An investigation into the Black-Litterman model

Author: Martin Felix Jørgensen

Supervisor: Niels Henrik Lehde Pedersen

COPENHAGEN BUSINESS SCHOOL
HANDELSHØJSKOLEN

Graduate Diploma in Business Administration in Finance - HD(F)
Department of Finance, ©2016.

Executive Summary

This report describes a practical application approach to the Black-Litterman model, which is developed on the basis of CAPM and modern portfolio-theory. The model allows investors to implement subjective socalled "views" or opinions on one or several assets/asset classes. Hence, the model is an asset-allocation model, developed in order to get closer to an optimal mathematical allocation-model, that supports investors to better manage investment portfolios. Of particular financial interest is model uncertainties and the influence of central variables on the final results. All figures or tables without external references are the author's own. This report is additionally filed as concluding 15 ECTS point project (thesis) for the Graduate Diploma in Business Administration in Finance at CBS.

Forord

Denne rapport beskriver en praktisk anvendelses tilgang til Black-Litterman modellen, som en videre-udvikling til CAPM og moderne portefølje-teori. Modellen tillader investorer at implementere subjektive såkaldte "views" eller meninger om et eller flere aktiver/aktivklasser. Modellen er således en aktiv-allokeringsmodel, udviklet for at komme tættere på en optimal matematisk allokerings-model, der kan støtte investorer til bedre at forvalte investerings-porteføljer. Af særlig finansiel interesse er model usikkerheder og betydningen af centrale variabler på de færdige resultater. Alle figurer eller tabeller uden eksterne henvisninger er forfatterens egne. Nærværende rapport er herudover indleveret som afsluttende 15 ECTS point projekt (afhandling) for HD studiet i finansiering på CBS.

Contents

1 Introduction

The mathematical foundation of modern portfolio theory (Modern Portfolio Theory (MPT)) was developed in the 1950s through the early 1970s. Two pioneers of the theory behind MPT are Harry Markowitz and William Sharpe, Nobel Prizes (see e.g. Bodie et al. (2014)). The ideas are based on risk/return trade-off such that high risk assets should be priced higher than low-risk assets. The theory of efficient and free markets ("efficient market hypothesis") will lead to an equilibrium such that if the price of high-risk assets or securities is considered to be too high, then the price of these assets or securities will begin to fall causing the expected return to increase until buyers and sellers agree on a given price (market forces). The expected returns $E(r)$ is defined as (Bodie et al., 2014, p.128):

$$E(r) = \sum_{s} p(s)r(s) \tag{1}$$

where s is an asset or securities, $p(s)$ is the weight or probability of something to happen and $r(s)$ is holding-period return calculated as (Bodie et al., 2014, p.128):

$$r(s) = \frac{\text{Ending price of share - Beginning price + Cash dividend}}{\text{Beginning price}} \tag{2}$$

For a time-series, a sequence of numbers of stock returns can be calculated using $E(r) = \frac{1}{n}\sum_{s=1}^{n} r(s)$ where n is the number of observations and $r(s)$ is a discrete (sampled) rate of return values, e.g. daily, weekly or montly. In such cases each "sample" has equal weight and therefore, that is only a minor modification of (1).

MPT also provides methods for calculating the effect of diversifaction of a portfolio consisting of a number of securities, using statistical and mathematical methods (mean-variance methods). Investors, investment managers or funds use different strategies and methods for trying to beat the market and earn the highest returns. However "the market" is very complicated to define as it ideally should include all kinds of assets, including something like investments in real estate portfolios. Such figures are not always easy to obtain or define as each investment by definition is very different. It is common to define a benchmark to use as index for investment portfolios instead of comparing returns to an existing index. Investment strategies can be split into either passive or active management strategies:

- Passive management is generally the cheapest as the strategy is to hold a highly diversified portfolio without monitoring it. The idea is that if one believes in the efficient market hypothesis, then it is virtually impossible to beat the market (the prices are already correct) so diversification is very important.

- Active management is more expensive as portfolio managers continously monitor the development of each security and actively tries to be over- or under-weighted in certain securities. Active management requires investors to use resources for analyzing accounts, read market news, watch the performance and to re-balance investment portfolios.

The passive strategy is easy to implement. Instead of hand-picking individual stocks one can choose to invest in one of the major already existing indices, e.g. Dow Jones, Nasdaq, VIX/Volatility S&P 500, S&P 500 (all US), Shanghai Composite/Nikkei (both Asia), OMXC20 (Denmark), OMXS30 (Sweden), FTSE (UK) or DAX (Germany) – and many more. The passive strategy is not as interesting to investigate as it is limited in how many ways we can act, if the idea is just to follow a benchmark (or construct a very diversified portfolio which is then mostly left untouched).

1.1 Portfolio selection and historical overview

Assets are typically said to be stocks, bonds and cash (money market securities). Each asset class can be divided into e.g. large-cap, mid-cap and small-cap stocks. Benchmarks exist for many asset classes, e.g. for emerging markets, national or international securities. Treasury bills (T-bills) have different maturities so an "optimal investment model" is probably difficult to define, there are many (numerous, countless) combinations and strategies. Investors must make their own decision about what they think provide the maximum return, while minimizing risk.

Conservative portfolio strategies have low risk and low returns. They typically consist of many bonds or fixed income securities and few stocks/equities. Aggressive portfolio strategies consist of many stocks/equities and few bonds or fixed income securities. Re-balancing the portfolio involves selling those portions of the portfolio that have increased significantly and to buy cheaper assets or assets, where the price has decreased sufficently until the expected returns again increase to acceptable levels.

Finally, the theoretical concept of the risk free interest rate and risk premiums should be introduced. The risk free rate is defined as the rate of return that can be earned, by investing in risk-free assets such as T-bills, money market funds or leaving the money in the bank (Bodie et al., 2014, p.129). For a given investment horizon (e.g. 1 year), the maturity of the risk-free interest rate should match. Currently the interest rates are very low. If we assume an expected index fund return of 8% and a risk-free interest rate of 1% to be used as proxy for the market returns, we define the risk premium as the difference, i.e $E(R_M) = E(r_M) - r_f = (8\% - 1\%) = 7\%$. The 7% is the price compensation we then require for taking this investment risk. Because there are no guarantees, investing in this fund does not necessarily lead to a profit of 7% in addition to r_f which is risk-free. It could lead to a profit of 4%, 7%, 10% or we could lose maybe 5-25%.

Risk is traditionally calculated using the variance or standard deviation, assuming sample data follows a normal distribution. Some investors are more than willing to put on additional risks while others try to avoid risk. The behaviour of investors are traditionally quantified by using a risk aversion coeffecient. The higher risk aversion, the higher risk premium is required and vice-versa. Expected returns can be calculated with (1) and risk or standard deviation can also easily be calculated, once time-series have been downloaded and imported. Portfolio selection on risky assets can be performed using the famous framework described in Markowitz (1952). Markowitz wrote the process can be divided in two stages:

1. Observation and experience, which ends with beliefs about future performances.

2. Beliefs about future performances end with the choice of portfolio.

Later, James Tobin (1958) added risk-free assets to the analysis and framework of Markowitz. This allowed leveraged or deleveraged portfolios and became the basis of a super-efficient portfolio and the capital market line. The articles by Sharpe (1964) and Lintner (1965) introduced the Capital Asset Pricing Model (CAPM). This is an important theory that explains the expected returns as a function of the systematic risk. Sharpe showed the market portfolio on the efficient frontier is the same super-efficient portfolio Tobin introduced. CAPM has later turned out to be very important and widely used. Investors should hold the market portfolio, either leveraged or de-leveraged with the risk-free asset. Later more advanced models such as Intertemporal Capital Asset Pricing Model (ICAPM) and Arbitrage Pricing Theory (APT) was introduced in the 70's by Merton and Ross. Together with Sharpe (1964) and Lintner (1965), Mossin (1966) is also recognized for observing that ideally investors should choose portfolios as a linear combination of the risk-free asset and the market portfolio.

The Black-Litterman model came many years later (see e.g. Black and Litterman (1992)). The model allows investors to implement their own beliefs, in order to help decide which weights for each asset class or security, should be used for optimal portfolio construction. Managers opinions are implemented through "views". The model is the main topic for this report, both from a theoretical and a practical point of view. Figure 1 illustrates the structure of the report.

1.2 Research questions

The Black-Litterman model introduces more complexity than known from traditional MPT. The objective is to integrate quantitative methods with qualitative or subjective estimates, however there are many uncertainties to consider. The following main research question have been asked:

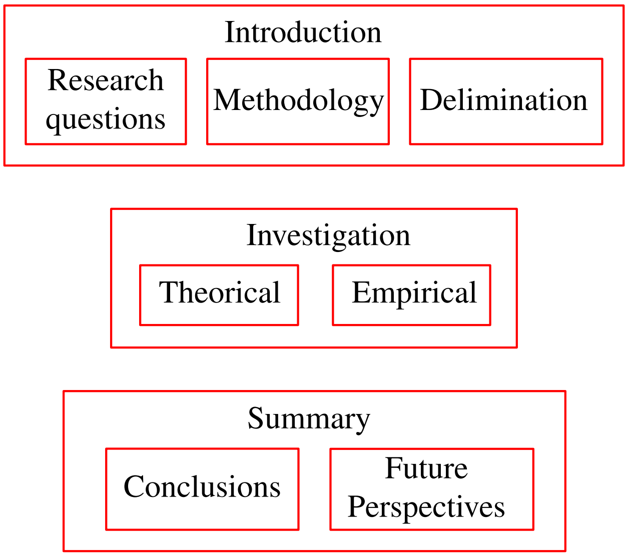


Figure 1: Graphical depiction of project structure and report chapters. The empirical part of the project should also be understood as the practical part, i.e. it is the basis for the investigation of the research questions.

