{\pi} - I\hat{\boldsymbol{\beta}}_{\boldsymbol{\pi}})'\Sigma^{-1}(\boldsymbol{\pi} - I\hat{\boldsymbol{\beta}}_{\boldsymbol{\pi}}). \tag{4.11}$$

Where $\hat{\boldsymbol{\beta}}_{\boldsymbol{\pi}}$ is the estimator of $\mathrm{E}(\mathbf{r})$ when only the equilbrium returns are known. In our case $n = k$ and thus the formula cannot be used.

4.2.6 Summary

The Black-Litterman model combines views of the investor and the market equilibrium on the expected return of assets in one formula. This formula should be a better approximation of the expected returns. These expected returns, or more precisely the estimator of the expected return, could then be used in a mean-variance optimizer.
The Black-Litterman model can be summarized by the following points:

1. The market consists of n assets. The assets have a return $\mathbf{r} \in \mathbb{R}^n$, with variance Σ and expected return $\mathrm{E}(\mathbf{r})$. The expected return $\mathrm{E}(\mathbf{r})$ is an unknown and normally distributed random variable, it is assumed to have mean $\boldsymbol{\pi}$ and variance $\tau\Sigma$.

2. The first source of information about $\mathrm{E}(\mathbf{r})$ are the equilibrium returns $\boldsymbol{\pi}$. The equilibrium returns are found by $\boldsymbol{\pi} = \delta\Sigma\mathbf{w}_m$, where δ is a (world) risk aversion coefficient or a ratio of the (world) market

portfolio and $\mathbf{w}_m$ are the market weights. They can be represented as $\boldsymbol{\pi} = \mathrm{E}(\mathbf{r}) + \mathbf{u}$, with $\mathbf{u} \sim \mathrm{N}(0, \tau\Sigma)$, where τ is some proportionality constant.

3. The second source of information are the k views of the investor. The views are expressed as $P\mathrm{E}(\mathbf{r}) = \mathbf{q} + \boldsymbol{\epsilon}$, where $P \in \mathbb{R}^{n\times n}$, $\mathbf{q} \in \mathbb{R}^n$, and $\boldsymbol{\epsilon} \sim \mathrm{N}(0, \Omega)$, Ω is a diagonal $(k \times k)$-matrix.

4. Combination of these two sources of information leads to $\mathrm{E}(\mathbf{r})$ being normally distributed with mean $[(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1}[(\tau\Sigma)^{-1}\boldsymbol{\pi} + P'\Omega^{-1}\mathbf{q}]$ and variance $[(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1}$.

5. This mean can be used in a mean-variance optimization process to obtain a mean-variance efficient portfolio.

There are a few subjects that could be explained more thoroughly, especially about the view matrix P, the matrix Ω and the parameter τ. Also, Black and Litterman meant the model to be used in a global context and to allocate not only equity, but also bonds and currency. In this global context the model should be modified slightly. This we will do in the next section.

4.3 A global allocation model

The Black-Litterman model is not only a model to allocate equity, but it can also be used to allocate bonds and currency. Investing in more asset classes is desirable, as it gives the possibility to combine the characteristics of both classes. Also, it gives a wider choice of assets which in turn widens the possibilities for diversification and thus reducing risk.
Furthermore, it is meant as a global allocation model, which makes it possible to diversify even more. Unfortunately, it also opens the investor up to a new source of risk: exchange risk. The investment in a foreign asset could perform well, but if the exchange rate of the currency drops relative to the domestic currency, a substantial loss on the investment could occur.
This problem can be alleviated by supplementing the portfolio with a suitable amount of foreign currency, which corresponds to the market capitalization of the assets held in that currency (Black, 1989). This procedure is known is as *hedging*. The question then remains, how much to hedge?

4.3.1 Universal hedging

Black (1989, 1990) has determined a very simple optimal hedging formula. The formula requires three inputs: the expected return of the world market portfolio, the standard deviation (volatility) of the world market portfolio and the average exchange rate volatility. In turn the formula yields three results: hedge your foreign equity, hedge the equities by the same ratio for all countries and do not hedge the total amount of foreign equity, it is only necessary to hedge a ratio of the foreign equity.

Theorem 4.3.1 (Universal hedging formula)**.** *Assume there are no taxes, trading cost or other barriers to international investment or disinvestment. Assume that every investor is a mean-variance investor and can borrow or lend without fear of default. Assume that the risk aversion coefficient* λ_i *is the same for every investor, hence* $\lambda_i = \lambda$*.*
Then each investor should hedge his global asset portfolio by holding a total short position of $1 - \lambda$ *in each currency of:*

$$1 - \lambda = \frac{\mu_m - \sigma_m^2}{\mu_m - \frac{1}{2}\sigma_e^2}.$$

Where μ_m *is the average world market portfolio expected excess return.*
Where σ_m^2 *is the average world market portfolio excess return variance.*
Where σ_e^2 *is the average exchange rate variance.*

The proof of this can be found in Black (1990). The result can be explained from Siegel's paradox. The basic idea is that, because investors in different countries measure returns in different units, each will gain some expected return by taking some currency risk. Investors will accept currency risk up to the point where the additional risk balances the expected return. Under certain simplifying assumptions, the percentage of foreign currency risk hedged will be the same for investors of different countries, therefore called "universal hedging".
The formula seems very simple, but it is difficult to determine exactly the right value for the universal hedging constant, primarily because the expected excess return on the market portfolio is difficult to estimate. Nevertheless, Black and Litterman (1992) "feel that universal hedging values between 75% and 85% are reasonable".

4.3.2 Mathematics in a global context

The mathematical description of the BL-model changes slightly in this global context. The expected return vector ($\boldsymbol{\pi}$) now represents the expected excess returns of a global market portfolio that is demonetized in the domestic currency.
The parameter δ becomes a global constant. Black and Litterman state that δ "is a proportionality constant based on the formulas in Black (1989)". It is probable that they mean the world risk aversion coefficient (λ) from Black (1990) instead.
The vector of weights now not only contains equity, but also bonds and currency. When the benchmark portfolio is the market portfolio as in the (1991a) article by Black and Litterman, then the weight of the currency

position is determined by the weight of the equity and bond position in that country and the universal hedging constant: $w_i = \lambda w_j^c$, where w_j^c is the country weight, i.e. the sum of the market weights for bonds and equities in the j'th country.

4.4 In-depth analysis

4.4.1 The parameter τ and the matrix Ω

The parameter τ is one of the least understood parameters in the model, as little indication is given on how to set it. Black and Litterman introduce it as a proportionality constant to scale the variance of the expected return. They first assume that the return is a random variable with mean $\mathrm{E}(\mathbf{r})$ and variance Σ, and next assume that the mean itself is a random variable with a probability distribution centered at the equilibrium returns and variance proportional to the variance of the return. They choose τ close to zero because the uncertainty about the mean is much smaller than the uncertainty of the return itself. The only indication they give is that it should be chosen close to zero.
The parameter has therefore been source of much confusion. Satchell and Scowcroft (2000) for instance state the opposite of what Black and Litterman advice, i.e. that the parameter τ it is often set to one. However, they also found a method to circumvent the problem of giving an estimate for τ: they assume that it is stochastic and derive a new form of the Black-Litterman formula in that case. The new formula, however becomes so complicated that a non-mathematical investment manager can no longer use it.
The parameter τ also plays a role in the BL-formula $\widehat{\mathrm{E}(\mathbf{r})} = [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1}[(\tau\Sigma)^{-1}\boldsymbol{\pi} + P'\Omega^{-1}\mathbf{q}]$. He and Litterman (1999) claim that Ω and τ need not be specified separately as only the ratio ω/τ enters into the formula. After a reworking of the original formula it can been seen that this is correct:

$$\begin{aligned}
\widehat{\mathrm{E}(\mathbf{r})} &= [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1}[(\tau\Sigma)^{-1}\boldsymbol{\pi} + P'\Omega^{-1}\mathbf{q}] \\
&=^1 [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1}\tau^{-1}\tau[(\tau\Sigma)^{-1}\boldsymbol{\pi} + P'\Omega^{-1}\mathbf{q}] \\
&=^2 [\tau(\tau\Sigma)^{-1} + \tau P'\Omega^{-1}P]^{-1}[\tau(\tau\Sigma)^{-1}\boldsymbol{\pi} + \tau P'\Omega^{-1}\mathbf{q}] \\
&= [\Sigma^{-1} + P'(\tau^{-1}\Omega)^{-1}P]^{-1}[\Sigma^{-1}\boldsymbol{\pi} + P'(\tau^{-1}\Omega)^{-1}\mathbf{q}].
\end{aligned}$$

$=^1$ *The identity $\tau^{-1}\tau = 1$ is inserted.*
$=^2$ *The matrix property $A^{-1}B^{-1} = (BA)^{-1}$ is used.*

He and Litterman explain in a footnote that "the confidence level on a single view $p\mathrm{E}(\mathbf{r}) = q + \epsilon$ is calibrated such that the ratio between the variance ω and parameter τ is equal to the variance of the portfolio in the view, $p'\Sigma p$." Usually one holds more than one view, hence the procedure should be generalized to multiple views. For multiple views this would imply that the

matrix Ω/τ should be calibrated such that it equals $P\Sigma P'$. It is assumed (assumption (A3)) that the views of the investor are independent, therefore the matrix Ω needs to be diagonal. This can be accomplished by taking only the variances and deleting the covariances, thus $\Omega/\tau = \text{diagonal}(P\Sigma P')$.
The rationale for this relation probably stems from the assumption that $\mathrm{E}(\mathbf{r})$ is normally distributed with variance $\tau\Sigma$. If the views of the investors $\mathbf{q} = P\mathrm{E}(\mathbf{r})$ are seen as a transformation of the expected return vector ($\mathrm{E}(\mathbf{r})$), then it follows from property (2.8) that the variance of these views Ω is also a transformation of $\tau\Sigma$, and hence $\Omega = P\tau\Sigma P'$, so Ω/τ equals $P\Sigma P'$. He and Litterman, however do not perform this transformation completely. As Ω needs to be a diagonal matrix only the variance of the transformed matrix $P\Sigma P'$ is used.
There is a substantial advantage: the problem of specifying τ is solved and there is a clear method to obtain an estimate for this parameter and for the uncertainty matrix Ω.
However, it is strange to see the views of the investor as a simple transformation of the equilibrium returns. It would actually imply that the investor has no additional information, as it would be logical that one should then also see the vector $\mathbf{q}$ as a transformation of the equilibrium return and hence should be equal to $P\boldsymbol{\pi}$.
Additionally, this procedure to determine τ and Σ takes away much flexibility from the BL procedure, while one of the attractions was that an investment manager could indicate a sense of certainty in her view. Moreover the conceptual relationship of τ as a parameter to scale the matrix Σ is lost in this way. By specifying Ω/τ in this manner the relationship between the parameter τ and the equilibrium distribution is non-existent, while this was the reason to introduce the parameter in the first place. Finally, it is strange that the two parameters were first specified separately and it is now possible to combine them into one parameter.

4.4.2 Specification of the view matrix P

There are three points of view on specifying the matrix P, which are not completely distinct, but slightly overlap. The simplest method is to specify a view on a single asset, just as one would do in mean-variance optimization. The second method is to specify a view on a portfolio of assets. This method is less intuitive because it is difficult for a person to give an estimate for the expected return of a portfolio.
Finally, a popular method is to specify relative views, i.e. that one asset, or a set of assets will outperform another set of assets. Effectively this means specifying a zero-investment portfolio. This method is popular due to the intuitiveness of it: it feels natural to express that asset A will outperform asset B. However, the value of the elements of P could then become a matter of importance. Idzorek (2004) has written a series of articles on this

subject.
For example, a view could be that the expected return of asset A will outperform asset B by 5 %. The specification is done in the form $P\text{E}(\mathbf{r}) = \mathbf{q}+\boldsymbol{\epsilon}$. Black and Litterman (1991a) specify the P matrix in this example as $P = \begin{pmatrix} 1 & -1 & 0 \end{pmatrix}$. This example is fairly straightforward, but it becomes more complicated when one wants to specify views about more assets. Idzorek (2004) has a nice example of this, which also features in example 5. This example will be continued now.

Example 6 (continuation of example 5[7])**.** *The third view is most important to us:*
View 3: US large growth and US small growth asset will outperform US large value and US small value by 2%.
The view matrix can be specified as was done in example 5:

$$P = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & -0.5 & 0.5 & -0.5 & 0 & 0 \end{pmatrix} \tag{4.12}$$

Satchell and Scowcroft use an equal weighting scheme to specify the weights in P, as seen in the last row of the matrix of equation (4.12). Under this system, the weightings are proportional to 1 divided by the number of respective assets outperforming or underperforming. View 3 has two nominally underperforming assets, each of which receives a -0.5 weighting. View 3 also contains two nominally outperforming assets, each receiving a $+0.5$ weighting.
However, the view matrix can also be specified differently. The previous weighting scheme ignores the market capitalization, the price of the equity times the number of outstanding equity of the assets involved in the view. The market capitalizations of the US Large Growth and US Large Value asset classes are nine times the market capitalizations of US Small Growth and Small Value asset classes. However, the method of Satchell and Scowcroft affects their respective weights equally, causing large changes in the two smaller asset classes.
Idzorek (2004) prefers to use a market capitalization weighting scheme, that takes in consideration the market capitalization of the assets. He and Litterman (1999) probably also use such a scheme but they do not state this explicitly. More specifically, the relative weighting of each individual asset is proportional to the assets market capitalization divided by the total market capitalization of either the outperforming or underperforming assets of that particular view.

[7]Derived from Satchell and Scowcroft (2000).

Example 7 (continuation of 6). *The market capitalization of the assets can be found in table 4.2. The relative market capitalization weights of the*

	Market capitalization (billions)	Relative weight
Outperforming		
US Large Growth	5,174	90%
US Small Growth	575	10%
Underperforming		
US Large Value	5,174	90%
US Small Value	575	10%

Table 4.2: Market capitalization weights.

nominally outperforming assets are 0.9 for US Large Growth and 0.1 for US Small Growth, while the relative market capitalization weights of the nominally underperforming assets are -0.9 for US Large Value and -0.1 for US Small Value. These figures are used to create a new matrix P:

$$P = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.9 & -0.9 & 0.1 & -0.1 & 0 & 0 \end{pmatrix}.$$

4.4.3 Alternative models

Scowcroft and Sefton (2003), of the bank UBS, use a model that is very similar to the Black-Litterman model. However, they make some different assumptions, remove some of the parameters and most noticeably they do not estimate the expected return of an asset but the return. The investor no longer has to specify views on the expected return of assets, but this is changed to the difference between the return and the long run return (equilibrium returns in the terminology of Black and Litterman).