> *How to use and construct portfolios (how to mix assets) with traditional portfolio optimization theory and also by using the Black-Litterman model?*
> *What is the impact or effect from individual components of the Black-Litterman equation?*

The following additional questions have been formulated, in order to help answering the main research questions:

- How to calculate historical covariance matrices and perform reverse optimization on the assets using MPT, for generation of a mean-variance portfolio?
- Which elements or components do the Black-Litterman model consist of?
- What happens when the views are changed and everything else is kept constant?
- What is the meaning of uncertainty on the view portfolios and how to deal with the unknown variables?
- What is the difference between views made with large or small confidence? How to implement such a method in the Black-Litterman model framework?

A systematic approach will be tried out, where only a single (or a few parameters) will be changed at the same time. The effect on the output will be studied and illustrations and figures will be used to draw conclusions on or to verify or reject any hypotheses.

1.3 Delimination

Everywhere normal distributions are used and everywhere it is assumed that the variance is constant. The alternative is to think of variances as e.g. time-dependent and estimating such an effect from ordinary least squares (that is definately outside the scope of this work).

It is assumed that the reader knows the most important basics of financial theory, possibly also some basic statistics/mathematics which is also outside the scope of this report.

2 Introduction to portfolio theory

The foundation of MPT is the Markowitch mean-variance portfolio optimization theory. That is also the foundation for the more advanced Black-Litterman model.

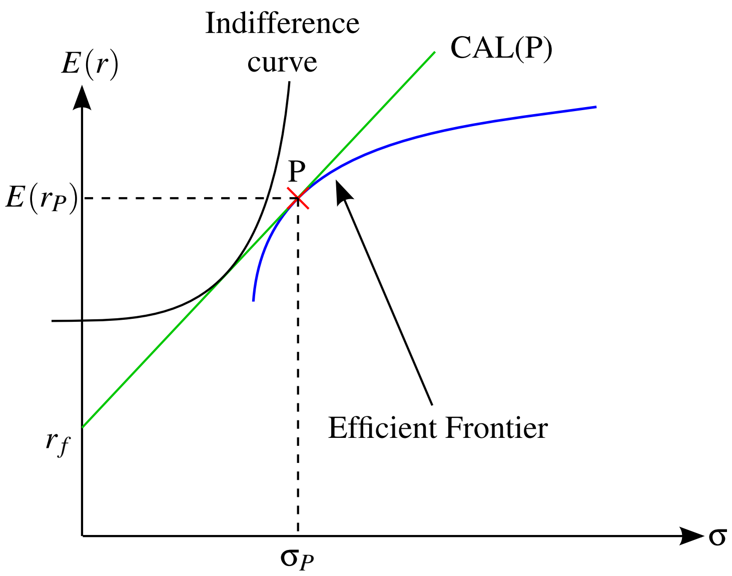


Figure 2: Illustration of the CAL, .

Figure 2 illustrates the Capital Allocation Line (CAL) from r_f to a place on the efficient frontier. The indifference curve can be moved. The green CAL-line is used to decide how much of the risk-free assets, the individual investor should buy, together with the risky portfolio. The risky portfolio here is illustrated by the point P. The CAL could be drawn as a straight line from r_f to anywhere on the efficient frontier and then the indifference curve should be moved to touch the CAL. The higher slope of the CAL, the higher risk/reward-ratio and the higher potential outcome of the investment.

If the Markowitz optimization process is used by all investors to hold securities, the CAL is replaced by a Capital Market Line (CML) as shown in Figure 3. The market portfolio is not easy to determine so for real applications we use a proxy, a benchmark or fund index. The market portfolio should contain an aggregation of all possible risky portfolios, made of up all assets in the investable universe. This leads to a discussion of the CAPM line. The relationship between the risk and return of a portfolio can be described using the Capital Asset Pricing Model (CAPM), again based on Markowitz Portfolio Theory:

$$R_i = r_f + \beta(r_m - r_f) \tag{3}$$

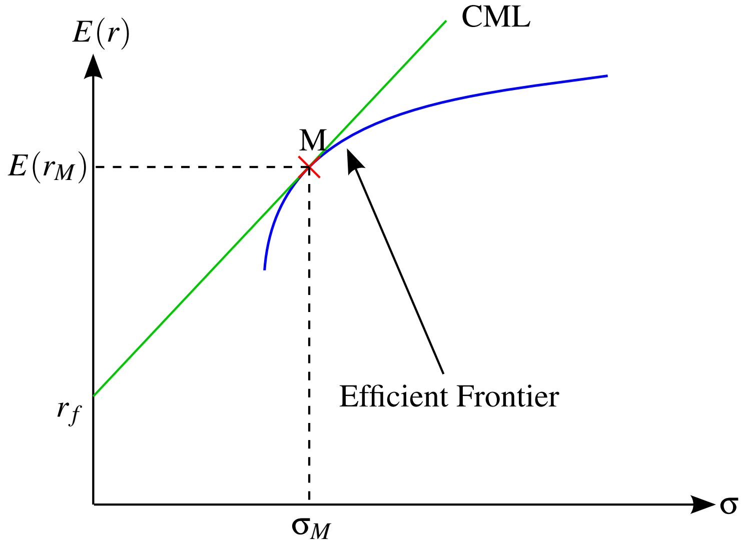


Figure 3: Illustration of the CML, .

where

$$r_i = \text{Rate of return for stock } i. \tag{4}$$

$$r_m = \text{Rate of market return.} \tag{5}$$

$$r_f = \text{Risk-free interest rate.} \tag{6}$$

$$\beta = \frac{cov(R_i, R_m)}{var(R_m)} \tag{7}$$

$$\tag{8}$$

Richard Roll introduced what later became known as Roll's critique which was described in Roll (1977). One key conclusion is that the market portfolio is unobservable. We cannot include all assets in our model such as e.g. investments in real estate, precious metals, jewelry and many other things. For this reason, it's impossible to test the CAPM. In spite of this, it's a popular model.

Assuming that **r** is a vector of asset returns (typically excess returns) and CAPM is valid, the $n \times 1$ equilibrium excess returns vector is $\Pi = \beta(\mu_m - r_f)$ where μ_m is the return on the global market and β is an $n \times 1$ vector of asset betas (Satchell and Scowcroft, 2000, p.139):

$$\beta = \frac{\text{Cov}\left(\mathbf{r}, \mathbf{r}^{\mathbf{T}}\mathbf{w}\right)}{\sigma_m^2} \tag{9}$$

2.1 Markowitz mean-variance model

Figure 4 illustrates that for a portfolio of many risky assets, ideally we want to invest in portfolios on the efficient frontier. These portfolios maximize the expected return $E(r)$ while at the same time, they have minimum risk (standard deviation σ). Returns can be calculated at time t using either discrete returns, i.e. $(r_{i,t} - r_{i,t-1})/r_{i,t-1}$, continously compounded returns ($\ln(x_{i,t}/x_{i,t-1})$ or with dividends: $\ln(x_{i,t} + D_{i,t})/x_{i,t-1}$. The risk-free interest rate is illustrated with r_f, which is relevant for the understanding CAPM.

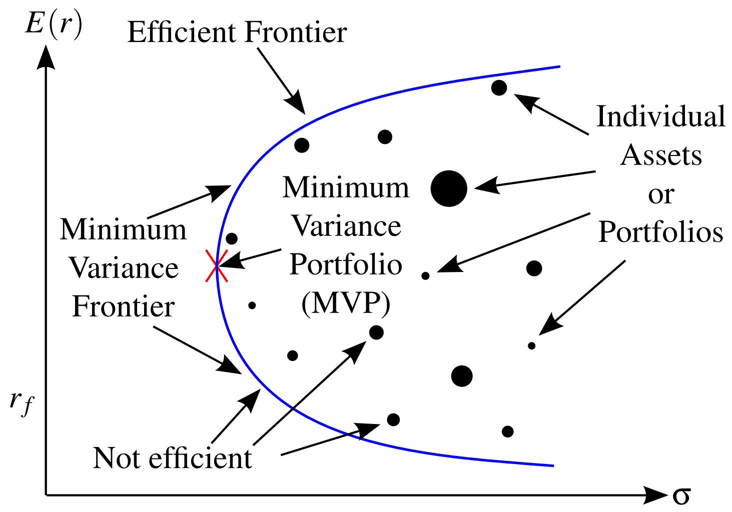


Figure 4: Illustration of the concept for the minimum-variance frontier for risky assets. The blue curve shows possible envelope portfolios (above MVP are efficient envelope portfolios, but those below are not efficient).

The optimal portfolio that can be constructed with the highest expected returns $E(r)$ given the risk σ, is the same portfolio that has the highest Sharpe ratio or the steepest slope from the origin, to the portfolio on the Markowitz efficient frontier (Bodie et al., 2014, p.134 and p.216):

$$\text{Sharpe ratio} = \frac{\text{Risk premium}}{\sigma(\text{excess return})} = \frac{E(r_p) - r_f}{\sigma_p} \tag{10}$$