The first assumption is that next periods return ($\mathbf{r}_t$) is the sum of the long run equilibrium returns $\boldsymbol{\mu}$ and next periods stochastic return $\boldsymbol{\epsilon}_t$, i.e. $\mathbf{r}_t = \boldsymbol{\mu} + \boldsymbol{\epsilon}_t$. The stochastic return (in other settings often called an error term) is normally distributed with zero mean and covariance matrix Σ. Therefore, the next periods return $\mathbf{r}_t$ could also be said to be normally distributed: $\mathbf{r}_t \sim N(\boldsymbol{\mu}, \Sigma)$.

The investor has views on the return of a single asset or a portfolio assets. These portfolios are formed by the rows in the matrix $P \in \mathbb{R}^{k \times n}$. The next period investor's views ($\mathbf{f}_t$) are distributed around the final realized vector of returns with an error $\boldsymbol{\nu}_t$. The error is normally distributed with covariance matrix Ω and mean zero.

Thus, the investors views are given by: $\mathbf{f}_t = P(\mathbf{r}_t - \boldsymbol{\mu}) + \boldsymbol{\nu}_t$, where $\boldsymbol{\nu} \sim N(\mathbf{0}, \Omega)$. It is assumed that the error vector in the investor's views are independent of the stochastic return, i.e. $\mathrm{E}(\boldsymbol{\nu}_t'\boldsymbol{\epsilon}_t) = 0$. The variance of the investor's views is a combination of the variance fo the return r and that of the error $\boldsymbol{\nu}_t$, thus: $\mathrm{var}(\mathbf{f}_t) = P\Sigma P' + \Omega$.
The investor does not forecast the absolute return of a portfolios of assets, but rather she can forecast the excess return over the long run equilibrium vector $(\mathbf{r}_t - \boldsymbol{\mu})$ of the portfolio. More precisely, $\mathbf{r}_t - \boldsymbol{\mu}$ is actually the stochastic return vector $(\boldsymbol{\epsilon}_t)$.

The combined return The two sources of information can be combined by the mixed estimation method of Theil, or the Bayesian statistics method. Either manner lead to the following result:

$$\begin{aligned} \mathrm{E}(\mathbf{r}_t|\mathbf{f}_t) - \boldsymbol{\mu} &= (\Sigma^{-1} + P'\Omega^{-1}P)^{-1}P\Omega\mathbf{f}_t \\ \mathrm{var}(\mathbf{r}_t|\mathbf{f}_t) &= (\Sigma^{-1} + P'\Omega^{-1}P)^{-1} \end{aligned}$$

Comparison of BL and UBS The Black-Litterman model and the UBS model differ mainly in their approach to forecasting the return vector. They both agree that it is difficult to forecast such a vector. The solution of Black and Litterman is to forecast an expected return vector ($\mathrm{E}(\mathbf{r})$) instead of the absolute return. The problem with this approach that one needs a distribution of the expected return vector and therefore they introduce the parameter τ. The parameter τ greatly complicates the problem as it is unintuitive and there is no guidance on setting it.
Satchell and Sefton take a different approach and solve the problem by requiring an estimate of the return relative to the long run equilibrium returns. The main advantage of the UBS model, is that it has no parameters that are difficult to understand. It might be interesting to research if it is really easy for a person to forecast the return in the method they specify. Furthermore, it would be interesting to compare the performance of the models in an empirical study.

4.5 Advantages and disadvantages

The Black-Litterman approach to asset allocation, is a step forward. The investor is now no longer obliged to specify views on all assets, but can specify views only if she holds one. The manner of specifying view is more natural, because it is possible to represent the views in a relative manner. Also it is possible to express how certain one is about the view. These additions make the model more versatile. However, the extension of the model also opens up new problems.
The model contains a few parameters that are difficult to specify, these are

the parameters τ and δ. Several ideas exist about how to specify them, but these contradict each other. The recommended values of the parameter τ varies between 0 and 1. According to Black and Litterman τ should be close to zero, according to Satchell and Scowcroft τ should nearly always be set to 1. The parameter δ can be seen as fraction of the excess return of the (world) market portfolio and its variance or as a (world) risk aversion constant.

The method combines the best of both the quantitative as well as the qualitative world. Two sources of information are combined in the BL-formula that were previously considered to be disconnected.

The Black-Litterman approach does not solve all problems of mean-variance optimization. The optimizer is still the driving force behind the model. Although it has new estimates of the expected return and variance and should therefore be better behaved, inherently nothing has been changed about the optimization procedure.

A major drawback of the model is that it assumes that the returns have a normal distribution, which is often not the case. It would be an improvement to develop the Black-Litterman formula under the assumption of other distributions that better reflect the return distribution. This has been explored in a discussion paper by Giacometti et al. (2006).

Chapter 5

Empirical Research

5.1 Introduction

The Black-Litterman is at the moment a very popular asset allocation model, many articles have been written on the subject. In this empirical study we will investigate if this popularity is justified by comparing the performance of the BL-model to mean-variance optimization.
In order to compare these models, we will first perform a sensitivity analysis of the parameters in the BL-model. The parameters and inputs to the model, to a large extent determine the output of the model. Therefore, it is important to investigate the influence of the parameters on the portfolios formed by the model. When the impact of the parameters is understood it will be possible to calibrate the value of the parameters. The Black-Litterman formula has the following inputs: the parameter τ, the covariance matrix Σ, the view matrix P, the view return vector $\mathbf{q}$, the uncertainty in view covariance matrix Ω and the equilibrium returns $\boldsymbol{\pi} = \delta\Sigma\mathbf{w}_m$, where δ has various definitions and $\mathbf{w}_m$ is the benchmark portfolio or the market portfolio. To recall, the BL-formula to the compute the new expected returns is:

$$\widehat{\mathrm{E}(\mathbf{r})} = [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1}[(\tau\Sigma)^{-1}\boldsymbol{\pi} + P'\Omega^{-1}\mathbf{q}].$$

The inputs Σ, P and $\mathbf{w}_m$ will be determined in a straightforward manner, as will be explained in section 5.2. The parameters τ, and δ are open to calibration. Finally, the parameter $\mathbf{q}$, the expected returns of the view portfolios of the investor will be computed via a new method. They will not be estimated by the investor, but they will be forecasted from a regression analysis on factors that describe the economic climate. The covariance matrix Ω can in this context be computed from the error in the regression analysis. We will also explore the alternative of taking $\Omega/\tau = P\Sigma P'$.
The performance of the BL-model and the MV-model will subsequently be compared in a three asset universe, where the three assets will be popular

zero-investment strategies. An investment strategy is a set of guidelines, behaviors or procedures, designed to maximize the expected return of an investment portfolio.
The outline of the chapter is as follows, first we will explain the three asset universe further. Next, we will move to the methods with which our topic is investigated, subsequently the data will be evaluated and finally the results of the empirical study will be discussed.

5.1.1 Investment universe: Three zero-investment strategies

The investment universe will consist of three zero-investment strategies. Each strategy has a certain characteristic on which the equity is categorized, from this division into groups it is possible to determine zero-investment strategies and to form portfolios based on these strategies. Various reasons are offered on why these strategies provide added return, these explanations will be discussed.

HML The first strategy divides equity in three categories: value, neutral and growth or glamor equity. Value equity, equity that is undervalued on the market in comparison to its accounting value, are shown to have a higher than average return than growth equity (Fama and French, 1992).
Stattman (1980) and Rosenberg, Reid, and Lanstein (1985) are among the first to examine the performance of value equity. They find that average returns on U.S. equity are positively related to the ratio of a firm's *book value of common equity* (BE) to its *market value* (ME)(often shortened to book-to-market value or BE/ME). The book value of the firms equity is the value of the equity in the accounting books of the firm at the date of the last balance sheet. It can for instance be the value determined by the accountant at January 2006. The market value of a company is the value of all the outstanding ordinary equity on the stock market, i.e. the number of equity outstanding times the price of the equity.
A high book-to-market ratio means that the value of the firm in the books is higher than the value on the stock market. Thus, it could be said that a firm is undervalued by the stock market. The derived strategy is known as *high-minus-low* (*HML*), which entails to go long in high book-to-market value equity and short in low book-to-market equity.

Reason Different reasons are proposed on the profitability of the HML strategy, not one has been reached consensus upon. The main argument of Fama and French (1992) centers around the efficient market hypothesis. The efficient market hypothesis says markets incorporate instantaneously all information on the assets that are traded on the market. Thus, if a strategy has a higher return, there must also be some additional risk attached to that asset.

Chan and Lakonishok (2004) however found no such higher risk after measuring risk with a variety of indicators. Therefore, they conclude that it is unlikely that the superior performance of value equity can be attributed to a higher risk.
Furthermore, they state that the sharp rise and decline in the nineties of technology and other growth equity calls into question the argument that growth equity is less risky than value equity. They do not find evidence for additional risk, while using a variety of indicators.
According to Chan and Lakonishok (2004) "the value premium can be tied to ingrained patterns of investor behavior or the incentives of professional investment managers. In particular, in the markets of the nineties (as in numerous past episodes in financial history), investors extrapolated from the past and became excessively excited about promising new technologies, like Internet, telecommunications and ICT. They overbid the prices of apparent "growth" stocks while the prices of value stocks dropped far below their value based on fundamentals. Because these behavioral traits will probably continue to exist in the future, patient investing in value stocks is likely to remain a rewarding long-term investment strategy."

SMB The second strategy differentiates among stock according to size. Banz (1981) shows that the size of a firm is an important indicator of future performance. At various holding horizons are small firms shown to outperform large firms, in the value of the risk-adjusted return.
Banz (1981) found that the size of a company, as measured by the market value of the common outstanding equity, is an indicator of the future performance. He shows that in the period from 1936 to 1975 "the common stock of small firms had, on average, higher risk-adjusted returns than the common stock of large firms", which he calls the 'size effect'.
The strategy entails to go long in small size equity and short in large size equity. Fama and French (1993) describe this strategy as *small-minus-big (SMB)*.

Reason Banz (1981) already warned that the size effect is not stable through time and for the lack of theoretical foundation for the effect. It is not certain to Banz whether size is the factor that causes the effect or if it is only correlated with the real factor and that size serves as a proxy for this other factor.
He gives a possible explanation for the effect: if there is little information on the equity, then few investors will hold these. Furthermore, Banz has shown in his PhD thesis that "securities sought by only a subset have higher risk-adjusted returns than those considered by all investors." Therefore the higher returns in small stock could be related to a higher risk.
Baker and Wurgler (2006) relate the sentiment of investors to the perfor-

mance of stocks. They measured sentiment by a regression analysis on certain factors. They find a relationship between sentiment and the return of small size equity. They find that returns are relatively high for small size equity when sentiment is low. The smallest stock had a monthly return of 2.94% as opposed to a return of 0.92% for the largest equity.
Finally, several researchers, among those Keim (1983), find that the size effect is mainly a January effect. "Nearly fifty percent of the average magnitude of the size anomaly over the period 1963-1979 is due to January abnormal returns. Further, more than fifty percent of the January premium is attributable to large abnormal returns during the first week of trading in the year, particularly on the first trading day."
Jegadeesh and Titman (2001) observe that the size effect is not observed after the sample period of 1965 to 1981. In the 1982 to 1998 sample period the Fama-French size factor is only -0.18% on average per month with a t-statistic of -1.01, as opposed to 0.53% per month with a t-statistic of 2.34 in the 1965 to 1981 period.

WML Finally, the momentum of the stock, the performance of the past three to eighteen months, seems to have some predictive power over future returns. It is found that past winners (named so by Jegadeesh and Titman (1993)), equity with an above average return, tend to be winners for some time longer, alternatively past losers also tend to be losers for some time longer. The portfolio that is derived from the momentum strategy is called *winners-minus-losers*(*WML*), or up-minus-down (UMD) by Fama and French, it is a portfolio that goes long in winners and short in losers.
Jegadeesh and Titman (1993) show that buying past winners and selling past losers realizes significant abnormal returns over the 1965 to 1989 period. They examine the returns of zero-cost winners minus losers portfolio in the 36 months following the portfolios formation date. In the short to medium-term the winners minus losers portfolios realize significant returns. However, the longer-term performances of these past winners and losers reveal that half of their excess returns in the year following the portfolio formation date dissipate within the following two years.

Reason Jegadeesh and Titman (2001) reassess momentum strategies since their original 1993 study to find that their work still holds in the out of sample 1993-1998 period. In their new study they try to understand why momentum strategies are profitable and to this end they evaluate various explanations.
There are mainly two opposing strands of thought on the workings of momentum strategies. The behaviorist explain momentum as being due to a bias most investors have in interpreting information. The bias is explained further by Scowcroft and Sefton(2005): "Investors have a tendency to at-

tribute positive outcomes to skill and negative outcomes to bad luck. Following a decision to buy, investors exhibiting this bias are more likely to later buy more of the stock it they receive further good news than they are likely to sell if they receive bad news. This asymmetry causes prices to rise too far in the short term and correct themselves later." Realist try to seek an explanation in rational models and suggest that the profitability is due to a compensation for risk factors.
Jegadeesh and Titman (2001) compared a behaviorist model and a rational model from Conrad and Kaul (1998). The models make 'diametrically' opposed predictions about the returns of past winners and losers over the period following formation. Jegadeesh and Titman examined the returns of the winner and loser equity in the 60 months following formation date. The Jegadeesh and Titman momentum portfolios have a positive return in the first twelve months after formation date and a negative return in the 13 to 60 months after formation date. This return reversal effect is consistent with behavioral theories but, inconsistent with the Conrad and Kaul hypothesis. Therefore, they conclude that the behavioral models appear to describe the effect better, however caution is necessary. They find strong evidence of return reversal for small firms, but the evidence is 'somewhat weaker' for larger firms. Additionally, there is strong evidence of the return reversal in the period 1965-1981, but this is substantially weaker in the period 1982-1998. This finding is noteworthy because there are no distinguishable differences between the two periods in either magnitude or significance of the momentum profits. Therefore, the endorsement of behavioral finance theory to explain the momentum effect has to be done cautiously.

Summary The investment universe consists of three assets, a HML portfolio, a SMB portfolio and a WML portfolio. The Black-Litterman portfolio will be a portfolio that is a combination of these three portfolio, in effect it is a superportfolio.

5.2 Methods

The Black-Litterman model has been implemented in Scilab, an open source matrix computation software package. The source code of the program used in this analysis can be found in Appendix A.
The research into the effectiveness of the BL-model is performed by compiling a new portfolio every month from January 1997 to December 2006 one portfolio with the BL-model and one with mean-variance optimization. The period from January 1978 to December 1996 will be used to calibrate the parameters of the BL-model. Each month the return of the portfolios will be computed, which allows the comparison of the models via various measures and to graphically display the return of the portfolios in time.

The inputs to the model are an important factor and therefore they will be discussed one by one. Also the manner in which the investors views are obtained, will be explained further.

Choice of benchmark The choice of benchmark to a large extend determines the performance of the BL-model, the portfolios formed by the BL-model vary around the benchmark. Therefore, it is important to choose a benchmark that has a reasonable Sharpe ratio under all circumstance. A Sharpe ratio starting at 0.7 is found to be reasonable.
The three strategies perform well under different circumstances and the circumstances under which they perform well are not clearly defined. Therefore the benchmark consists equally of all three portfolios. This should allow a reasonable Sharpe ratio under all circumstances, thus $\mathbf{w}_m = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$.

Estimate covariance from 'RiskMetrics' The covariance matrix will be estimated from daily return data of the three assets. The method of 'RiskMetrics' (2001) is used to estimate the covariance matrix. Every month a new matrix will be constructed. It is stated that recent data is more important for covariance modeling and therefore a decay factor is introduced, wherewith recent data has a larger weight in the covariance model than data farther in the past. The covariance matrix at time t (Σ_t) is computed as follows.

$$\Sigma_t = \frac{1-\lambda}{1-\lambda^n} \sum_{k=1}^{n} \lambda^{k-1} (\mathbf{r}_{t-k} - \bar{r}) \, (\mathbf{r}_{t-k} - \bar{r})'$$

$\mathbf{r}_{t-k}$ is the vector of return at time $t-k$

$\bar{r}$ is the average return over the time period $t-1$ to $t-n$

RiskMetrics advices to take the decay factor λ equal to 0.97 for estimating a monthly covariance matrix. The number of days incorporated in the estimate is to a certain extend not very relevant. The decay factor lessens the importance of later dates, for example $\lambda = 0.97$ and a forecast that uses 227 days results in a $\lambda^{n-1} = 0.97^{226} \approx 0.10\%$, thus the 227$^{\text{th}}$ day has a very small weight in the sum. Therefore, we choose to use one business year of information in the forecast, i.e. 256 days.

Delta The definitions of δ vary, as discussed in section 4.2.2. The coefficient δ can be taken to equal the expected excess return of the market portfolio divided by its variance. Then, $\delta = \frac{\mathrm{E}(\mathbf{r}_m) - r_f}{\mathrm{var}(\mathbf{w}_m)}$, where the excess return on the market is obtained from the data library of French and the variance of the market portfolio is computed from the excess return minus the risk-free rate.

Alternatively, the coefficient δ can be seen to represent a risk aversion coefficient, in this form it can be set to a certain level. Both alternatives to setting δ will be explored.

P The investor can specify views on single assets, portfolios of assets or the relative performance of two assets, see section 4.4.2. The forecasting procedure to produce views provides a forecast for every individual asset, therefore the view matrix will be chosen accordingly equal to the identity matrix.

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Omega/tau The confidence in view matrix Ω can be chosen as suggested by He and Litterman (1999) to equal $\frac{\Omega}{\tau} = \text{diagonal}\,(P\Sigma P')$, see section 4.4.1. This removes the difficulty of stating a weight on the view of the investor and of calibrating the parameter τ.
Alternatively, the matrix Ω can be chosen to represent the error or uncertainty that results from the forecasting procedure. In that case τ has to be specified separately, τ is said to vary between zero and one.

q The vector of expected returns of the investor ($\mathbf{q}$) is a transformation of the expected returns that follows from the forecasting procedure ($\widehat{\mathrm{E}(\mathbf{r})}_{\mathbf{q}}$) and the view matrix P,

$$\mathbf{q} = P\widehat{\mathrm{E}(\mathbf{r})}_{\mathbf{q}} \tag{5.1}$$

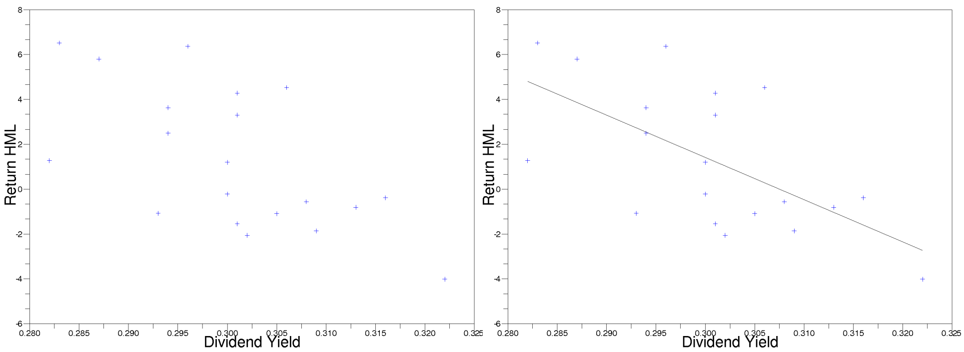
in this case equals P the identity matrix and therefore no transformation takes place, i.e. $\mathbf{q} = \widehat{\mathrm{E}(\mathbf{r})}_{\mathbf{q}}$.

5.2.1 Economic factors

It is commonly believed that asset prices react sensitively to economic news. Chen et al. (1986) state that "daily experience seems to support the view that individual asset prices are influenced by a wide variety of unanticipated events and that some events have a more pervasive effect on asset prices than do others."
Chen et al. (1986) have investigated which economic factors have an influence on the return of assets. They find that the most important factors are industrial production, changes in the risk premium, twists in the yield curve, and, somewhat more weakly, measures of unanticipated inflation. These factors and some others that describe the economic climate, will be used to forecast the asset return. The explanation of these factors will be omitted,

as it caries to far for this thesis. Explanation of these factors and others, can be found in the works of Chen et al. (1986), Chen (1991), Pesaran and Timmermann (1995) and Fama and French (1989).
The general consensus among these researchers is that asset return can to some extend be predicted by a regression analysis on economic factors. The idea of regression analysis and in the manner in which it can be used in our context will be discussed in the next paragraph.

Regression analysis to forecast q We assume there is a relationship in time between the return of assets and certain economic factors, but are uncertain about the nature of the relationship. To examine this relationship one could plot the data of the asset return and for example dividend yield in one graph, see figure 5.1. Every point in the graph is the asset return and the dividend yield at a time instance, where the return is at time t and the dividend yield is at time $t-1$.



(a) Data points. (b) Data points with regression line.

Figure 5.1: The relationship between the return of a HML portfolio and dividend yield.

Examining the graph, one expects a linear relationship between dividend yield and return. This is our next assumption, there is a linear relationship between the asset return (y) and the economic factor (x) of the form:

$$y_t = a + bx_{t-1} + \epsilon_t.$$

The dependent variable y_t, in this case the asset return, will be explained from the dividend yield x_{t-1}, also known as the explanatory variable. The asset return at time t equals some intercept a plus the dividend yield at time $t-1$ times some unknown coefficient b plus some unknown error ϵ. Furthermore, it is assumed that this unobservable error ϵ_t is normally distributed $\epsilon_t \sim N(0, \sigma^2)$. Typically, there is more than one observation of y and x

to make the analysis on, therefore the formula is often written in matrix notation, the formula then becomes:

$$\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}.$$

$$\mathbf{y} = \begin{pmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-n+1} \end{pmatrix}, X = \begin{pmatrix} 1 & x_{t-1} \\ 1 & x_{t-2} \\ \vdots & \vdots \\ 1 & x_{t-n} \end{pmatrix}, \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}, \boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_t \\ \epsilon_{t-1} \\ \vdots \\ \epsilon_{t-n+1} \end{pmatrix}$$

To estimate this linear relationship it is necessary to estimate the coefficients a and b in the vector $\boldsymbol{\beta}$. This can be accomplished by choosing the coefficients $\boldsymbol{\beta}$ in such a way, that they minimize the distance between the observed data $(\mathbf{y}, X)$ and the lines $\mathbf{y} = X\boldsymbol{\beta}$. The coefficients that minimize the distance form $\hat{\boldsymbol{\beta}}$, the estimate of $\boldsymbol{\beta}$.

$$\hat{\boldsymbol{\beta}} = (X'X)^{-1}X'\mathbf{y} \tag{5.2}$$

The proof of this formula can be found in any good book on econometrics, for example Theil (1971). The analysis can be expanded to incorporate more explanatory variables, in our case we would also like to incorporate among others the term spread. The regression formula with m explanatory variables then becomes $y_t = \beta_1 + x_{t-1,2}\beta_2 + \ldots + x_{t-1,m}\beta_m + \epsilon_t$. This can be easily accomplished, by expanding the matrix of explanatory variables X to encompass m columns of explanatory variables and by expanding the coefficient vector of $\boldsymbol{\beta}$ to a m column vector.
The relationship in time between the explanatory variables and the dependent variable can be used to make a prediction of the return at time $t+1$. At that time the formula is $y_{t+1} = \mathbf{x}_t'\boldsymbol{\beta} + \epsilon_{t+1}$, at time t are the explanatory variables in $\mathbf{x}_t$ known and the coefficients $\boldsymbol{\beta}$ can be estimated from equation (5.2). Thus the forecasted return for time $t+1$ on basis of the economic factors is given by:

$$\hat{y}_{t+1} = \mathbf{x}_t'\hat{\boldsymbol{\beta}}. \tag{5.3}$$

So far we have not discussed the error term ($\boldsymbol{\epsilon}_t$) in much detail, it is assumed to have a normal probability distribution with zero mean and covariance matrix $\sigma^2 I$. The value of σ^2 is unknown but, can be estimated to be:

$$\hat{\sigma}^2 = \frac{\mathbf{y}'(I - X(X'X)^{-1}X')\,\mathbf{y}}{n-m}, \tag{5.4}$$

the proof of this formula can be found in Theil (1971).
Knowing the error in $\boldsymbol{\epsilon}_t$ it is now possible to comment on the error in our forecast $\hat{y}_{t+1}$, however we will first need to discuss the error in the coefficient estimate $\hat{\boldsymbol{\beta}}$ before the error in $\hat{y}_{t+1}$ can be discussed.

Proposition 3. *The uncertainty in the estimated coefficient vector* $\hat{\boldsymbol{\beta}}$ *is:*

$$\mathrm{var}(\hat{\boldsymbol{\beta}}) = \sigma^2 (X'X)^{-1}. \tag{5.5}$$

Proof. The proof of this proposition uses that $\hat{\boldsymbol{\beta}} = (X'X)^{-1}X'\mathbf{y}$, furthermore X is a non-stochastic matrix and therefore the variance of $\hat{\boldsymbol{\beta}}$ depends on the variance of $\mathbf{y}$. Finally, a little reworking of the formulas is needed to obtain the end result.

$$\begin{aligned}\mathrm{var}(\hat{\boldsymbol{\beta}}) &=^1 \mathrm{var}((X'X)^{-1}X'\mathbf{y}) =^2 (X'X)^{-1}X'\mathrm{var}(\mathbf{y})((X'X)^{-1}X')' \\ &=^3 (X'X)^{-1}X'\mathrm{var}(\mathbf{y})X(X'X)^{-1} =^4 (X'X)^{-1}X'\sigma^2 IX(X'X)^{-1} \\ &= \sigma^2(X'X)^{-1}(X'X)(X'X)^{-1} \\ &= \sigma^2(X'X)^{-1}\end{aligned}$$

$=^1$ $\hat{\boldsymbol{\beta}} = (X'X)^{-1}X'\mathbf{y}$ *from equation* (5.2).
$=^2$ X *and* $\mathbf{y}$ *are non-stochastic and therefore equation* (2.8) *applies:* $\mathrm{var}(PX) = P\mathrm{var}(X)P'$.
$=^3$ $(X'X)'^{-1} = (X'X)^{-1}$
$=^4$ $\mathrm{var}(\mathbf{y}) = \mathrm{var}(\boldsymbol{\epsilon}) = \sigma^2 I$

□

Now that the variance of the coefficients $\hat{\boldsymbol{\beta}}$ is known, it is possible to compute the variance of the forecast $\hat{y}_{t+1}$.

Proposition 4. *The uncertainty in the forecast* $\hat{y}_{t+1}$ *is:*

$$\mathrm{var}(\hat{y}_{t+1}) = \sigma^2 \mathbf{x}_t'(X'X)^{-1}\mathbf{x}_t.$$

The variance σ^2 *can then be estimated from formula* (5.4)*, this makes the estimate:*

$$\frac{\mathbf{y}'(I - X(X'X)^{-1}X')\,\mathbf{y}}{n-m}\mathbf{x}_t'(X'X)^{-1}\mathbf{x}_t. \tag{5.6}$$

Proof. The proof of this proposition follows along the same lines as the proof of previous proposition.

$$\begin{aligned}\mathrm{var}(\hat{y}_{t+1}) &=^1 \mathrm{var}(\mathbf{x}_t'\hat{\boldsymbol{\beta}}) =^2 \mathbf{x}_t'\mathrm{var}(\hat{\boldsymbol{\beta}})\mathbf{x}_t \\ &=^3 \mathbf{x}_t'\sigma^2(X'X)^{-1}\mathbf{x}_t = \sigma^2\mathbf{x}_t'(X'X)^{-1}\mathbf{x}_t \\ \text{and thus } &{}^4\frac{\mathbf{y}'(I - X(X'X)^{-1}X')\,\mathbf{y}}{n-m}\mathbf{x}_t'(X'X)^{-1}\mathbf{x}_t\end{aligned}$$

$=^1$ $y_{t+1} = \mathbf{x}_t'\hat{\boldsymbol{\beta}}$ *from equation* (5.3).
$=^2$ $\mathbf{x}_t$*is non-stochastic and therefore equation* (2.8) *applies:* $\mathrm{var}(PX) = P\mathrm{var}(X)P'$.
$=^3$ $\mathrm{var}(\hat{\boldsymbol{\beta}}) = \sigma^2(X'X)$ *see equation* (5.5).
$=^4$ $\hat{\sigma}^2 = \frac{\mathbf{y}'(I - X(X'X)^{-1}X')\mathbf{y}}{n-k}$, *see equation* (5.4).