By solving a set of equations with the constraint that the Sharpe ratio should be maximized one arrives at the solution, which is a vector $\mathbf{w}$. This vector contains the optimal weights of each of the i individual securities, such that the sum of all weights equal 100%. Depending on risk aversion, fund managers or investors might want to combine the risky investment by also investing in risk-free assets with interest rate r_f ($\sigma = 0$ as illustrated). Investors could also borrow money and leverage their investments (make use of financial gearing). It is an important concept, that the optimal (efficient) portfolio can have less risk when adding more and more assets, due to the effect of diversification. The expected return of a portfolio of i securities is found to be exactly as in (1) with variance as in (Bodie et al., 2014, p.222):

$$E(r_p) = \sum_{i=1}^{n} x_i E(r_i) \tag{11}$$

$$\sigma_p^2 = \sum_{i=1}^{n} \sum_{i=1}^{n} x_i x_j \text{Cov}(r_i, r_j) \tag{12}$$

where for each security or asset i, x_i is the weight, $E(r_i)$ is the expected return and $\text{Cov}(r_i, r_j)$ is the covariance of securities i and j. The covariance matrix of expeced returns for securities (r_i and r_j) can be calculated using different methods, one of them is by using the sample covariance method:

$$\text{Cov}(r_i, r_j) = \sigma_{i,j} = \frac{\sum_{k=1}^{n} (r_{ik} - \bar{r}_i)(r_{jk} - \bar{r}_j)}{n-1} \tag{13}$$

where n is the number of data values, $\bar{r}_i = E(r_i)$ and $\bar{r}_j = E(r_j)$ is the average (expected value) of the vector of return values (this makes the vector centered or we could also call the terms, excess returns terms). If the excess returns matrix has been calculated, the covariance matrix is one vector multiplied with the other vector transposed and then divided by number of observations, n minus 1. Division can also be made with n instead of $n-1$, but in that case it should be assumed that all the discrete data available is from the whole population of data and that normally does not apply for stock or security data. However the difference is negligible for large number of data sets. To illustrate this (Benninga et al., 2008, p.292):

$$A = \text{Matrix of excess returns} = \begin{bmatrix} r_{11} - \bar{r_1} & \dots & r_{N1} - \bar{r_N} \\ r_{12} - \bar{r_1} & \dots & r_{N2} - \bar{r_N} \\ \vdots & & \vdots \\ r_{1M} - \bar{r_1} & \dots & r_{NM} - \bar{r_N} \end{bmatrix} \tag{14}$$

The covariance matrix for n periods is

$$\mathbf{\Sigma} = \sigma_{i,j} = \frac{A^T A}{n-1} \tag{15}$$

A problem with estimation of covariances based on historical data, is that historical data usually is a bad predictor for future covariances. Therefore, using a mechanical method as traditional MPT suggests doing, is a bad idea. The method can produce unrealistic results and unrealistic weights. Sometimes, however, it might help a little to add a constraint that prevents short-sales. A socalled single-index model for constructing the covariance matrix could also be considered. So-called “shrinkage” methods that may produce more reliable covariances also exists, however those methods are not considered to be of particular importance here. Throught this report, when referring to variance, standard deviation or mean returns, that is given in the context of a normal distribution. The normal distribution probability[1] around the mean μ with standard deviation σ is written $\sim \mathcal{N}(\mu, \sigma^2)$ and given by

$$f(x|\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad -\infty < x < \infty \tag{16}$$

where σ^2 is the variance or squared standard deviation.

2.2 Optimal (efficient) portfolio weights using Black's method

As described in (Benninga et al., 2008, Ch.8.5, p.250), the efficient portfolio is one that solves:

$$\min \sum_i \sum_j x_i x_j \sigma_{ij} = \mathrm{Var}(r_p) \tag{17}$$

Subject to

$$\sum_i x_i r_i = \mu = E(r_p) \tag{18}$$

$$\sum_i x_i = 1 \tag{19}$$

As shown in Black (1972) we can find the whole efficient frontier by choosing any two efficient portfolios, x and y consisting of N risky assets. If a is a constant, the portfolio Z is also efficient:

$$z = ax + (1-a)y = \begin{Bmatrix} ax_1 + (1-a)y_1 \\ ax_2 + (1-a)y_2 \\ \vdots \\ ax_N + (1-a)y_N \end{Bmatrix} \tag{20}$$

If the expected returns for security i is $E(r_i)$ such that $\mathbf{E}(\mathbf{r}) = \{E(r_1), E(r_2) \ldots, E(r_N)\}^T$, it follows from (Benninga et al., 2008, Ch.9.2, p.262), that we can define a constant c and setup the equations $\mathbf{E}(\mathbf{r}) - c = \sigma_{ij}\mathbf{z}$ where $\mathbf{z}$ is unknown and must be solved for:

$$\mathbf{z} = \boldsymbol{\sigma}_{ij}^{-1}\{\mathbf{E}(\mathbf{r}) - c\} \tag{21}$$

$$\mathbf{x} = \{x_1, x_2 \ldots, x_N\}^T \tag{22}$$

[1] https://en.wikipedia.org/wiki/Normal_distribution

where

$$x_i = \frac{z_i}{\sum_{j=1}^{N} z_j} \tag{23}$$

It also follows from (Benninga et al., 2008, Ch.9.2, p.265) that if y is the first envelope portfolio and x is any other portfolio (not necessarily an envelope portfolio), we have the relationship:

$$E(r_x) = c + \beta_x[E(r_y) - c] \tag{24}$$

where

$$\beta_x = \frac{\text{Cov}(x,y)}{\sigma_y^2} \tag{25}$$

which is also known as Black's zero-beta CAPM, from the paper Black (1972). This was a very quick summary of something that is not a main topic for the report. The methods described above will shortly be used from page 25.

2.3 Introduction to the Black-Litterman model

The Black-Litterman approach was published and refined in the early 1990's by Fisher Black and Robert Litterman from Goldman-Sachs, see Black and Litterman (1992). Mechanical optimization based on historical returns is normally not a good estimate for future returns. Traditional portfolio theory can lead to e.g. large short positions and historical data is not always a good indication for future returns. Optimization constraints can limit unrealistic large short or long positions but there are other alternatives. Traditionally MPT is known to produce unrealistic results. The Black-Litterman method tries to solve some of the problems. In 1998 a paper by Goldman-Sachs in greater detail discussed how to implement the method (see e.g. Bevan and Winkelmann (1998)):

> *Investors should take risk where they have views, and correspondingly, they should take the most risk where they have the strongest views.*

As described in Benninga et al. (2008), the idea is that in the absence of other information, the benchmark cannot be outperformed. Instead of having an optimal portfolio as output from input data, it is assumed that a given portfolio is already optimal. The result of the BL approach is the expected returns of each of the components of the benchmark. If an investor disagrees with the expected returns of one of more components of an portfolio, he can incorporate his views. Otherwise that person should just buy the portfolio (or benchmark) and there is no need for using the model.

According to Walters et al. (2014) it is an essential assumption of both mean-variance optimization as well as for the Black-Litterman model, that asset returns are normally

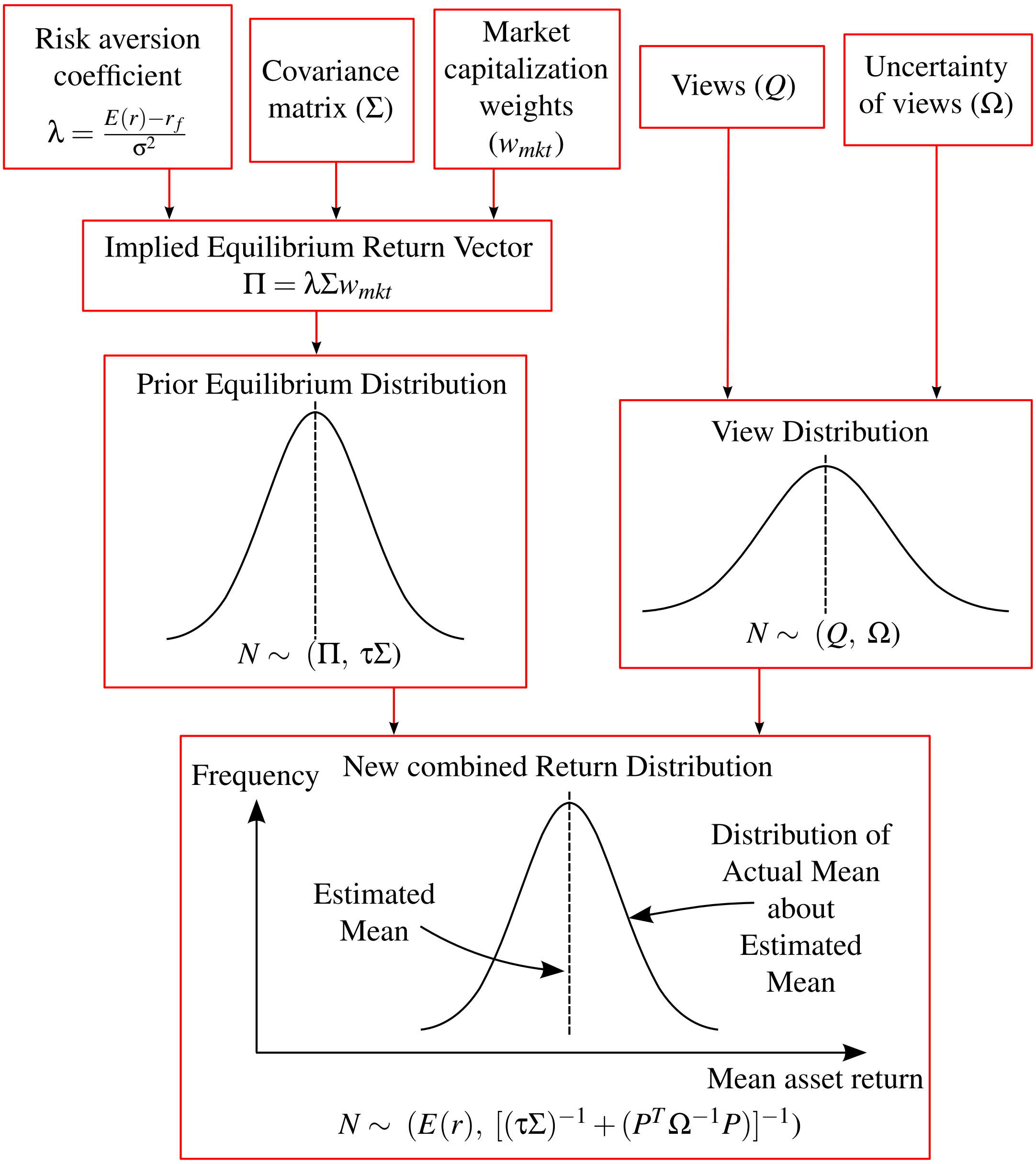


Figure 5: Basics of the model as illustrated in the book (Satchell, 2011, Fig.2.1) with minor additions. According to same book, the variance of the new combined return distribution is derived in Satchell and Scowcroft (2000).

distributed. Because the input data are assumed to follow normal distributions, the output (posterior) will also follow a normal distribution. Another key assumption is that "unknown mean and known variance" is used and for that reason we talk about "the variance of the unknown mean about the actual mean". Figure 5 illustrates the concepts of the model. Given that the following is known:

1. λ (sometimes also written as δ): The risk aversion coefficient of the market portfolio (scalar). It is either specified or computed using (Satchell and Scowcroft, 2000, p.139)
$$\lambda = \frac{\mu_m - r_f}{\sigma_m^2} \tag{26}$$

2. τ: A number that specifies the uncertainty of the prior estimate of the mean returns (this parameter will be discussed in greater detail in section 2.5 from page 16).