□

Summary The forecasting procedure will be performed for all three assets and this will result in the view vector of the investor ($\mathbf{q}$). The elegance of this procedure is that it provides a natural choice for the covariance matrix Ω. The matrix Ω should convey the variance of the investors views. Ordinarily this is estimated by the investor, but it is difficult for a person to give an estimate of a variance. With this procedure the variance follows from the error in the forecasting procedure as given by equation (5.6).

5.3 Data

The empirical study requires a considerable amount of data, for example to compute the covariance matrix and to predict the return of the assets. The data that is needed for this empirical study will be taken from various sources.

5.3.1 Data of French is used for the assets

Fama and French have written some articles on size and value strategies (see for example Fama and French (1989, 1992, 1995)), as part of that research Kenneth French maintains a data library on his website where HML, SMB and WML portfolios are formed with various holding horizons.
The monthly portfolios, available from July 1926 to the present, will be used to compute in each month the realized return of the Black-Litterman and the mean-variance portfolio. The monthly data will also be used to perform the regression analysis on the economic factors and thus forecast the asset return.
Finally, the daily data on asset returns of French will be used to estimate the covariance matrix with the RiskMetrics procedure, the data is available from July 1, 1963.

Construction of the strategy portfolios French compiles portfolios since 1926 based on the three strategies SMB, HML and WML. These portfolios will be used as assets in the Black-Litterman optimization process to construct a superportfolio.
The construction of the three portfolios follow a similar pattern.[1] French constructs the portfolios by focusing on six portfolios, a simple sort divides the firms in two groups on size and three groups on value, the intersection of these form six portfolios.
The equity in the portfolio are all listed on either the New York Stock Exchange (NYSE), the American Stock Exchange (AMEX) or the NASDAQ. The first division on size is made by simply listing the firms on the value of

[1]The explanation of the formation is taken from the website of Kenneth French: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

their market equity, the median of the NYSE is used to divide the equity in two groups, a group of small size equity and a group of large size equity.
The second division is made by ranking the stocks according to the book-to-market value, the breakpoint of the bottom group is the 30%-mark of the NYSE, the middle group consists of the next 40 % and the top group consists of the last 30 %. The bottom group is considered *growth equity* and the top group *value equity*.
The intersection of these groups form six portfolios, see table 5.1.

	Median ME	
	Small value	Big value
P70 BE/ME		
	Small neutral	Big neutral
P30 BE/ME		
	Small growth	Big growth

P70 is the 70^{th} percentile, P30 is the 30^{th} percentile

Table 5.1: Six portfolios formed on size and book-to-market value.

The small value portfolio, for example consists of the equity that fall in the small size category and are considered value equity, thus are in top BE/ME segment. This portfolio is also denoted as the S/V portfolio, small in size and high in value. The big growth portfolio is hence called the B/G portfolio.
In total there are six portfolios, three small portfolio S/V, S/M and S/G also there are three big portfolios B/V, B/M and B/G. These six portfolios are used to construct the size strategy (SMB) and the value strategy (HML). The momentum strategy (WML) is constructed in a similar fashion, but instead of a sorting on book-to-market value there is a sorting on past return.

Compiling the HML portfolio The high-minus-low portfolio is an average of a long position in the small value portfolio and the big value portfolio and a short position in the small growth and big growth portfolios: $\mathbf{w}_{HML} = \frac{1}{2}(SV + BV) - \frac{1}{2}(SG + BG) = \frac{1}{2}\begin{pmatrix}1 & 1 & -1 & -1\end{pmatrix}$.

Compiling SMB SMB is the difference between the returns on small size and big size equity portfolios with about the same weighted-average book-to-market equity: $\mathbf{w}_{SMB} = \frac{1}{3}(SV + SN + SG) - \frac{1}{3}(BV + BN + BG) = \frac{1}{3}\begin{pmatrix}1 & 1 & 1 & -1-1 & -1\end{pmatrix}$.

Compiling WML The momentum strategy is compiled from six portfolios sorted on size and prior two to twelve month returns. The portfolios, are the intersections of two portfolios formed on size and three portfolios formed on prior (two to twelve month) return. The monthly size breakpoint is the median NYSE market equity, the monthly prior return breakpoints are the 30th and 70th NYSE percentiles.
WML is the average return on the two high prior return portfolios minus the average return on the two low prior return portfolios: $\mathbf{w}_{WML} = \frac{1}{2}(SH + BH) - \frac{1}{2}(SL + BL) = \frac{1}{2}\begin{pmatrix}1 & 1 & -1 & -1\end{pmatrix}$.

5.3.2 Economic factor model data

The data for the economic factor model is acquired from various sources, all data is available in the period January 1978 until December 2006.

UTS One of the economic factors should capture the shape of the term structure, which can be measured by the difference between the long term government bonds (10-year T-bond rate) and short term government bonds (3-month T-bill rate), this factor is called UTS. The data is taken from the "Real Time Dataset for Macroeconomists" of the Federal Reserve in Philadelphia.

MP The change in monthly industrial production (MP) is defined as: $\mathrm{MP}(t) = \log\left(\frac{\mathrm{IP}(t)}{\mathrm{IP}(t-1)}\right)$, where $\mathrm{IP}(t)$ is the industrial production at time t. The industrial production figures are obtained from the "Real Time Dataset for Macroeconomists" of the Federal Reserve in Philadelphia.

UI The unanticipated inflation (UI) is defined by Chen (1986) as the difference between the inflation and the expected value of the inflation, where the inflation is defined as the log of the change in the consumer price index (CPI) and the expected inflation is taken from a study by Fama and Gibson (1984). The definition of unanticipated inflation is adapted to the change in inflation from the previous month to the present month: $\mathrm{UI}(t) = \log\left(\frac{\mathrm{CPI}(t)}{\mathrm{CPI}(t-1)}\right) - \log\left(\frac{\mathrm{CPI}(t-1)}{\mathrm{CPI}(t-2)}\right)$. The figures on consumer price index are obtained from Shiller.[2]

DEF The default spread (DEF), has been defined by Fama and French (1989) as the difference between the yield on a market portfolio of 100 corporate bonds and the AAA yield. However, they claim the results are robust over changes in definition and the yield of BAA bonds could be substituted for the market portfolio.

[2]http://www.econ.yale.edu/˜shiller/

The difference between BAA bond yield and AAA bond yield is computed on basis of LEHMAN US AGGREGATE data.

DP The dividend yield (DP) is defined by Fama and French (1989) as the sum of the dividends on the portfolio for the year preceding time t and divided by the value at time t. This data is not available for the period 1974-2007 it is only available until 2004, hence it is approximated by the dividend yield of 'S&P 500 COMPOSITE - DIVIDEND YIELD' which is available until 2006. The two data sets are perfectly correlated for the period 1966-2004, therefore it is assumed that this change will not give any problems.

5.4 Results

The results of the investigation into the performance of BL-optimization will be presented in this chapter and compared to the performance of MV-optimization. Before, the BL-model can be compared to MV-optimization we will determine the values of the input parameters Ω, τ and δ. These values will be determined by performing a sensitivity analysis and by calibrating the model. The calibration will be performed over the period January 1978 to December 1996, subsequently will the models be compared over the period January 1997 to December 2006.
Previous to the sensitivity analysis we will look into the uncertainty matrix Ω.There are two options to specify the uncertainty matrix Ω, both will be compared and the one with the best result, i.e. reasonable values of Ω and a good Sharp ratio of the resulting portfolio, will be used in the remainder of the study.
Next we will move to the illusive parameter τ, we will will investigate how τ influences the weight distribution of the portfolio. When that relation is understood, we can calibrate the parameter to improve the performance of the BL-portfolio. Finally, we will look into the sensitivity of the parameter δ and calibrate it in the same manner as the parameter τ.
After the calibration of the parameters for the BL-model it will be possible to perform the empirical study. The BL-model will compile portfolios from the BL-optimized returns and the covariance matrix Σ. The MV-portfolios will use the views of the investor($\mathbf{q}$), that are undampened by the equilibrium returns, and the corresponding covariance matrix Ω to compute the portfolio. The models will be compared under a full investment constraint, that is that the weights in the portfolios have to sum to one and thus that all portfolios are of equal size.
The portfolios will be compared over the ten year period from January 1997 to December 2006. Every month portfolios will be formed via both methods. The performance will, among others, be measured via the Sharpe ratio,

the mean return and the standard deviation of return over the investment period, all measures are annualized.

5.4.1 The parameters of the BL-model

The parameters of the BL-model will be varied one by one, to determine the sensitivity and the optimal value. While varying one parameter we will keep the others constant to a default value. The default value of the parameter δ will be 4, δ could be seen as a risk aversion parameter in that context this is a reasonable value. The parameter τ has two diametrically opposed interpretations, Black and Litterman (1991a) say its near to zero, Satchell and Scowcroft (2000) on the other hand state that it is often set to one. Therefore, we take a position in the middle, the default setting is $\tau = 0.5$. The uncertainty of views matrix is most often specified in the manner of He and Litterman (1999) they take $\Omega/\tau = \text{diagonal}(P\Sigma P')$, this therefore is our starting point. When Ω/τ is specified in this manner it is no longer necessary to specify τ separately. The purpose of the investigation is to one by one determine the sensitivity of the parameters and subsequently to calibrate the parameter to the optimal value.
The model will be calibrated such that the parameter values yield a model with a reasonable Sharpe ratio, but also such that it mixes both sources of information and such that the month-on-month turn-over in the portfolio becomes not to large. The turn-over in the portfolio is defined as the sum of the variations in the weights between two subsequent months, in formulas this can be written as turn-over $= \sum_{i=1}^{n} |\mathbf{w}_{i,t} - \mathbf{w}_{i,t-1}|$. We will calibrate the model such that the average turn-over during the sample period is at least 5% and maximally 10%.

Omega and tau The first parameters that will be varied are the uncertainty on views matrix Ω and the parameter τ. He and Litterman (1999) suggest to specify the parameters together and take $\Omega/\tau = \text{diagonal}(P\Sigma P')$, this will be the first option that will be investigated. For the moment we take the risk aversion parameter δ equal to 4, its default setting. These parameter choices lead to a series of portfolio. The average performance of the portfolios can be observed in table 5.2, all measures in the table, except for the turn-over, are annualized.

The annualized Sharpe ratio of the thus formed BL-portfolio is equal to 0.79, this is acceptable, but not very good. Additionally, a very large turn-over in the portfolio takes place, on average changes every month 64% of the portfolio, this would result in very high transaction costs. Including the transaction costs of 50 bp per percentage of change in the Sharpe ratio, would dramatically reduce the profitabiliy of the BL-model, it would result in a Sharpe ratio of 0.09.

	Sharpe	turn-over(%)	μ	σ
$\Omega/\tau = \text{diag}(P\Sigma P')$	0.79	64.05	12.74	16.06

Table 5.2: The performance of the BL-portfolio with $\Omega/\tau = \text{diag}(P\Sigma P')$ and $\delta = 4$.

Aside from the performance of the model, it is also important to note the value of the parameters, the thus formed values of Ω should represent reasonable choices.
The values of Ω/τ equal the diagonal elements of $P\Sigma P'$, in this experiment P equals the identity matrix, therefore $\Omega/\tau = \text{diag}(\Sigma)$. It is difficult to inspect all 227 individual matrices of Ω/τ, therefore we will examine the average size of the matrices. The size of the matrices will be measured with the Frobenius norm of the matrix $\|\Omega\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^n \omega_{ij}^2}$. The result can be found in table 5.3.

	mean	σ	max	min
$\|\frac{\Omega}{\tau}\|_F$	0.39	0.47	5.58	0.09

Table 5.3: The size of the values of $\Omega/\tau = \text{diagonal}(P\Sigma P')$.

The table shows that the average Frobenius norm is 0.39, with a standard deviation of 0.47. Small values of the norm of Ω mean that the variance in the views of the investor is also small and therefore that the investor is thus very certain about her views.
To discuss the reasonableness of these values we will try to compare them with other values of Ω found in literature. Idzorek (2004) gives an example with

$$\|\Omega\|_F = \|\frac{1}{1000}\begin{pmatrix} 0.709 & 0 & 0 \\ 0 & 0.141 & 0 \\ 0 & 0 & 0.866 \end{pmatrix}\|_F = 0.0011281,$$

and $\tau = 0.025$, this results in $\|\Omega/\tau\|_F = 0.04512$. In comparison, the size of our matrix is much larger, it is approximately 9 times as large.
Satchell and Scowcroft (2000) present an example with a single view, that has variance 0.05 and another example where the view has variance 0.025, in both examples is $\tau = 1$. Thus, $\|\Omega/\tau\|_F = 0.05$ and 0.025 respectively, this is larger than in the previous example but not as large as our findings.

Finally, Koch (2004) presents an example with two views, the first view has variance $(0.61\%)^2$ and the second view has variance $(0.91\%)^2$, the parameter τ in this example equals 0.3, this results in $\|\Omega/\tau\|_F = 0.03026$.
The figures we obtain look large in comparison to these figures, but the parameter τ makes a significant impact. The norm of $\|\Omega/\tau\|_F = \|\Omega\|_F/\tau$, this means that for values of τ close to zero the norm will grow rapidly. The influence of τ makes it difficult to compare the figures. For example, the matrix Ω/τ of Satchell and Scowcroft with a τ of 0.01 instead of 1, would lead to norm of 0.5; this number is much closer to our findings. Therefore, we cannot draw a conclusion on the appropriateness of the values of Ω.
The matrix Ω could also be specified differently, in a way that has a greater connection to the certainty in the views that are expressed. The views of the investor are computed from a linear regression forecasting analysis, this regression forecasting exercise also results in a variance of each forecasted return. The idea is now to take these variances as the diagonal elements of the uncertainty matrix Ω, the off-diagonal elements of Ω are taken zero as is required by Black and Litterman, see assumption A3. The matrix Ω can then be specified as

$$\begin{aligned}\omega_{ii} &=^1 \hat{\sigma}_{ii}^2 \mathbf{x}_t'(X'X)^{-1}\mathbf{x}_t \\ &= \frac{\mathbf{y}_i'(I - X(X'X)^{-1}X')\,\mathbf{y}_i}{n-k}\mathbf{x}_t'(X'X)^{-1}\mathbf{x}_t\end{aligned} \tag{5.7}$$

$=^1$ *The derivation of this formula can be found in paragraph 5.2.1.*

Having specified Ω separately from τ, leaves the determination of τ. For the moment it will equal its default setting 0.5. BL-optimization gives the following results, that can be seen in table 5.4.