3. Information about views: If m is the number of views and n the number assets, the model views are specified using the P-matrix (if more than 1 view) and Q-vector:
   - P: is a $m \times n$-matrix that identifies which assets are involved in either absolute or relative views (the Q-vector).
   - Q: is the "views", a vector of $n \times 1$ elements.

4. Σ: The size $n \times n$ variance-covariance matrix of excess returns (in the following, it is designated as the covariance matrix but the diagonal contains the variances for each asset).

5. w_{eq}: Equilibrium market capitalization weights (a vector of $n \times 1$ weight factors for all assets that totals to 100 %). It can be calculated with an unconstrained mean-variance reverse optimization process as described in (Idzorek, 2005, Eq.2): max $w^T\mu - \lambda w^T \Sigma w/2$, using the definition of the implied excess equilibrium return vector Π ($n \times 1$-vector), the risk aversion coefficient λ and Σ[2]:
$$\Pi = \lambda \Sigma w_{eq} \tag{27}$$
   where the $n \times 1$ vector of implied excess equilibrium returns (prior returns) is Π. Substituting Π in (27) with any other vector of excess returns μ yields
$$w_{eq} = (\lambda \Sigma)^{-1} \mu \tag{28}$$
   The risk-aversion coefficient λ can be considered a scaling factor for the reverse optimization estimate of excess return and is also a measure of the expected risk-return tradeoff.

The posterior estimate of returns can be calculated using:

[2] Other authors, e.g. Walters et al. (2013) use $\delta = \lambda$ so the equation becomes: $\Pi = \delta \Sigma w_{eq}$

1. The standard BL-equation (often called the “master formula”):

$$E(r) = \left[(\tau\Sigma)^{-1} + P^T\Omega^{-1}P\right]^{-1} \left[(\tau\Sigma)^{-1}\Pi + P^T\Omega^{-1}Q\right] \tag{29}$$

The first factor (inverted, the “denominator”) can be considered as a normalization factor (though it is a matrix) and the second factor (the “numerator”) is a vector composed of something involving equilibrium returns Π and estimates Q. In the second factor, it can be seen that the first term has $\tau\Sigma)^{-1}$ is a weighting factor and $P^T\Omega^{-1}$ acts as a weighting factor for the second term. It is also seen that if there are no views:

$$E(r) = \left[(\tau\Sigma)^{-1}\right]^{-1} \left[(\tau\Sigma)^{-1}\Pi\right] = \Pi \quad \text{(equil. returns)} \tag{30}$$

If there is no estimation error, $\Sigma^{-1} \to \infty$, hence

$$E(r) \approx \left[P^T\Omega^{-1}P\right]^{-1} \left[P^T\Omega^{-1}Q\right] = P^{-1}Q \quad \text{(view returns)} \tag{31}$$

2. Another not so common version of the BL-equation is shown in Walters et al. (2014):

$$E(r) = \Pi + \tau\Sigma P^T(P\tau\Sigma P^T + \Omega)^{-1}(Q - P\,\Pi) \tag{32}$$

with covariance matrix of returns M:

$$M = \left((\tau\Sigma)^{-1} + P^T\Omega^{-1}P\right)^{-1} \tag{33}$$

and posterior covariance (most authors do not compute this):

$$\Sigma_p = \Sigma + M = \Sigma + \left((\tau\Sigma)^{-1} + P^T\Omega^{-1}P\right)^{-1} \tag{34}$$

For each of the m expressed views, Ω is an $m \times m$ diagonal covariance matrix of the error that represents the uncertainty in each view. As described in e.g. Idzorek (2005) (and many other places):

$$\Omega = \begin{bmatrix} \omega_1 & 0 & 0 \\ 0 & \omega_2 & 0 \\ 0 & 0 & \omega_3 \end{bmatrix} = \begin{bmatrix} \tau(p_1\Sigma p_1^T) & 0 & 0 \\ 0 & \tau(p_2\Sigma p_2^T) & 0 \\ 0 & 0 & \tau(p_3\Sigma p_3^T) \end{bmatrix} \tag{35}$$

ω_i are the variances of the views scaled by the factor τ (the variance on the error terms of the views). The inverse of the variances are also known as the precision, hence the inverted covariance-matrix is a precision-matrix. Because there is uncertainty in the views, error terms ε are added where $\varepsilon \sim \mathcal{N}(0,\Omega)$.The sum can be expressed by a vector of unknown mean returns:

$$P \cdot E(r) = Q + \varepsilon = \begin{bmatrix} Q_1 \\ \vdots \\ Q_m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_m \end{bmatrix} \tag{36}$$

where $E(r)$ is a vector of expected returns of all assets and Q is a vector of estimated returns to each of the combined views, i.e. the product $P\mu$.According to e.g. Idzorek (2005), if an investor is 100% confident in the expressed view, the i'th error term $\omega_i = 0$ and otherwise it is not 0. Satchell and Scowcroft (2000) describes that larger ω_i represents a larger degree of disbelief represented by γ_i where $\gamma = \mathbf{P}E(\mathbf{r})$. Therefore Ω represents the uncertainty of the views and larger ω_i values represents greater uncertainty. Methods for determining the error terms is considered to be one of the most complicated aspects of the model and it will therefore be discussed later, in greater detail from page 16.

2.4 Implementation of investor views

The following is a short example about expressing the views using the P-matrix and the Q-vector. Consider an example of the 6 assets: USA, Germany, Hong Kong, France, UK, China with 3 views, involving assets as shown below:

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0,3 & 0 & 0 & -0,3 \\ 0,8 & -0,2 & 0 & -0,2 & -0,2 & -0,2 \end{bmatrix} \tag{37}$$

The Q-vector could be e.g.

$$Q = \begin{bmatrix} 5\% \\ 2,5\% \\ 7,5\% \end{bmatrix} \tag{38}$$

This example has been taken to involve all assets (no zero-columns). The first row in the P-matrix tells how assets are involved for the first view and so on for the following rows. Therefore there exists m rows in the P-matrix and the number of n columns must be equal to the number of assets. An explanation to each of the views follows:

View 1 : This view only involves the first asset (USA), with the factor 1. If the row summates to 0, it is a relative view. For this reason, the view is an absolute view. As the first entry in the Q-vector is 5% this view tells, that our investor believes that USA will have an excess return of 5%.

View 2 : The sum of factors in row 2 is 0, hence it is a relative view. If a factor is positive, the asset outperforms other (negative factor) assets (assuming that the scalar entry in the Q-vector is also positive, which it normally is). Entry number 2 in the Q-vector is 2,5%, i.e. USA outperforms Germany with a factor 2,5% and Hong Kong outperforms China as the the last asset which underperforms equally much by 0,75% (30% of 2,5%).

View 3 : View 3 is again relative and similar to view 2 except only USA outperforms 4 of the other assets. As the last element of the Q-vector is 7,5%, this view tells that our investor believes that USA outperforms 4 other assets by 6% (80% of 7,5%), while each of the 4 other assets (Germany, France, UK, China) underperforms by 1,5%.

The next section addresses the level of confidence of investor views.

2.5 Dealing with uncertainties in investor views

The matrix Ω as calculated by (35) is the covariance matrix of the error term and a measure of the uncertainty of the views. According to He and Litterman (2002) each diagonal element of Ω is a function of τ, i.e. $\omega_i = \tau(p_i \Sigma p_i^T)$. This implies that the factor τ is directly connected to uncertainty of the views and therefore also the uncertainty in investors prior estimate of the returns. However, different authors construct the Ω-matrix differently, as an example:

- In He and Litterman (2002), $\tau = 0.05$ because this number:

  ... corresponds to the confidence level of the CAPM prior mean if it was estimated using 20 years of data.

- In Satchell and Scowcroft (2000) they write:

  [A2]... and τ is a (known) scaling factor often set to 1.

- Another approach without the use of τ in Ω is described in Meucci (2010) where $\Omega \equiv \frac{1}{c} P \Sigma P^T$, where c is a positive scalar that represents an overall confidence level in the views. In Meucci (2005) this version is given for the "uncertainty matrix":

$$\Omega \equiv \left(\frac{1}{c} - 1 \right) P \Sigma P^T \tag{39}$$

If $c \to \infty$ the view distribution will be infinitely disperse so the views will have no impact. If $c \to 1$ the distribution will be infinitely peaked implying that the investor is trusted completely. If $c = 1/2$ the investor and market model is trusted equally. Therefore it is said that the Black-Litterman model smoothly blends the the market model and the investors opinion.

A key difference between using high and low confidence in the views is illustrated in Figure 6. A few conclusions, illustrated by the figure:

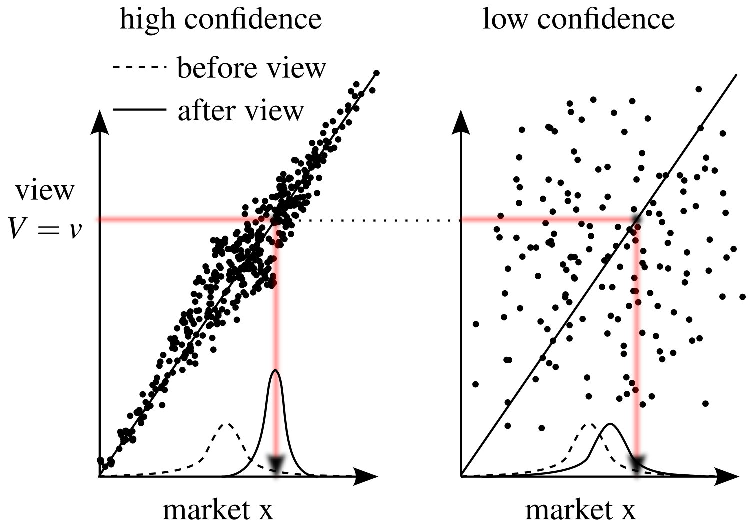


Figure 6: Illustration as in Meucci (2005) of daily returns for market X, which is also normally distributed s.t. $X \sim \mathcal{N}(\mu, \sigma^2)$. The view V is a conditional distribution $(V|x)$ and reflects the confidence in the view.

- If an investor is very insecure about the views (low confidence), the views (the conditional distribution) will have a large standard deviation σ and vice-versa.
- If the confidence is low, the figure shows a very scattered cloud of returns and vice-versa.
- Views with low confidence doesn't significally change the mean returns μ but with high confidence the market distribution will be affected.

There are several methods to calculate or asses the variance or uncertainty in the estimate of the mean. According to Walters et al. (2014) Ω can be calculated as:

1. Make it proportional to the variance of the prior estimate
2. Use a confidence interval
3. Use the variance of residuals in a factor model
4. Use Idzorek's method

2.5.1 Proportional to the variance of the prior estimate

These methods weights the investors views and the market equilibrium weights. The method in He and Litterman (2002) and other authors employs τ as a proportionality factor and the method is the same as shown in (35):

$$\Omega = \text{diag}(P(\tau\Sigma)P^T) \tag{40}$$

This method is the most common method used in the litterature. Most authors agree that τ is between 0 and 1. From the article Black and Litterman (1992):

> *... Because the uncertainty in the mean is much smaller than the uncertainty in the return itself, τ will be close to zero.*

From Lee (2000), *tau* should be between 1% and 5%. Idzorek (2005) wrote

> *Lee, who has considerable experience working with a variant of the Black-Litterman model, typically sets the value of the scalar (τ) between 0.01 and 0.05, and then calibrates the model based on a target level of tracking error.*

In the footnote after that sentence is written: "*This information was provided by Dr. Wai Lee in an e-mail.*" In Blamont and Firoozy (2003) $\tau = 1/N$ because the authors interpret $\tau\Sigma$ as the standard error of the implied equilibrium return vector Π, i.e. the factor depends on the number of N observations. Another method is the one already described by (39), which do not care about diagonalization.

2.5.2 Confidence interval

If an investor believes that an asset has a 3% mean return and expects that 68% of returns will likely be in the interval 2% - 4%, then $\mu = 3\%$ and 68% of a normal distribution is close to 1 standard deviation. The variance in this case becomes $(1\%)^2$ and this value will be used in Ω.

2.5.3 Variance of residuals in a factor model

The return of an asset r using a factor model is described as:

$$r = \sum_{i=1}^{n} \beta_i f_i + \varepsilon \tag{41}$$

where β_i is the factor loading, f_i is the return from the factor and ε is a normally distributed residual value. Then the variance of ε can be computed as part of a regression as the squared standard error. According to Walters et al. (2014) the mixing will be more robust, if only diagonal elements will be used. In Beach and Orlov (2007) GARCH[3] style factor models are used to generate views.

[3] Generalized ARCH, see e.g.: https://en.wikipedia.org/wiki/Autoregressive_conditional_heteroskedasticity

2.5.4 Idzorek's method

As described in Litterman et al. (2003), there is no "universal answer" on how to specify the diagonal elements of Ω, representing the uncertainty of the views. Remember (35) that specified:

$$\Omega = \begin{bmatrix} \omega_1 & 0 & 0 \\ 0 & \omega_2 & 0 \\ 0 & 0 & \omega_3 \end{bmatrix} \tag{42}$$

Idzorek (2005) provides a method where the confidence is a weight factor or level of confidence from 0% to 100%. This is more intuitive in relationship with the variance of the view number k ($p_k \Sigma p_k^T$). If all diagonal elements of Ω are zero, that is equivaluent to specifying 100% confidence in all views. Idzorek uses the BL-equation from (32) which is repeated below and modified slightly

$$E_{100\%}(r) = \Pi + \tau\Sigma P^T (P\tau\Sigma P^T + \cancelto{0}{\Omega})^{-1}(Q - P\,\Pi) \tag{43}$$

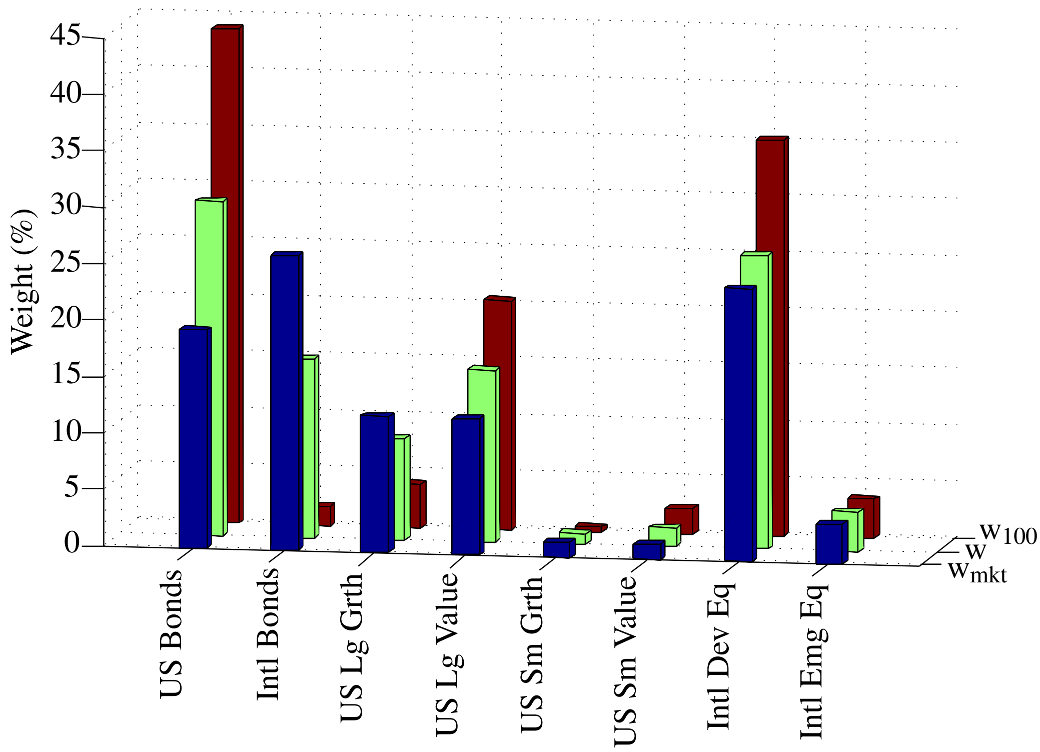


Figure 7: Illustration of the weights w_{eq}, w and $w_{100\%}$ using Idzorek's method. The figure is a reproduction of (Idzorek, 2005, Fig.2.2) which uses w_{mkt} as notation for w_{eq}.

where the subscript 100% is added to the new combined return vector $E(r)$ and Ω is cancelled out, because there is no uncertainty in the views. The weight vector based on 100% confidence is denoted $w_{100\%}$. As the whole optimization problem starts out with a vector of market capitalization weights w_{eq}, Idzorek introduces the "implied confidence level" as the fraction

$$I = \frac{w - w_{eq}}{w_{100\%} - w_{eq}} \tag{44}$$

where the denominator expresses the maximum possible weight difference. Figure 7 illustrates the 3 different weights w_{eq}, w and $w_{100\%}$ and it can be seen that even though there are only k views, it is possible to calculate an implied confidence level for each asset (confidence level in asset space). As there can be many views but only 1 asset, I is the "aggregate confidence level". Equation (44) must only be used for those assets that are involved in any of the views. In the example shown by Idzorek (2005), there are no views involved for the last asset. Therefore there is no implied confidence for the last asset. Suppose the confidence level is specified to be $C = 65\%$ for ω_k. The diagonal element of Ω is then

$$\omega_k = \alpha P \Sigma P^T, \qquad \text{where} \quad \alpha = \frac{1-C}{C} \tag{45}$$

This has the effect that if confidence is 100% then $\alpha = 0\%$. If confidence on the other hand is 0% then $\alpha \to \infty$ like ω_k. Idzorek proposes a more intuitive non-mathematical method for specifying the diagonal elements of Ω. He introduces the user-specified confidence level C from 0-100% such that for view number k of n views:

$$Tilt_k \approx (w_{100\%} - w_{eq}) \cdot C_k \tag{46}$$

where $Tilt_k \approx (w_{k,\%} - w_{eq})$ is also the difference between the target weight vector $w_{k,\%}$ and the equilibrium weights. It is the tilt caused by view number k. The procedure is to calculate $E(r_{k,100\%})$, $w_{k,100\%}$ and the difference from market equilibrium $D_{k,100\%} = (w_{k,100\%} - w_{eq})$. The tilt becomes the user-specified confidence C_k multiplied by $D_{k,100\%}$ and the target weight vector is $w_{k,\%} = (w_{eq} + Tilt_k)$. The diagonal elements ω_k of Ω is the solution to the minimum least-square problem:

$$\min \sum (w_{k,\%} - w_k)^2 \tag{47}$$

subject to $w_k > 0$ where

$$w_k = (\lambda\Sigma)^{-1} \left[(\tau\Sigma)^{-1} + p_k^T \omega_k^{-1} p_k \right]^{-1} \left[(\tau\Sigma)^{-1}\Pi + p_k^T \omega_k^{-1} Q_k \right] \tag{48}$$

One should repeat the steps for all n views and finally $E(r)$ for all views can be calculated. Idzorek points out that despite the complexities in each step, the key advantage is to be able to specify the confidence using a more intuitive scale ranging from

0-100%. Alternative methods are more abstract for specifying the diagonal elements of Ω. Idzorek finishes by concluding that his new method should increase the intuitiveness and usability by helping users realize the benefits of the Black-Litterman model.