	Sharpe	turn-over(%)	μ	σ
$\Omega = \hat{\sigma}^2\mathbf{x}_t'(X'X)^{-1}\mathbf{x}_t$	1.09	8.89	6.10	5.62

Table 5.4: The performance of the BL-portfolio with $\Omega = \hat{\sigma}^2\mathbf{x}_t'(X'X)^{-1}\mathbf{x}_t$, $\tau = 0.5$ and $\delta = 4$.

The Sharpe ratio of the thus formed BL-portfolio is 1.09, which is a great improvement over the previous portfolio, also the turn-over in the portfolio has greatly diminished to 8.89% of the portfolio.
The sample statistics of the values of Ω can be found in table 5.5. The values of Ω are significantly larger than in the previous example, this reflects a greater uncertainty in the views of the investor, also there is more variability

	mean	σ	max	min
$\|\|\Omega\|\|_F$	3.47	1.87	10.43	0.70

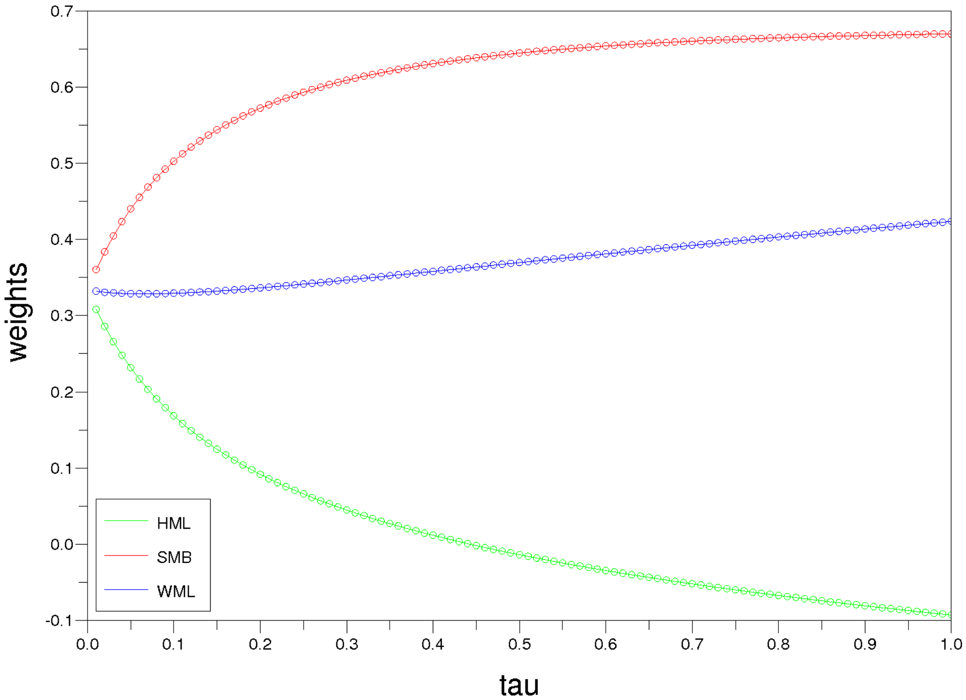
Table 5.5: The size of the values of $\Omega = \hat{\sigma}^2 \mathbf{x}_t'(X'X)^{-1}\mathbf{x}_t$.

in the values of Ω ($\sigma = 1.87$).
As the latter method for specifying Ω gives a better Sharpe ratios and less turn-over in the portfolio, we will adopt this method in the remainder of our empirical study. Further variation of the parameter τ should result in different portfolios with possibly an even higher Sharpe ratio.

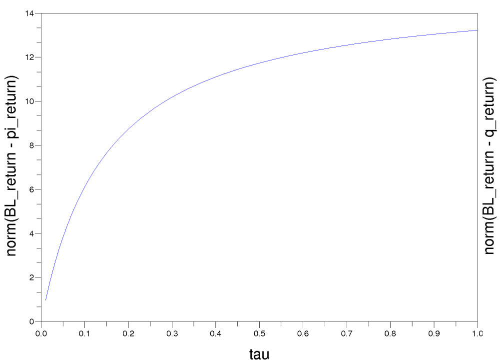
tau The parameter τ could be the most mysterious parameter of the BL-model, originally it is used to specify the relation between the distribution of the asset returns and the distribution of the mean of the asset returns. Black and Litterman assume that the variance in the mean of the return is smaller than the variance in the return itself and therefore τ is chosen close to zero, see section 4.2.4. However, according to Satchell and Scowcroft (2000) it should be close to one. Koch (2004) on the other hand takes a position somewhat in the middle and finds that values of $\tau = 0.3$ are reasonable. All these differing opinions require an investigation into the value of τ.

Sensitivity of $\boldsymbol{\tau}$ First, we will reexamine the sensitivity analysis of Drobetz (2002). He performed an analysis of the change in the weights of a portfolio caused by increasing the parameter τ. For each value of τ between zero and one we will compile a BL-portfolio for the month April 1982, the distribution of the weights will be plotted against the change in τ.
Under normal circumstances it is only possible to observe a part of the behavior. The effect of τ can be enlarged by scaling the covariance matrix Σ by a factor of ten.
The effect that τ has on a portfolio is portrayed in figure 5.2. For $\tau = 0$, one can observe that the portfolio almost equals the equilibrium portfolio, and as τ increases to one, one sees that weights of the portfolio diverge. It is interesting to investigate why the weights diverge from the equilibrium weights and if they converge to some portfolio, as seems plausible from observing the graph.

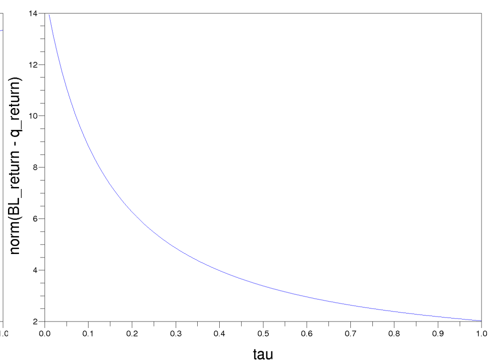
The parameter τ is a scaling parameter that affects the covariance matrix Σ. The assumption is (see assumption A5) that the expected value of the returns is normally distributed with a mean equal to the equilibrium returns and a variance of $\tau\Sigma$, i.e. $\mathrm{E}(\mathbf{r}) \sim N(\boldsymbol{\mu}, \tau\Sigma)$. A larger value of τ, thus means a larger covariance in the equilibrium expected returns and more uncertainty



(a) Influence of τ on weight distribution in the BL-portfolio.



(b) The difference between the BL-return and the return of $\boldsymbol{\pi}$ becomes larger.



(c) The difference between the BL-return and the return of $\mathbf{q}$ becomes smaller.

Figure 5.2: The influence of τ on the weights in the portfolio and the cause for the change in weights. The variance Σ has been enlarged 10 times to make the effect more pronounced.

on the equilibrium portfolio. Therefore, in an allocation the weights should be diverted from this risky position and thus from the equilibrium portfolio. This is exactly what can be observed in the graph.
The Black-Litterman portfolio is a combination of the equilibrium portfolio and the investors's view portfolios. If the weight distribution moves away from the equilibrium portfolio, then one should expect that they move to-

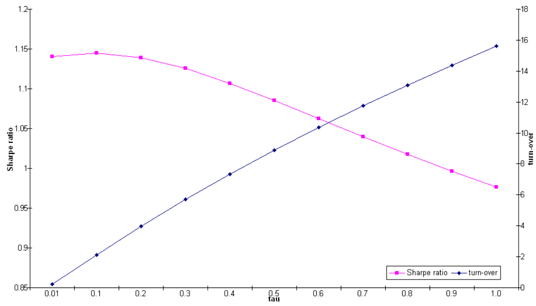
wards the investor's views portfolio.
The investor specifies views on multiple portfolios, therefore there is not one investor portfolio to compare the BL-portfolio to. However, it is possible to compare the BL-optimized returns to the equilibrium returns ($\boldsymbol{\pi}$) and those of the investor ($\mathbf{q}$). The difference between the BL-returns and the equilibrium returns can be found in figure 5.2b.
The difference between the equilibrium returns and the BL-returns is very small for $\tau = 0$, and increases as τ increases to one. The difference between the BL-optimized returns and those of the investor are compared in figure 5.2c. For $\tau = 0$ the difference between the BL-returns and those of the investor are the largest, this corresponds to the weights as the BL-weights are almost equal to the equilibrium weights. As τ increases to one, one sees that the difference becomes smaller. Therefore, it seems plausible the BL-weights converge to the investor's portfolio as τ grows larger.

Calibration of $\boldsymbol{\tau}$ The parameter τ will be calibrated over the period January 1978 to December 1996. The purpose of the calibration is to obtain a value of τ that mixes both sources of information, results in portfolios with a reasonable Sharpe ratio and finally causes an average turn-over in the portfolio of maximally 10%.
The parameter δ will be taken to its default setting for the moment. That means that the risk aversion coefficient δ is taken equal to 4. The uncertainty in the views matrix (Ω) will be estimated by the error in the forecast. The result of the optimization can be found in table 5.6. The Sharpe ratio and the average turn-over are given for the different values of τ. These results have also been displayed graphically in table 5.6b. Examining the graph there seems to be an optimal value of τ, for which the Sharpe ratio is maximal. The Sharpe ratio of the portfolio increases steadily as τ grows smaller and reaches its maximum around τ equal to 0.1, from there on the Sharpe ratio starts diminishing again. However, for this value of τ there is little turn-over in the portfolio, on average only 2.10% of the portfolio changes every month. There is little mixing between both sources of information.

In figure 5.3 is the cumulative return of the portfolio in the time plotted. The cumulative return is defined in this framework as the sum of the monthly returns. In a zero-investment strategy each month one starts with no funds and the funds that are acquired at the end of each month are not reinvested, but they are accumulated. It can be seen from figure 5.3 what it means when there is little turn-over in the portfolio, effectively the BL-portfolio follows the benchmark portfolio. The BL-portfolio differs in some cases from the benchmark portfolio, often with a positive result. The average difference between the BL-portfolio and the benchmark portfolio is now only 7.94%.

τ	Sharpe	turn-over (%)
0.01	1.14	0.23
0.1	1.14	2.10
0.2	1.14	3.98
0.3	1.13	5.71
0.4	1.11	7.35
0.5	1.09	8.89
0.6	1.06	10.36
0.7	1.04	11.76
0.8	1.02	13.10
0.9	1.00	14.39
1.0	0.98	15.63

(a) Sharpe ratio and turn-over.



(b) τ versus the Sharpe ratio and turn-over.

Table 5.6: The influence of τ on the Sharpe ratio and the average turn-over in the portfolio.

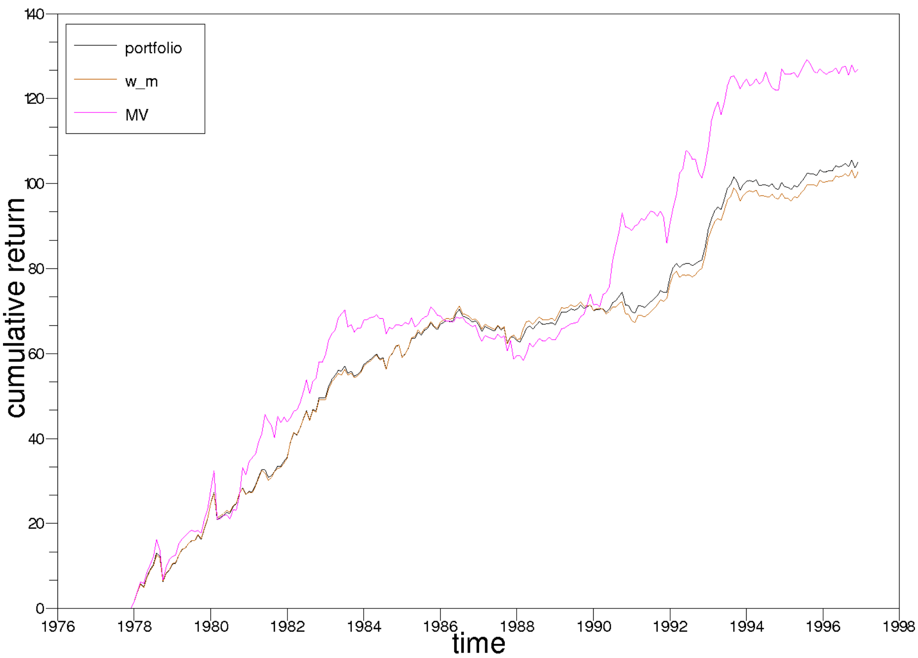


Figure 5.3: For $\tau = 0.1$ does the BL-portfolio almost equal the benchmark portfolio.

For the moment we will keep τ equal to 0.3, the variation in the portfolio is then less than 10% and the Sharpe ratio is reasonable at 1.13.

delta The last parameter that can be calibrated is the parameter δ. The parameter has two alternative explanations as detailed in section 4.2.2. The different interpretation hinge on the manner in which the equilibrium returns are computed. If the equilibrium returns are reverse optimized from some well chosen benchmark, then δ is a risk-aversion coefficient. The parameter, in that case, is set at the start of the optimization procedure. Typical values of the risk aversion parameter vary around three. The equilibrium returns could also be computed from the CAPM. Then δ is a ratio of the market portfolio that varies on every allocation and is specified as $\delta = \frac{\mathrm{E}(r_m)-r_f}{\mathrm{var}(r_m)}$. First, we will investigate the sensitivity of the weights to changes in the risk aversion parameter δ, subsequently will the parameter be calibrated over the period January 1978 to December 1996.

Sensitivity of δ The sensitivity of δ will be investigated by forming a portfolio in the month April 1982 for values of δ ranging from 1 tot 10. The results can be found in figure 5.4.