2.6 Derivation of the risk aversion and prior returns equations

Equations (26) for the risk aversion coefficient and Equation (27) for prior returns or implied excess equilibrium returns are simple and easy to derive by introducing the quadratic utility-function U as in e.g. Walters et al. (2014):

$$U = w^T \Pi - \left(\frac{\lambda}{2}\right) w^T \Sigma w \tag{49}$$

where the objective for the “reverse” optimization is to maximize the convex utility function U. With no constraints, there will be a closed form solution meaning that the expression can be solved for analytically using a finite number of elementary functions. By differentiating (49) and setting it to 0:

$$\frac{dU}{dw} = \Pi - \lambda\Sigma w = 0 \quad \Rightarrow \quad \Pi = \lambda\Sigma w_{eq} \tag{50}$$

which is exactly the solution vector of excess returns shown in (27). Multiplication of both sides of this equation by w^T yields

$$w^T \Pi = w^T \lambda\Sigma w \tag{51}$$

where $w^T \Pi = r_m - r_f$ where r_m is the total market return. Because the variance of the market portfolio is $\sigma_m^2 = w^T \Sigma w$, the equation can be rewritten to

$$r_m - r_f = \lambda\sigma_m^2 \quad \Rightarrow \quad \lambda = \frac{r_m - r_f}{\sigma_m^2} = \frac{\text{Sharpe ratio}}{\sigma_m} \tag{52}$$

where the Sharpe ratio from (10) has been employed. This is exactly the same as shown in (26) because r_m is also μ_m (any vector of excess returns).

2.7 Bayes theorem in the context for portfolio construction

The Black-Litterman model is based on a socalled Bayesian methodology as described in Satchell and Scowcroft (2000). In 1761 the english clergyman and statistician Thomas Bayes[4] died 59 years old. Two years later his friend Richard Price send his essay (it was written in the late 1740’s but never published): “*An Essay towards solving a Problem in the Doctrine of Chances*” together with many amendments and additions, to John Canton of the Royal Society of London, for publication. The phrase "doctrine of chances"

[4] https://en.wikipedia.org/wiki/Thomas_Bayes

means “the theory of probability”[5]. It took many years to recognize the work as important for inverse probability and in the past century the work of Thomas Bayes became recognized and is considered a cornerstone.

To begin with, we define all possible outcomes of an experiment as a sample space S. As an example, the sample space for rolling a die would be $\{1, 2, 3, 4, 5, 6\}$ and for tossing a coin it would be $\{\text{heads, tails}\}$ or $\{0, 1\}$. Any subset of a sample space is called an event, e.g. tossing a coin causes the event 0 or 1. Typically events can be expressed in terms of other events, e.g. A and B or C and D which are all sets of the sample space S. Then unions, intersections and complements are formed in the following way (see e.g. Johnson (2000)):

1. $(A \cup B)$ is the subset of S that contains all elements either in A, in B or in both. Therefore this is the notation for their union.

2. $(A \cap B)$ is the subset of S that contains all elements that are in both A and B. Therefore this is the notation for their intersection.

3. $\overline{A}$ is the complement of A, i.e. it is all elements of S that are not in A.

A conditional probability is the probability for something to happen, given that another event happens at the same time. Bayes theorem is used to calculate conditional probabilities and it is therefore closely related to statistics and probability theory. Bayes theorem states that (also in a version applied to portfolio theory):

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \qquad \Rightarrow \qquad P(E(\mathbf{r})|\boldsymbol{\Pi}) = \frac{P(\boldsymbol{\Pi}|E(\mathbf{r}))P(E(\mathbf{r}))}{P(\boldsymbol{\Pi})} \tag{53}$$

where $P(A)$ and $P(B)$ are the mutually exclusive observed probabilities that events A and B happens (without regard to each other, hence they’re unconditional probabilities). $P(A|B)$ is the (joint) probability that event A happens given that even B is true (conditional probability) and vice-versa for $P(B|A)$. An explanation for the validity of the rule is (using “set theory”[6]) that the probability for A to intersect B is

$$P(A \cap B) = \begin{matrix} P(B) \cdot P(A|B) \\ P(A) \cdot P(B|A) \end{matrix} \tag{54}$$

In order words, the probability that both events A and B happens is the probability for B to happen multiplied by the probability of A given B (or vice versa, it’s the same and therefore two expressions are shown on the right-hand-side, typically abbreviated RHS).

[5] https://en.wikipedia.org/wiki/An_Essay_towards_solving_a_Problem_in_the_Doctrine_of_Chances

[6] https://en.wikipedia.org/wiki/Set_theory

Setting the two expressions on the RHS in (54) to be the same and dividing by $P(B)$ to isolate for $P(A \cap B)$, it follows that the resulting equation is the same as (53).

When applied to portfolio theory, the lefthand side $P(A|B)$ is considered the posterior distribution. On the righthand side, $P(B|A)$ is formally known as the sampling distribution or the conditional distribution, $P(A)$ is the prior distribution and $P(B)$ is a normalizing constant:

1. $P(E(\mathbf{r})$ is a representation of the PDF (Probability Density Function) that expresses the prior views of the investor or fund manager.
2. $P(\boldsymbol{\Pi})$ is a representation of the PDF of equilibrium returns.
3. $P(E(\mathbf{r})|\boldsymbol{\Pi})$ is the result of Bayes theorem and it should be seen as a posterior forecast or mathematically, as a conditional PDF of equilibrium returns, given the prior views.

Consider the following simple example:

- P(rain) = 10%.
- P(cloud) = 25% (at a certain moment in the morning, there is 25% probability of having clouds).
- P(cloud|rain) = 75% (given that at a certain time there is rain, there is a 75% chance that there has been clouds in the morning).

Given that there is clouds, P(rain|cloud)=30% is the probability of getting rain:

$$P(\text{rain}|\text{cloud}) = \frac{P(\text{cloud}|\text{rain})P(\text{rain})}{P(\text{cloud})} = 75\% \cdot \frac{10\%}{25\%} = 30\% \tag{55}$$

This is an alternative explanation of Figure 5, but the Black Litterman model combines a prior equilibrium PDF for asset class returns with a the probability function for the views. Given that the views are correct, the model results in new combined returns, hence it mixes quantitative and subjective input data. The quantitative input data is based on historical data and therefore beliefs (=views) of the model should be considered as “adjustment factors”. Bayesian spam filtering can be used as a learning algorithm to detect spam emails based on something historically (anterior data) and “adjustment factors” (some words like “viagra” and “buy” clearly increase the probability of detecting messages as spam, i.e. the probability P(spam|buy Viagra)), just to name another example.

2.8 Partial conclusion on the theoretical background to portfolio optimization

This previous sections and pages introduces the Black-Litterman model and the whole framework, including important historical events from the litterature. Many of the equations provided are basic and elementary, but they have been introduced to ensure that the reader understands e.g. important differences between traditional modern portfolio optimization as we shall soon go into details about the Black-Litterman model.

No important results have been presented yet. This partial conclusion only summarizes all the necessary background information that is required to understand, for getting the best outcome from the following pages.

3 Example of data-analysis using traditional portfolio theory

Following is an illustration of the concepts of the classical Markowitz portfolio optimization method. A Python-script that downloads quotes from Yahoo Finance has been made, using Yahoo.com (2016) as data source. Figure 8 illustrates the downloaded data that is used to make up a portfolio. Selected data range is from Februar 1st 2012 to Februar 26th 2016 (one/some of the symbols didn't have data earlier than around 2012 and therefore this was chosen to be the start date). The data has been exported and later imported into Excel (daily "adj. close" has been used).

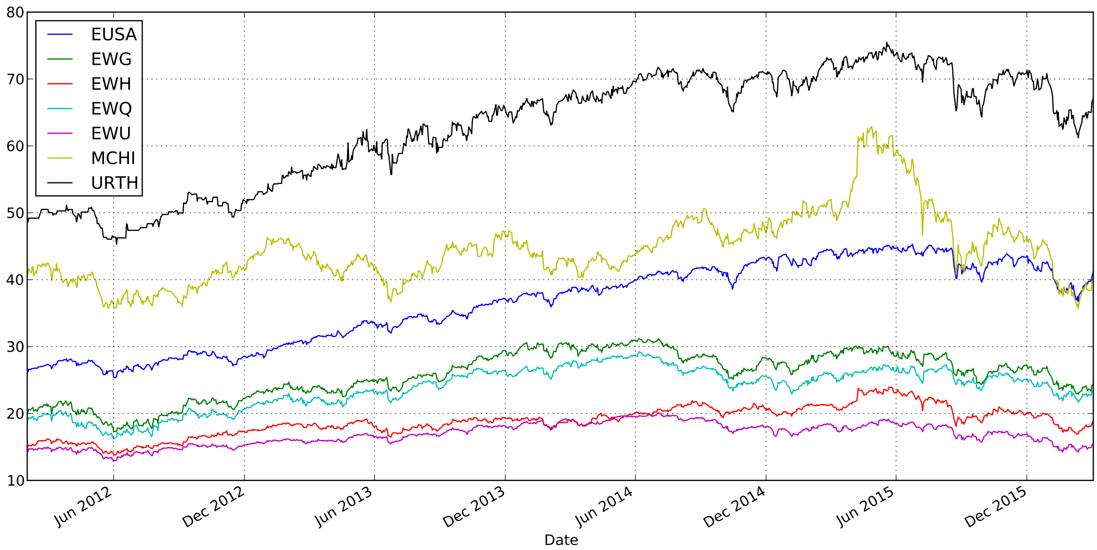


Figure 8: Representation of "raw" downloaded Yahoo Finance quotes.