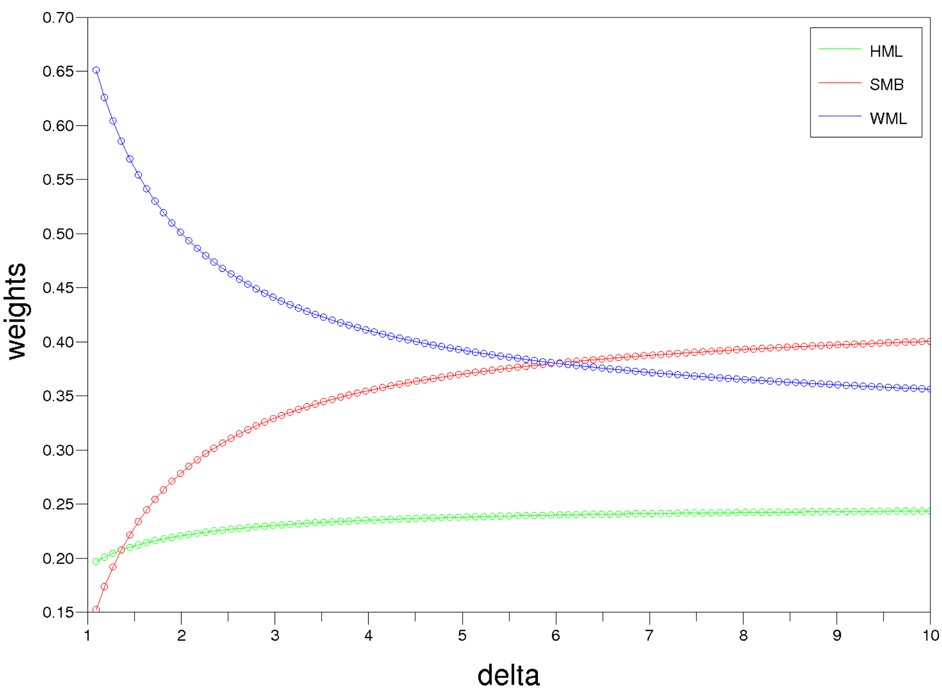


Figure 5.4: The difference between the weights diminish as delta grows larger.

It is visible in this figure that increasing δ has the opposite effect as increasing τ. When increasing δ, the weights seem to move towards each other, the difference between the weights diminishes. This behavior can be understood from the meaning of δ as a risk-aversion coefficient. A larger value of δ means that one becomes more risk-averse, one would like to take on less risk and hence less extreme positions.
That δ truly is a risk-aversion parameter is also supported from evidence of the volatility of the BL-portfolio. The volatility of BL-portfolio formed between January 1978 and December 2006 has been computed, the result can be found in table 5.8. The table shows that increasing delta leads to less volatility. Therefore, δ can be understood to be a risk-aversion parameter.

Calibration of $\boldsymbol{\delta}$ The parameter δ will be calibrated in order to obtain an optimal value. First, we will investigate δ when it is related to the CAPM, the values of δ will be examined by computing some summary statistics of these values. Defining δ as $\frac{\mathrm{E}(r_m)-r_f}{\mathrm{var}(r_m)}$ gives very erratic values in the calibration period, as can be seen from table 5.7. The mean value of δ, the standard deviation, the maximum and minimum value of the values of δ can be observed in the table.

δ	μ	σ	max	min
e	7.03	6.40	35.50	0.11

Table 5.7: The variation in δ when $\delta = \frac{\mathrm{E}(r_m)-r_f}{\mathrm{var}(r_m)}$.

The values of δ are fairly large and very erratic, they are not realistic when delta could also represents a risk aversion coefficient. The Sharpe ratio of the BL-portfolios, formed in the period, is good at 1.04, see table 5.8. Before we draw any conclusion we will investigate δ as a risk aversion coefficient.
The parameter δ could also been seen as a risk aversion coefficient and fixed to a certain value. A range of values of δ will be tried in order to obtain the best Sharpe ratio and variability in the portfolio, these are detailed in table 5.8.

Values of δ around 10 seem to give a good Sharpe ratio of 1.17, then the portfolio varies on average 2.6%. When increasing δ even more, the Sharpe ratio increases some more and the variability in the portfolio diminshes further. The portfolio approximates the benchmark portfolio. Furthermore, as δ should represent a risk aversion coefficient the values then becomes uncharacteristically large. Therefore, we will keep the parameter at four. The

δ	Sharpe	turn-over(%)	volatility
e	1.04	18.98	7.08
1.0	0.81	22.24	8.64
2.0	1.02	11.19	6.12
3.0	1.09	7.52	5.46
4.0	1.13	5.71	5.18
5.0	1.14	4.64	5.03
10	1.17	2.62	4.79
20	1.17	1.73	4.70
50	1.18	1.33	4.65
100	1.18	1.27	4.64

Table 5.8: The performance of the BL-portfolio for $\tau = 0.3$ and various δ.

Sharpe ratio, for the sample period 1978-1996, is 1.13 and the turn-over in the portfolio is on average 5.71%.

Summary The covariance matrix Ω can best be approximated by $\Omega = \hat{\sigma}^2 \mathbf{x}_t'(X'X)^{-1}\mathbf{x}_t$, this gives far superior Sharpe ratios and less turn-over.
The sensitivity analysis of the parameter τ shows that is determines how near the BL-portfolio is either to the equilibrium portfolio or the investor's portfolio. The parameter is calibrated to $\tau = 0.3$.
The sensitivity analysis of δ shows that the parameter behaves like a risk-aversion parameter. Increasing the value of δ leads to less volatile portfolios. The parameter has been calibrated to $\delta = 4$.

5.4.2 BL vs MV

The Black-Litterman model can now be compared to mean-variance optimization with a full investment constraint. The performance of both models will be compared over the period from January 1997 to December 2006. They will be compared via the following measures: the turn-over in the portfolio, the hit ratio, the kurtosis, the mean return, the skewness, the volatility of the return and the Sharpe ratio, additionally are available the minimal and the maximal return. The hit ratio is defined as the percentage of time that the return of a portfolio is positive.

Inspection of table 5.9 learns that the BL-portfolio performs better than the MV-portfolio in almost all respects. The BL-portfolio displays less average month on month turn-over, less volatility and the downward peak in the return is less pronounced. Furthermore, it has a higher hit ratio and most

	BL	MV
mean return	6.93	8.23
volatility	8.87	13.38
Sharpe	0.78	0.62
hit ratio	0.64	0.62
min return	-8.15	-12.26
max return	8.69	23.02
skewness	0.09	1.29
kurtosis	4.63	13.08
turn-over(%)	3.39	12.08

Table 5.9: The BL-portfolio has a higher Sharpe ratio than the MV-portfolio.

importantly a better Sharpe ratio. If we would take transaction costs into consideration, MV would perform even worse, due to the high turn-over in the portfolio. The BL-portfolios have a skewness near to zero, this means that the returns are symmetrically distributed around the mean. Moreover, the kurtosis is 4.63, where the normal distribution has a kurtosis of 3.
As the BL-model performs much better than traditional mean-variance optimization, it is also interesting if it performs better than the individual strategies or than an equal weighting scheme (in this case this equals the equilibrium portfolio). Therefore, the performance has been compared to these portfolios, the result can be found in table 5.10.

	BL	HML	SMB	WML	MV	equal weights
mean return	6.93	6.53	3.67	9.19	8.23	6.46
volatility	8.87	13.51	14.93	20.00	13.38	8.74
Sharpe	0.78	0.48	0.25	0.46	0.62	0.74
hit ratio	0.64	0.61	0.50	0.61	0.62	0.63
min return	-8.15	-12.66	-16.58	-25.05	-12.26	-7.93
max return	8.69	13.71	21.87	18.40	23.02	9.20
skewness	0.09	-0.03	0.77	-0.59	1.29	0.18
kurtosis	4.63	5.06	8.93	6.59	13.08	4.94
turn-over(%)	3.39	0	0	0	12.08	0

Table 5.10: The BL-portfolio performs better than the individual strategies.

Table 5.10 shows that the BL-portfolio does perform better than the indi-

vidual strategies of which the portfolio is composed, it has a higher Sharpe ratio and it performs better than a equal weighting scheme.
The cumulative return of the strategies are plotted in an overview graph of the period (figure 5.5), the graph shows the erratic behavior of the MV-portfolios. It also clearly shows how the BL-portfolio moves between the return of the MV-portfolio and the equilibrium (equal weights) portfolio.

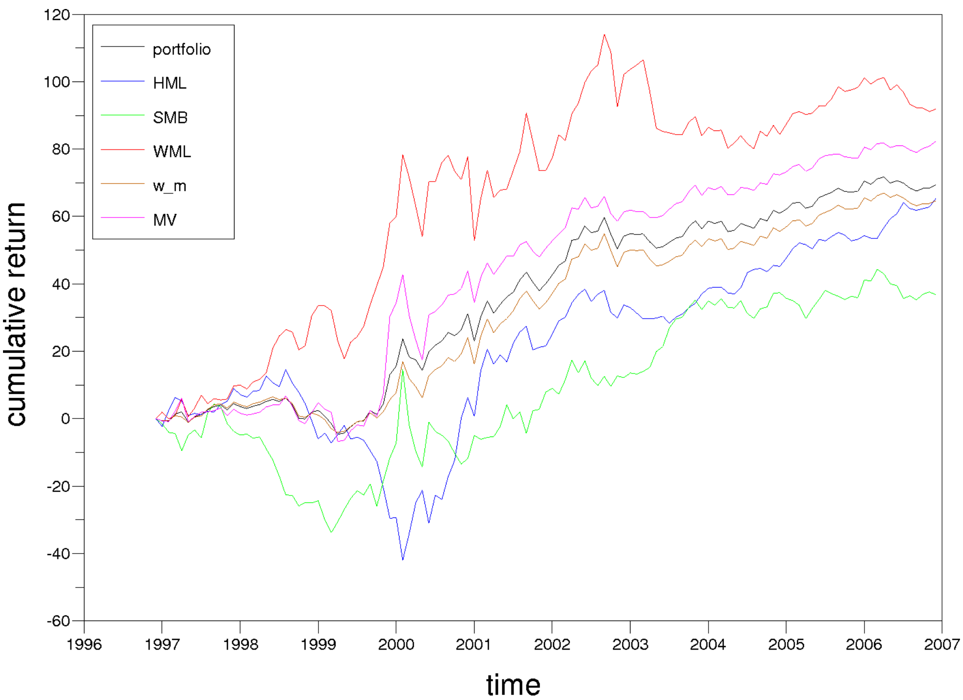


Figure 5.5: Cumulative return of the different portfolios between 1997 and 2006.

The performance of the BL-portfolio tracks the performance of the benchmark portfolio, the average difference between the Bl-portfolio and the benchmark portfolio is 12%. It might be interesting to investigate if the good performance of the BL-portfolio is due to the benchmark, that is whether a change of benchmark would still offer good results.
Therefore, we will try three different benchmarks, the first benchmark is long in HML and WML and leaves out SMB, $\mathbf{w}_m = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$. The second benchmark is long in WML, $\mathbf{w}_m = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$ and the third portfolio is long in SMB, $\mathbf{w}_m = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}$.
From table 5.11 it becomes clear that the BL-portfolios have in all these instances a larger Sharpe ratio then the benchmark portfolio ($\mathbf{w}_m$), although it is not much larger. Altering the benchmark does have a side effect, the BL-portfolios performs worse in comparison to the MV-portfolios in two

	BL	MV	$\mathbf{w}_m$	BL	MV	$\mathbf{w}_m$	BL	MV	$\mathbf{w}_m$
	$\mathbf{w}_m = \begin{pmatrix}\frac{1}{2} & 0 & \frac{1}{2}\end{pmatrix}$			$\mathbf{w}_m = \begin{pmatrix}0 & 0 & 1\end{pmatrix}$			$\mathbf{w}_m = \begin{pmatrix}0 & 1 & 0\end{pmatrix}$		
μ	8.12	8.27	7.86	9.05	8.49	9.19	4.56	8.14	3.67
σ	11.57	13.38	11.75	18.49	13.38	20.00	13.69	13.39	14.93
S	0.70	0.62	0.67	0.49	0.63	0.46	0.33	0.61	0.25
h	0.63	0.62	0.63	0.61	0.63	0.61	0.51	0.62	0.50
t	3.29	12.22	0	3.83	12.25	0	3.7	12.25	0

Table 5.11: Performance of the BL-portfolio with different benchmarks.

cases. It might be possible to gain some better results for the BL-model, by calibrating the parameters. However, this example does show the sensitivity of the BL-model to the chosen benchmark. The BL-model in combination with the original benchmark resulted in a Sharpe ratio of 0.78, this is much better than the worst result of 0.3.

5.5 Conclusion

The empirical study allows several conclusion about the BL-model to be drawn, some conclusion concern its calibration other concern the performance of the BL-model.

The new method of specifying Ω has some good results. It could be argued that this is due to the relatively large values of the variance obtained in this manner. The large uncertainty in the views steers the BL-returns towards the equilibrium returns, which are chosen in such a way that it has under most circumstances a favorable return.

The sensitivity analysis has shed some more light on the parameter τ. The parameter scales the confidence in the benchmark or equivalently the uncertainty in the expressed views. However, it still remains an unintuitive parameter. Knowing that it steers the BL-returns towards the one or the other does not give an intuitive idea of the value of the parameter.

The calibration of τ showed there was certainly an optimal value of τ, figure 5.6b displays the optimal value around 0.1. The relative small value of τ indicates again, that it is best to remain close to the benchmark portfolio.

The sensitivity analysis shows that δ can be seen as risk aversion coefficient, increasing δ leads to less volatility in the portfolio return, which indicates a less risky position. This interpretation of δ additionally, allows more flexibility , then when δ is chosen to equal $\frac{\mathrm{E}(r_m)-r_f}{\mathrm{var}(r_m)}$. Additionally, it offers a possibility to calibrate the parameter.

The performance of the BL-model has been observed over the period 1997 to 2006, inspecting the cumulative return graph (figure 5.5) it becomes clear that the return of the BL-portfolio varies between the benchmark return and the mean-variance return. The BL-model has a higher Sharpe ratio than the mean-variance portfolio and the benchmark portfolio. The BL-portfolios do not differ much from the benchmark, on average 12% of the allocations differ, but when it differs from the benchmark it succeeds in producing extra return.

Therefore, the performance of BL has been evaluated under different benchmarks, this leads to lower Sharpe ratios. Fortunately, the BL-model still performs better than its benchmark, although this is only 2 or 3 basis points. This leads to the conclusion that the BL-model is truly an enhanced indexing model, it improves slightly on the benchmark portfolio. When the BL-model moves away from the benchmark, it does make the right choices and improves on the benchmark portfolio. However, the difficulty in choosing the parameters, the dependency of the BL-model on the benchmark portfolio and the slight improvement over the benchmark would suggest that it maybe easier to choose a benchmark portfolio well and invest in this.

Chapter 6

Conclusion and further research

6.1 Conclusion

The purpose of this thesis has been to investigate the popular Black-Litterman model for asset allocation. The model is designed to alleviate some of the flaws found in the traditional asset allocation model, mean-variance optimization.
Mean-variance optimization has since its origination been the most popular method to allocate assets, the popularity could be due to the understandable premise on which it sorts assets. The central premise in asset allocation is that one should balance risk and expected return, and to only take on more risk if one acquires also more expected return. The mean-variance model allocates assets exactly on this premise. Despite the well accepted underpinning does mean-variance optimization have its flaws.
In practice, the MV-portfolios are often very concentrated in only a few assets and do not reflect the views of the investor. In order to cope with these problems, investors often constrain the mean-variance model in such way that the possible portfolios lie in an bandwidth they are comfortable with.
Black and Litterman set out to alleviate these problems by making a model that would result in intuitive portfolios and a model that could be used by investors. They at least had the effect that the Black-Litterman model has become a very popular model and many papers are written on the subject. However, the papers mainly try to explain the model, as the mathematics of the model are originally not very clearly described. Black and Litterman did not go in to detail about how they obtained the BL-formula and what the parameters meant. Thus, in trying to alleviate the problems of the mean-variance optimization, they also created a few new problems.
Especially the parameter τ has been a source of confusion. It is used to

scale the variance matrix of the equilibrium returns, but how the matrix should be scaled is unclear and what the value of the scaling parameter should be or on what it should depend is unclear. Therefore, a popular method of circumventing this problem has been to choose the uncertainty in views covariance matrix Ω and τ together. The method solves the problem of specifying τ, but mathematically there is no consistent reason to specify these variables together. It seems that this method serves no other purpose than removing the difficult parameter τ from the model.

The Black-Litterman has been the subject of an empirical study in a three asset environment. The purpose of the study was to determine the performance of the Black-Litterman model in comparison to mean-variance optimization. The assets we have chosen are three popular zero-investment strategies, a momentum strategy, a size strategy and a book-to-market value strategy. All the input data, apart from the views of the investor, has been acquired or computed in a straightforward manner.

The views of the investor have been forecasted by a regression analysis on variables that describe the state of the macro economy. The regression analysis, additionally, produces an estimate of the variance of the views and thus an consistent manner to specify the uncertainty matrix Ω. This however, still leaves the parameter τ.

The BL-model and MV-optimization have been compared over the period January 1997 to 2006. The graph of the cumulative return in this period, clearly shows how the return of the BL-portfolio varies between the return of the equilibrium portfolio and those of the investor.

The Sharpe ratio of the BL-portfolio was higher in the observed period, than the Sharpe ratio of the MV-portfolio. The BL-portfolio performed also slightly better, 4bp, than the equilibrium portfolio. A change of the benchmark showed that the performance of the BL-portfolio strongly depends on the benchmark portfolio and the BL-model truly is an enhanced indexing model.

The concept of the BL-model is very good, investors often have difficulty estimating data, see Herold (2003), therefore it is good to combine the ideas of an investor with those of a quantitative source. The way the model was specified however, resulted in a few unintuitive parameters that are difficult to use. The problem of the parameter τ has been solved by specifying the parameter together with the uncertainty matrix Ω. It would be mathematically more consistent and probably easier to use if all the parameters in the model have a consistent definition and a method to calibrate them. Otherwise, it could be better to specify the model in a different manner as for example has been done by Scowcroft and Sefton at UBS.

6.2 Suggestions for further research

As speaks clearly from the conclusion, there is room for improvement in the Black-Litterman model. The suggestions can be divided in two part, first we will discuss suggestions on what could be researched on the BL-model and finally suggestions on models that could be based on the Black-Litterman model.

Suggestions for the BL-model A large problem of the BL-model are the parameters, for there is no comprehensive explanation of the meaning and the value of all parameters. Most notably the parameter τ. It would be good if the influence of the parameter τ on the weights, could be quantified. It has become clear in the empirical study, that τ determines how near the BL-returns are to either the equilibrium returns or the investors returns. If this relationship could quantified even more, for example that a $\tau = 0.3$ means that the weights are a mix of one third the equilibrium returns and two thirds the returns of the investor, then it would be easier to specify τ. The same could be investigated for the parameter δ, although there is less of a direct relationship between δ and the two return estimates.
Another interesting subject is the BL-optimized covariance matrix produced by combining the two sources of information with the Bayesian statistics approach, see equation (4.7). Usually, the covariance matrix Σ is used to optimize the portfolio with the Black-Litterman returns. It is interesting why the BL-covariance matrix is not used, as this is the matrix that corresponds to the BL-returns. A short exploration of using this matrix, shows that it produces portfolios that have a large month-on-moth turn-over. It could be interesting to see why this covariance matrix performs so badly, and if its performance could be improved.
Finally, the model could be improved by adding a method to compute the parameter τ. The original work of Theil (1971) does provide an estimate for τ, see equation (4.11), however it could not be applied to the problem at hand due to constraints on the BL-formulation. It could be a major improvement if the model could be adapted such that there is a method to compute τ.
An easy improvement could also be made by incorporating different probability distributions for the returns, it has be shown by that returns often are not normally distributed, see for example Embrechts et al. (2003). Giacometti et al. (2006) have started with the research into this subject.

Improvements on the BL-model Instead of tinkering with the BL-model, it might be better to start from their premise, the combination of two information sources, and build on that premise a new model. The new model should have parameters that can be understood and chosen by the investor or computed in a consistent manner. The model developed by

UBS, see paragraph 4.4.3, is an example of this. They started from the same premise, adapted the model and left out for example the parameter τ. It would interesting to compare the performance of the UBS model to the model of Black and Litterman.
The model could be improved on even further by making it easier for investors to provide forecasts of returns. Another method, also mentioned by Herold (2003), would be to provide a ranking of the asset performance which subsequently will be transformed to a forecast of the expected return and variance. This removes the difficult for the investor of providing an estimate of the probabilistic measures expected return and variance, but it leaves intact the advantage that a quantitative and a qualitative approach are combined.

Appendix A

Source Code

A.1 Main Program

```
function [weights, weights_MV,  return_totaal, NES] = ...
   BLproces(start, eind, wm, tau);

exec('data\return200701.sce'); exec('data\excess_return.sce');
exec('data\rf.sce'); exec('data\GB.sce'); exec('data\DP.sce');
exec('data\URP.sce'); exec('data\MP.sce');
exec('data\UI.sce');

return_strat = [return_HML, return_SMB, return_WML];
Y = [DP, UI, GB, URP, MP];
[n, m] = size(return_strat);
vmaand = 60;
pmaand = 6;
transaction_costs = 0; // 0.5;
printen = 0;
P = eye(m,m);
[Sigma] = standaardinput(start, eind);
delta = 4;
[d, q, OmegaTau] = calibreerbareinput(start, eind, n, m, tau, 1, P, delta);
[weights, weights_MV,  return_totaal, NES, pis] = ...
   berekenreeks(wm,  start, eind, P,  Sigma, d, q, OmegaTau, printen);

endfunction
```

A.2 Subroutines

A.2.1 'Standaard Input'

```
function [Sigma] = standaardinput(start, eind);

Sigma = zeros(m, m, n);
load('data\rd256.dat', 'rd256')
rd = rd256;
// grootte van de 3-d matrix met dag-date bepalen.
// w = aantal dagen, e = aantal maanden,
// r = aantal assets
[w, e, r] = size(rd);
decay_factor = 0.97;
s = zeros(w, r);
for i = start:eind
  s(:, :)  = rd(:, i, :);
  Sigma(:, :, i) = variancedecay(s, decay_factor, 'r');
end

endfunction
```

A.2.2 'Calibreerbare Input'

```
function [d, q, OmegaTau] = ...
   calibreerbareinput(start, eind, n, m, tau, O, P, delta_method);

d = zeros(n, 1);
q = zeros(n, m);
OmegaTau =zeros(n, m);
for i = start:eind
  d(i) = delta(i, delta_method); // delta, uitrekenen
    // qp voorspelling van return uitrekenen
    // Omega geschat op basis van fout in forecast
  [qp, Omega] = lse_forecast(return_strat(i-vmaand:i, :), Y(i-vmaand:i, :));
    // de views van de investeerder (q) is een transformatie van de
    voorspellingen (qp) met matrix P
  q(i, :) = (P*qp)';
  if O == 1 then
    OmegaTau(i, :) = diag(P*Sigma(:, :, i)*P')';
  else
    OmegaTau(i, :) = Omega/tau;
  end
end
```

```
endfunction
```

'lse forecast'

```
function [q, s2, bet] = lse_forecast(r, Y)

// maakt een voorspelling van r op t+1 doormiddel van lineare regressie op Y
// r_t = X_{t-1}*beta + e_t

// grootte van r en Y bepalen
[lr, mr] = size(r);
[lY, mY] = size(Y);

// output initialisatie
q = zeros(mr,1);
s2 = zeros(1, mr);

// fout in input afvangen
if lr <> lY then error('The matrices should have the same length.'); return, end;

// lineare regressie met constante term, dus vector van enen toevoegen aan matrix Y
X = [ones(lY, 1), Y];

// schatter beta maken
bet = zeros(mr, mY+1);
bet = lse( X(1:lr-1, :), r(2:lr, :), 'r' );

// tijdsrelatie constant veronderstellen en r op t+1 schatten
q = bet*(X(lr, :))';

// schatting van homescedatic variance van q, formule uit Theil(1971)
XX = inv(X'*X)
M = eye(lr, lr) - X*XX*X';
s = diag(r'*M*r)/(lr-mY-1);

for i=1:mr
// variantie matrix van de schatter voor de verschillende assets is:
  sig(:, :, i) = s(i)*XX;
// variantie matrix van de fout in voorspelling bepaald door waarnemingen op tijdstip
// lr en door fout in schatter
  s2(1, i) = X(lr, :)*sig(:, :, i)*X(lr, :)';
end

endfunction
```

A.2.3 'Berekenreeks'

```
function [weights, weights_MV,  return_totaal, NES, pis] = ...
   berekenreeks(w,  start, eind, P,  Sigma, d, q, OmegaTau, printen);

// 'berekenreeks' berekend de gewichten voor een BL-portfolio.
// De inputs zijn: w, start, eind, P, Sigma, OmegaTau, printen
//  w = benchmark portfolio
//  start = maand waarin optimalisatie proces start
//  eind = maand waarin proces eindigt
//  P = views matrix van investeerder
//  Sigma = serie van covariantie matrices
//  OmegaTau = covarariantie matrix, onzekerheid in investeerder's views
//  printen = booelean, 1 = save output
// De outputs:
//  [weights, weights_MV, return_totaal, NES, pis]
//  weights =  de gewichten na BL-optimalisatie,
//  weights_MV = gewichten na MV-optimalisatie
//  return_totaal = matrix van maandelijkse return van portfolio en assets
//  NES, pis =  de BL-expected return en de equilibrium returns

weights = zeros(return_strat);
weights_MV = zeros(return_strat);
NES = zeros(n, m);

// Benchmarkportfolio als start portfolio in vector van weights geplaatst.
weights(start, :) =  w';
weights_MV(start, :) = w';    // evenals in de MV-weights vector

// optimalisatie proces uitvoeren;
// de nieuwe expected return de gewichten bepalen
for index = start:eind-1
    // Berekenen van een Nieuwe Expected returns vector (NE)
    [NE,  pi] = BL(Sigma(:, :, index), P, q(index, :)', d(index), w, ...
                diag(OmegaTau(index, :)));
    // MV optimalisatie met fullinvestment constraint
    weights(index+1, :) =  (MVequal(NE, d(index), Sigma(:, :, index)))';
    // MV met de voorspellingen q en variantie Omega
    weights_MV(index+1, :) = (MVequal(q(index, :)', d(index), ...
                              diag(OmegaTau(index, :))*tau))';
    NES(index+1, :) = NE';
    pis(index+1, :) = pi';
end;
```

```
return_portfolio = diag(weights*(return_strat)') ....
                     - trading_costs(weights, transaction_costs, 'c');
return_MV = diag(weights_MV*(return_strat)')
             - trading_costs(weights_MV, transaction_costs, 'c');
return_totaal = [return_portfolio, return_strat, return_MV];

if printen == 1 then
  print('\resultaten\result070412\weights.txt', weights);
  print('\resultaten\result070412\weights_MV.txt', weights_MV);
end

endfunction
```

BL

```
function [NE, pi] = BL(Sigma, P, q, delta, w, OmegaTau)

//de functie berekend de BL geoptimaliseerde nieuwe vector van verwachte return.
// initialisatie output parameter
NE = zeros(w);
pi = zeros(w);
//pi reverse optimizen uit de benchmark portfolio
pi = (delta*Sigma)*w;
// de nieuwe expected return vector
NE = (inv(inv(Sigma) + P'*inv(OmegaTau)*P))*(inv(Sigma)*pi + P'*inv(OmegaTau)*q);

endfunction
```

MVequal

```
function [weight] = MVequal(pi, delta, Sigma)

// mean-variance optimization with a full investment constraint,
// i.e. the method of Lagrange multipliers is used.
// inputs:
// pi, the vector of expected return, delta the risk-aversion coefficient,
// sigma the covariance matrix.
// gam is the lagrange multiplier, A and b are the full investment constraints,
// C is a term that is introduced for convenience,

// initialisation of output parameter
weight = zeros(pi);

// determine size of pi
```

```
[n, m] = size(pi);

// full investment constraint A*w= b
A = ones(1, n);
b = 1;
C = delta^(-1)*inv(Sigma);
gam = inv(A*C*A')*(b - A*C*pi );
weight = C*(A'*gam + pi);

endfunction
```

Glossary

Symbols

$\mathsf{cov}(x_i, x_j)$ The covariance of x_i and x_j, this can also be denoted as σ_{ij}.