EUSA	iShares MSCI USA Equal Weighted
EWU	iShares MSCI United Kingdom
EWG	iShares MSCI Germany
EWQ	iShares MSCI France
EWH	iShares MSCI Hong Kong
MCHI	iShares MSCI China
URTH	iShares MSCI World

Table 1: iShares tickers from Yahoo Finance (inkluder reference). MSCI World is included as a proxy for a diversified "global" benchmark portfolio.

Table 1 shows Yahoo Finance ticker symbols, used to create a new very simple "market" portfolio which is the sum of each asset class (one for each geographic region but excluding MSCI World because this will be used for "visual comparison"). The expected returns from the new market portfolio were calculated in Excel using (1) with equal probability or weight factor $p(s) = 1$. One could argue it is a mistake to construct the market portfolio that simple, with equal weights. For that reason, it makes sense to compare the market portfolio with MSCI World which is generally accepted as a good proxy for a developed market (DM) portfolio because it[7]

> *... captures large and mid cap representation across 23 Developed Markets (DM) countries*. With 1,649 constituents, the index covers approximately 85% of the free float-adjusted market capitalization in each country... DM countries include: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Hong Kong, Ireland, Israel, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, the UK and the US...*

MSCI World is not a perfect proxy for a global market portfolio, because the perfect proxy doesn't exist. Maybe it would be better to find a proxy that includes emerging markets (could be "MSCI Emerging Markets Index"), also because e.g. China and Brazil are normally considered large important markets – or more advanced methods could be used. However, it was decided to make this example, as simple as possible.

Figure 9 shows a comparison of the market proxy and market portfolio. For illustrative purposes the new market portfolio has been multiplied by 50% (only in this graph) and in this simple comparison, the objective is to compare iShares MSCI World with the new simple market portfolio. MSCI World then acts as a global stock benchmark, but is this a reasonable justifiable assumption? Figure 9 shows general tendencies in both MSCI World and in the new market portfolio that are quite similar. If the two curves are seen as two "mountains", many important peaks and troughs appear close to each other. A comparison of these two curves could maybe be considered to be a kind of "validation" of the new, simple market portfolio. The newly constructed market portfolio is highly influenced by China - maybe or probably too much. The reason is that the quoted chart values for MSCI China are higher than it is for the other securities. A discussion of correlation of developed and emerging markets (China/Asia) and how they influence on each other, including on the theoretical "global market portfolio" could perhaps come into play here. But in order to keep this relatively simply, a further discussion of this is out of the scope as it is not deemed very interesting in relation to an introduction to Markowitz portfolio theory.

[7] https://www.msci.com/resources/factsheets/index_fact_sheet/msci-world-index.pdf

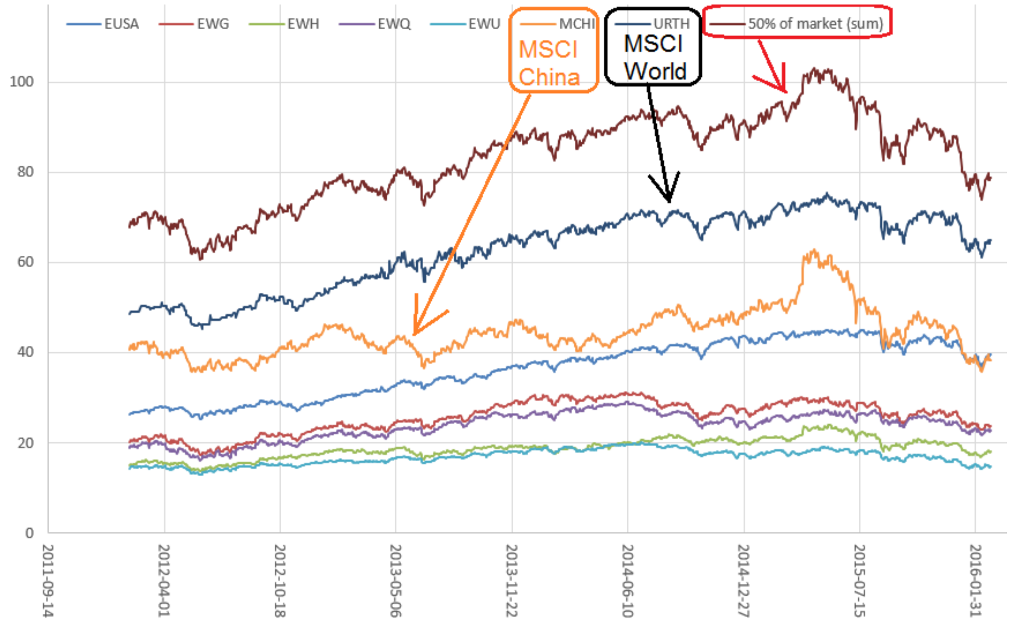


Figure 9: Construction of a very simple market portfolio, based on asset classes for geographic regions. Visually it seems like the benchmark or market proxy MSCI World is highly correlating with the new market portfolio, which is simply the sum of the other securities.

3.1 Risk premiums and risk-free interest rate, r_f

For estimation of the risk premiums, the risk-free interest rate is required. Initially risk-free interest rates were downloaded as "1-Year Treasury Constant Maturity Rate"[8] and adjusted such that invalid values became replaced by the average of the two adjacent values. However it was found that in addition to the fact that it is difficult to find good risk-free interest rate estimates:

- The 1-Year daily data didn't look correct (subjective), as in the end of 2015 and 2016 the rates increased very much.
- The overall interest levels seemed too high.
- It was deemed better to use either "3-month treasury bill"- or "effective federal funds rate" as input data for the risk-free interest rates (both curves are visually very similar) and calculate pro anno rates from these.

In the past few years we've seen clear effects on the financial crisis starting in end of year 2008 and therefore we know r_f is historically low. It has been decided to use "3-

[8] https://research.stlouisfed.org/fred2/series/DGS1#

month treasury bill" interest rates[9] as a qualified attempt to quantify values that should be used for estimation of the risk premiums. It is out of the scope of this report to use more advanced approaches, for obtaining the r_f-curve.

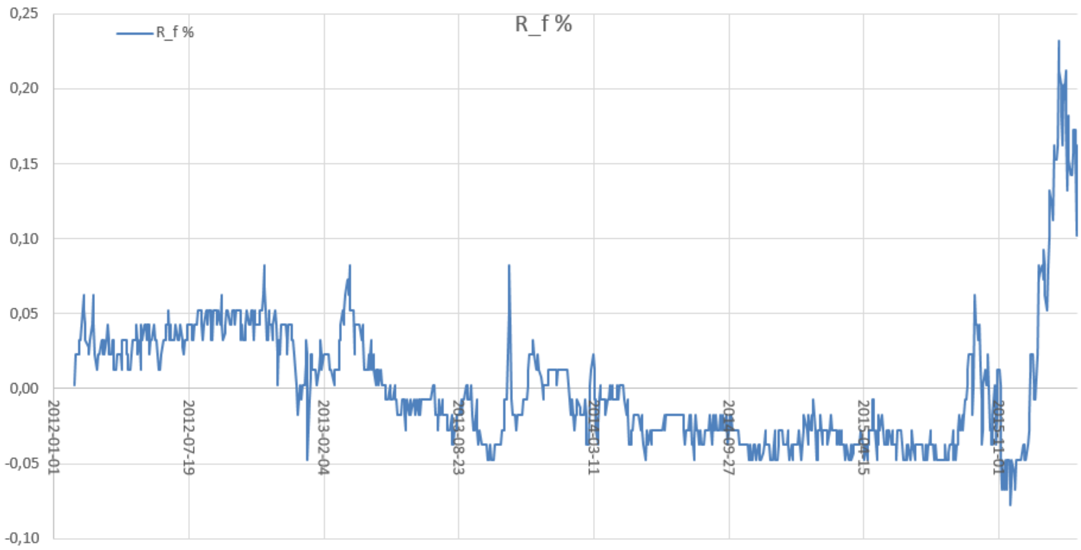


Figure 10: Representation of downloaded and adjusted into daily risk free interest rates. The increase in r_f around year 2016, towards the end seems too high. That was however neglected because it seems difficult to find good sources for risk free interest rates.

Figure 10 shows the resulting risk free-curve. It is assumed that the risk premiums on most securities should be around 5-15% p.a. but the downloaded 3-month treasury bill-rates gave negative risk premiums on securities when directly subtracted from the calculated returns. Hence the downloaded rates was divided by the number of days per 3 months (assuming 245 trading days or interest rate days per year divided by 4 corresponds to 61,25 days per 3 months). The result is that the 3-month interst rate converted to a "daily interest rate" in average is very close to 0% which is seen on the figure. This corresponds to the fact that very often these days we hear about governments or central banks launching new stimulus packages, in order to boost the economies in their regions and in order to try to prevent recession and other side-effects of the financial crisis. The idea with stimulus packages is based on the ideas of the british economist John Maynard Keynes who studied the effects of the great depression[10] in the 1930's. With low interest rates, central banks launched several quantitative easing packages[11], because they want to encourage people to invest more money in the economy.

[9] https://research.stlouisfed.org/fred2/graph/?id=DTB3,
[10] https://en.wikipedia.org/wiki/Great_Depression
[11] https://en.wikipedia.org/wiki/Quantitative_easing

The calculated risk premiums converted to pro anno, using this method where daily returns has been subtracted by daily risk premiums, multiplied by 245 trading days per year is in the interval between 0,97% (MCHI) and 10,62% (EUSA). The latter outperforms both the benchmark index (URTH with risk premium of 8,1%) and the new simple market portfolio (risk premium of 4,66%). These levels seem to be inside an acceptable range and also something inside this interval is expected for this period.