$\mathbf{q}$ $\in \mathbb{R}^k$ the vector that specifies the expected return of the investors view.

$\mathbf{r}$ $\in \mathbb{R}^n$ vector of random variables, that represents the return or excess return of a vector of assets. The standard assumption is that the returns are normally distributed, $\mathbf{r} \sim N(\mathrm{E}(\mathbf{r}), \tau\Sigma)$.

$\mathsf{var}(\cdot)$ The variance operator, often shortened to σ^2.

$\mathbf{w}$ $\in \mathbb{R}^n$ vector of weights of the assets in the portfolio. There could be a full investment constraint, that requires the weights to sum to one. Alternatively, it could be a zero-investment portfolio, then the weights sum to zero.

$\mathbf{w}_m$ The (world) market portfolio, this consists of all the assets in the market to the ratio of the respective market capitalizations. The market portfolio is often approximated by a well chosen benchmark.

$E(\cdot)$ Expected value operator.

P $\in \mathbb{R}^{k\times n}$ the view matrix specifies which assets are under consideration of the investor. A row of P can specify a view on a single asset or a portfolio of assets. Views are specified as $\mathbf{q} = P\mathrm{E}(\mathbf{r}) + \boldsymbol{\epsilon}$.

r_f Return of the risk-free asset.

r_p The return of a portfolio, $r_p = \sum_{i=1}^{n} w_i r_i$.

δ Black and Litterman call this the (world) risk-aversion coefficient, it is also said to equal $\frac{\mathrm{E}(r_m)-r_f}{\sigma_m^2}$.

$\boldsymbol{\epsilon}$ — $\in \mathbb{R}^k$ the vector of uncertainty in the expressed view, this vector is normally distributed with mean $\mathbf{0}$ and variance Ω.

λ — The risk aversion coefficient or a Lagrange multiplier.

μ — The mean value of a quantity.

Ω — $\in \mathbb{R}^{k\times k}$ the diagonal covariance of views matrix.

$\boldsymbol{\pi}$ — The equilibrium returns, these are often computed from $\boldsymbol{\pi} = \delta\Sigma\mathbf{w}_m$.

ρ — The correlation coefficient.

Σ — $\in \mathbb{R}^{n\times n}$ covariance of return matrix.

τ — A highly debated parameter, Black and Litterman use it to scale the covariance matrix Σ.

Words[1]

basis point (bp) A basis point is a percentage of a percentage, i.e. 100 bp = 1% or equivalently 1 bp = 0.01%.

benchmark portfolio (benchmark) A portfolio against which the performance of the investment manager can be measured, this could for example be an index portfolio like the AEX-index or the NASDAQ. A well diversified benchmark can serve as a proxy for a market portfolio.

bond A debt investment with which the investor loans money to a company or government that borrows the funds for a (defined) period of time at a specified interest rate.

book equity (BE) The accounting value of the common equity.

capital asset pricing model (CAPM) An economic model of the behavior of asset prices under conditions of risk. The model is intended to promote an equilibrium relationship for pricing the risk associated with holding assets. It predicts that the expected rates of return will be directly related to a single common factor: the return on the market portfolio.

equilibrium return The expected returns that would hold if the market is in an equilibrium, often CAPM returns are used.

[1] The definitions are obtained from various sources: Smullen and Hand (2005), Moles and Terry (1997)

equity Represents the ownership of a part of a company, synonyms are shares and in the US also stocks.

expected excess return (expected return) The expected return of an asset in the domestic currency minus the domestic risk free rate $\mathrm{E}(\mathbf{r}) - r_f$.

growth equity Equity that is categorized in the bottom BE/ME percentile. Common equity that is expected to increase, or has already increased, its earnings per share at a rate faster than that for the market as a whole.

hedge A transaction or position designed to mitigate the risk of other financial exposures.

high-minus-low (HML) A zero-investment strategy that entails to go long in value equity, with a high book-to-market value and to go short in a growth equity, equity with a low book-to-market value.

long Going long means to bet on a price increase by taking a positive position in the asset. The asset is bought to be sold a later time at a higher price.

market equity (ME) The value of common equity if it would be sold on the market.

mean-variance (MV) Evaluating uncertain investments in terms of their expected return and variance of outcomes.

MV-efficient portfolio A portfolio that has minimum variance of return for a certain level of expected return, or equivalently maximum expected return for a certain level of variance.

non-systematic risk The risk that is uncorrelated with the market and can be diversified, also known as specific risk.

posterior distribution In a Bayesian sense, it represents the probability distribution of the random variable A, given observations of the random variable B. It is the distribution that can be obtained after Bayesian inference.

prior distribution In a Bayesian sense, it represents the probability distribution of the random variable A, prior to any observations.

return Short for rate of return, which is the fraction of profit to invested money.

risk A characteristic of the return of an asset, that defines the chance that a certain estimate is not made.

risk aversion The investment philosophy that extra risk should only be taken on, if it also implies a higher expected return.

risk aversion coefficient The measure of risk averseness, when there is an utility function this can be defined mathematically.

short Shorting an asset, or going short means to bet on a price fall by taking a negative position in the asset. The asset is first sold and later bought back at a lower price.

small-minus-big (SMB) Investment strategy that entails to go long in firms with a small size and to go short in firms with a large size. The size of a firm is measured by its market equity.

systematic risk The risk associated with the market as a whole, this risk cannot be diversified.

utility theory The field of study concerned with analysis and ordering of preferences that is used to explain individual decision-making.

value equity Equity that is categorized in the top BE/ME percentile, equity of a company with solid fundamentals that are priced below those of its peers, based on analysis of price/earnings ratio, yield, and other factors.

volatility The standard deviation, the square root of the variance, is often called the volatility in finance.

weight The weight is the proportion an asset has in the portfolio.

winners-minus-losers (WML) A zero-investment strategy that entails to go long in equity that has performed well in the past 3 to 18 months and to go short in equity that has performed poorly in that period. WML is more commonly known as momentum, or UMD (up-minus-down).

zero-investment portfolio An investment portfolio where the long positions are financed by short positions in other assets, the weights in the portfolio sum to zero.

Bibliography

Baker, M. and J. Wurgler. 2006. *Investor sentiment and the cross-section of stock returns*, The Journal of Finance **41**, no. 4, 1645–1680.

Banz, R. W. 1981. *The relationship between return and market value of common stocks*, Journal of Financial Economics **9**, no. 1, 3–18.

Best, M. J. and R. R. Grauer. 1991. *On the sensitivity of mean-variance-efficient portfolios to changes in asset means: Some analytical and computational results*, The Review of Financial Studies **4**, no. 2, 315–342.

Black, F. 1989. *Universal hedging: Optimizing currency risk and reward in international equity portfolios*, Financial Analyst Journal **45**, no. 4, 16–22.

______. 1990. *Equilibrium exchange rate hedging*, The Journal of Finance **45**, no. 3, 899–907.

Black, F. and R. Litterman. 1991a. *Global asset allocation with equities, bonds and currencies*, Fixed Income Research, Goldman, Sachs & Co.

______. 1991b. *Asset allocation: combining investor views with market equilibrium*, The Journal of Fixed Income, 7–18.

______. 1992. *Global portfolio optimization*, Financial Analysts Journal **48**, no. 5, 28–43.

Chan, L. K. C. and J. Lakonishok. 2004. *Value and growth investing: Review and update*, Financial Analysts Journal **60**, no. 1, 71–86.

Chen, N.-F., R. Roll, and S. A. Ross. 1986. *Economic Forces and the Stock Market*, The Journal of Business **59**, no. 3, 383–403.

Chen, N.-F. 1991. *Financial investment opportunities and the macroeconomy*, The Journal of Finance **46**, no. 2, 529–554.

Conrad, J. and G. Kaul. 1998. *An anatomy of trading strategies*, Review of Financial Studies **11**, no. 3, 489–519.

Drobetz, W. 2002. *Einsatz des Black-Litterman verfahrens in der asset allocation: Handbuch asset allocation* (H. Dichtl, J. M. Kleeberg, and C. Schlenger, eds.), Uhlenbruch Verlag, Bad Soden.

Embrechts, P., F. Lindskog, and A. McNeil. 2003. *Modelling dependence with copulas and application to risk management: Handbook of Heavy Tailed Distributions in Finance* (S. T. Rachev, ed.), Elsevier, Amsterdam. available at: http://www.ma.hw.ac.uk/m̃cneil/ftp/DependenceWithCopulas.pdf.

Fama, E. F. and K. R. French. 1989. *Business conditions and expected returns on stocks and bonds*, The Journal of Financial Economics **25**, no. 1, 23–49.

______. 1992. *The cross-section of expected stock returns*, The Journal of Finance **47**, no. 2, 427–465.

______. 1993. *Common risk factors in the returns on stocks and bonds*, The Journal of Financial Economics **33**, no. 1, 3–56.

———. 1995. *Size and book-to-market factors in earnings and returns*, The Journal of Finance **50**, no. 1, 131–155.

Frankfurter, G. M., H. E. Phillips, and J. P. Seagle. 1971. *Portfolio selection: The effects of uncertain means, variances, and covariances*, Journal of Financial and Quantitative Analysis **6**, no. 5, 1251–1262.

Giacometti, R., M. Bertocchi, S. T. Rachev, and F. J. Fabozzi. 2006. *Stable distributions in the Black-Litterman approach to asset allocation*, technical report. available at http://www.pstat.ucsb.edu/research/papers/BLapproach2005.pdf.

Graham, A. 1987. *Nonnegative matrices and applicable topics in linear algebra*, Halsted Press, New York.

Hanoch, G. and H. Levy. 1969. *The efficiency analysis of choices involving risk*, Review of Economic Studies **36**, no. 3, 335–346.

He, G. and R. Litterman. 1999. *The intuition behind Black-Litterman model portfolio*, Investment Management Research, Goldman, Sachs & Co.

Herold, U. 2003. *Portfolio construction with qualitative forecast*, The Journal of Portfolio Management **30**, no. 1, 61–72.

Idzorek, T. M. 2004. *A step-by-step guide to the Black-Litterman model: Incorporating user-specified confidence levels*, Zephyr Associates, Inc, unpublished. available at: http://www.globalriskguard.com/resources/assetman/BL_Draft_with_Graphs.pdf.

Jegadeesh, N. and S. Titman. 1993. *Returns to buying winners and selling losers: Implications for stock market efficiency*, The Journal of Finance **48**, no. 1, 65–91.

———. 2001. *Profitability of momentum strategies: Evaluation of alternative explanations*, The Journal of Finance **41**, no. 2, 699–720.

Johnson, R. A. and D. W. Wichern. 1998. *Applied multivariate statistical analysis*, 4th ed., Prentice Hall, Upper Saddle River, NJ.

Keim, D. B. 1983. *Size-related anomalies and stock return seasonality: Further empirical evidence*, Journal of Financial Economics **12**, no. 1, 13–32.

Koch, W. 2004. *Consistent return estimates in the asset allocation process: The Black-Litterman approach.* presentation available at: http://www.globalriskguard.com/resources/assetman/blach-litterman.pdf.

Lee, P. M. 1997. *Bayesian statistics: An introduction*, 2nd ed., Arnold, London.

Lee, W. 2000. *Advanced theory and methodology of tactical asset allocation*, John Wiley & Sons, New York.

Lintner, J. 1965. *The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets*, The Review of Economics and Statistics **47**, no. 1, 13–37.

Luenberger, D. G. 1998. *Investment science*, Oxford University Press, New York.

Markowitz, H. 1952. *Portfolio selection*, The Journal of Finance **45**, no. 1, 31–42.

———. 1959. *Portfolio selection*, John Wiley & Sons, New York.

———. 1987. *Mean-variance analysis in portfolio choice and capital markets*, Basil Blackwell, New York.

Michaud, R. O. 1989. *The Markowitz optimization enigma: is 'optimized' optimal?*, Financial Analysts Journal **45**, no. 1, 31–42.

Moles, P. and N. Terry. 1997. *The handbook of international financial terms*, Oxford University Press, Oxford, Oxford Reference Online, accessed on 07-02-2007. available at http://www.oxfordreference.com/views/ENTRY.html?subview=Main&entry=t181.e4581.

Mossin, J. 1966. *Equilibrium in a capital asset market*, Econometrica **34**, no. 4, 768–783.

Pesaran, M. H. and A. Timmermann. 1995. *Predictability of stock returns: Robustness and economic significance*, The Journal of Finance **50**, no. 4, 1201–1228.

RiskMetrics Group. 2001. *Return to RiskMetrics: The evolution of a standard*, Technical document. available at http://www.riskmetrics.com/pdf/rrmfinal.pdf.

Rosenberg, B., K. Reid, and R. Lanstein. 1985. *Persuasive evidence of market inefficiency*, Journal of Portfolio Management **11**, no. 3, 9 –16.

Satchell, S. and A. Scowcroft. 2000. *A demystification of the Black-Litterman model: Managing quantitative and traditional portfolio construction*, Journal of Asset Management **1**, no. 2, 138–150.

Scowcroft, A. and J. Sefton. 2003. *Enhanced indexation* (S. Satchell and A. Scowcroft, eds.), Butterworth-Heinemann, Amsterdam.

———. 2005. *Understanding momentum*, Financial Analysts Journal **61**, no. 2, 64–82.

Sharpe, W. F. 1964. *Capital asset prices: A theory of market equilibrium under conditions of risk*, The Journal of Finance **19**, no. 3, 425–442.

———. 1970. *Portfolio theory and capital markets*, McGraw-Hill, New York.

Smullen, J. and N. Hand. 2005. *A dictionary of finance and banking*, Oxford University Press, Oxford, Oxford Reference Online, accessed on 14-12-2006. available at http://www.oxfordreference.com/views/ENTRY.html?subview=Main&entry=t20.e161.

Stattman, D. 1980. *Book values and stock returns*, The Chicago MBA: A journal of selected papers **4**, 24 –45.

Theil, H. 1971. *Principles of econometrics*, John Wiley & Sons, New York.

Theil, H. and A. S. Goldberger. 1961. *On pure and mixed statistical estimation in econometrics*, International Economic Review **2**, no. 2, 65–78.

Tobin, J. 1958. *Liquidity preferences as behavior towards risk*, Review of Economic Studies **25**, no. 3, 65–86.

Treynor, J. L. 1961. *Towards a theory of market value of risky assets*, unpublished.