The pro anno (p.a.) expected returns or risk premiums are calculated as the daily average times the number of trading days per year:

$$E(r)_{\text{year}} = E(r)_{\text{day}} \cdot 245 \tag{56}$$

The pro anno (p.a.) standard deviations of risk premiums are calculated as the daily standard deviation times the number of trading days per year:

$$\sigma_{\text{year}} = \sigma_{\text{day}} \sqrt{245} \tag{57}$$

	$E(r)$ (% daily)	$E(r)$ (% p.a.)	σ (% daily)	σ (% p.a.)
EUSA	0,04	10,81	0,82	12,87
EWG	0,02	5,75	1,23	19,25
EWH	0,02	5,89	1,10	17,26
EWQ	0,03	6,28	1,26	19,72
EWU	0,01	2,27	1,02	16,02
MCHI	0,00	1,16	1,45	22,66
Market	0,02	4,85	1,00	15,59
URTH	0,03	8,30	1,04	16,28

Table 2: Average ($E(r)$) and standard deviation (σ) of daily returns (including converted values to p.a.). Simple market portfolio and URTH (MSCI World) are more diversified.

Tables 2 and 3 summarize simple statistical results from Excel. From Table 2 it can be seen that if we want to construct a portfolio with higher returns, maybe we need to be over-weighted in EUSA as the expected returns are almost 11% p.a. (assuming future expectations are made on the basis of historic data, which is normally a bad but simple assumption). At the same time, by coincidence EUSA almost has the lowest historic risk (it is also a too simple assumption, to use standard deviation on historic data as a measure of future risk – but it is simple).

High expected returns and low risk contribute positively to increased Sharpe ratio, (10) of the portfolio. The pro anno standard deviations are in the interval between 12% and 23% which is deemed acceptable or realistic. There is not much difference between

	$E(r)$ (% daily)	$E(r)$ (% p.a.)	σ (% daily)	σ (% p.a.)
EUSA	0,04	10,62	0,82	12,89
EWG	0,02	5,56	1,23	19,27
EWH	0,02	5,70	1,10	17,27
EWQ	0,02	6,09	1,26	19,73
EWU	0,01	2,07	1,02	16,03
MCHI	0,00	0,97	1,45	22,69
Market	0,02	4,66	1,00	15,61
URTH	0,03	8,10	1,04	16,29

Table 3: Average ($E(r)$) and standard deviation (σ) of daily risk premiums (including converted values to p.a.). Simple market portfolio and URTH (MSCI World) are more diversified.

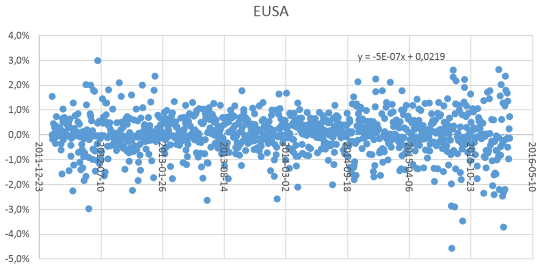
returns and risk premiums. The reason is that a 3-month interest rate has been converted to a pro anno interest rate, which is used a basis for a risk-free interest rate with 245 trading days per year. If the result should be different, the risk-free interest rate should be changed. With the risk premiums, regression has been made and the covariance matrix was calculated.

3.2 Markowitz portfolio optimization in practice

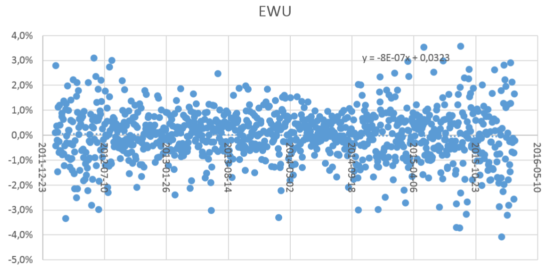
It would be a problem to using the normal distribution assumption, if the variance is e.g. small in the beginning and huge in the end. Figure 11 indicates that the time-dependency or variation of daily risk premiums is relatively constant. In other words, the variance is acceptable over the time-period, because it is relatively uniform over the considered time-frame from 2012 - 2016.

Figure 12 shows an example of regression analysis on the daily risk premiums for the price data. A linear regression is of the form $y = \alpha x + \beta$ where α and β are coefficients from the regression analysis. The α-values are close to 0 (they're insignificant) thus the expected returns heavily depend on the risk or volatility of securities in a given portfolio.

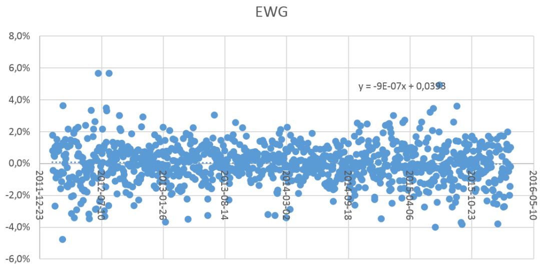
Table 3 gave the market premium risk of $E(R_M)$ =4,66% and from linear regression analysis, the volatility or contribution from β-values can be used to calculate new expected returns. The new expected returns consists of a constant contribution from α and a volatility or risk-depending contribution that linearly grows with increased risk-taking. Table 4 summarizes the contribution from α and β times the historic standard deviation, which follows the notion of the CAPM.



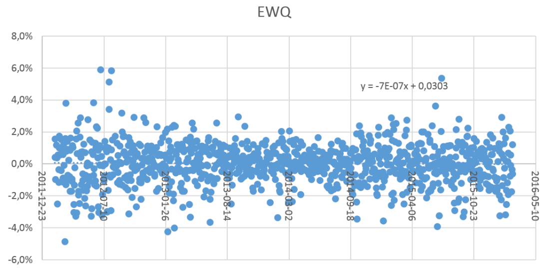
(a) USA Equal Weighted.



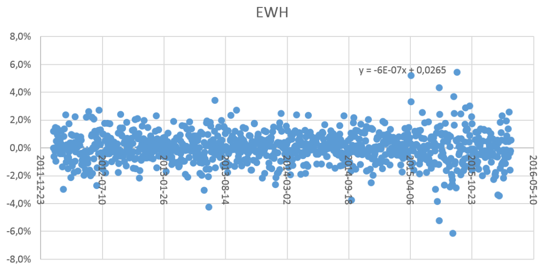
(b) United Kingdom.



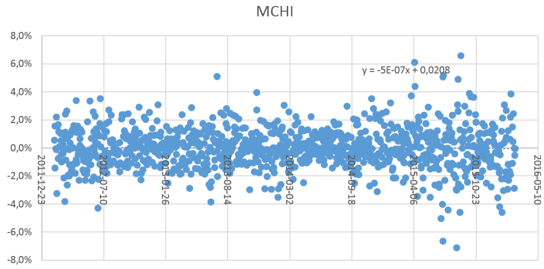
(c) Germany.



(d) France.



(e) Hong Kong.



(f) China.

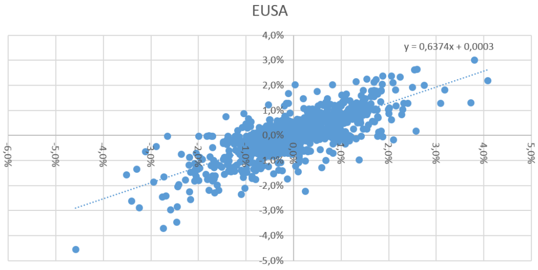
Figure 11: Daily risk premiums over the time period. The variance should not depend much on the time-period if the standard normal distribution assumption is used.

	EUSA	EWG	EWH	EWQ	EWU	MCHI	Market
α	0.031%	0.003%	0.006%	0.004%	-0.009%	-0.021%	0%
$\beta \cdot E(R_M)$	2.968%	4.937%	4.274%	5.071%	4.162%	6.028%	4.656%
$\sum$	2.999%	4.940%	4.280%	5.075%	4.153%	6.007%	4.656%

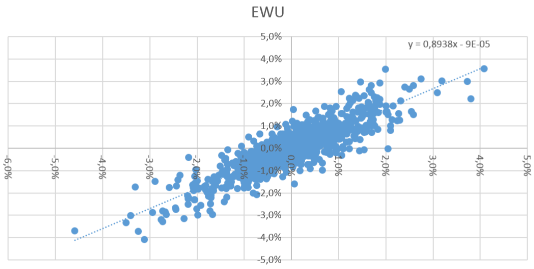
Table 4: Regression analysis on risk premiums for estimation of expected returns.

3.2.1 Calculation of optimal asset allocation weights

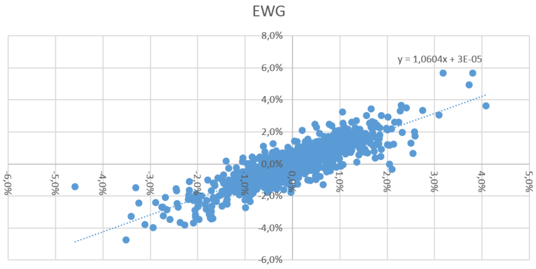
Black's method was described from page 10. Of great importance is the construction of the covariance matrix. Two methods have been used for calculating the covariance matrix:



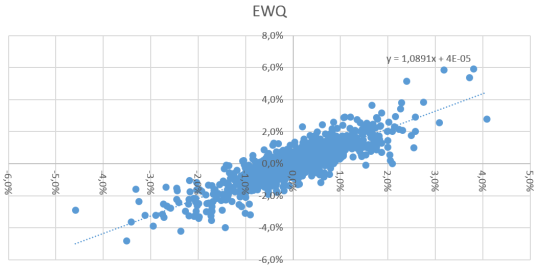
(a) USA Equal Weighted.



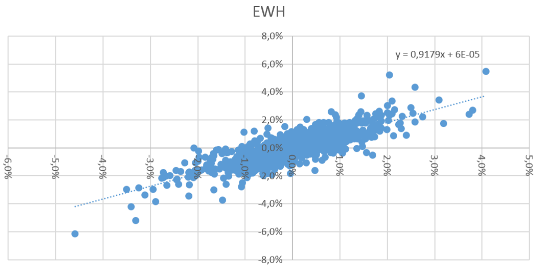
(b) United Kingdom.



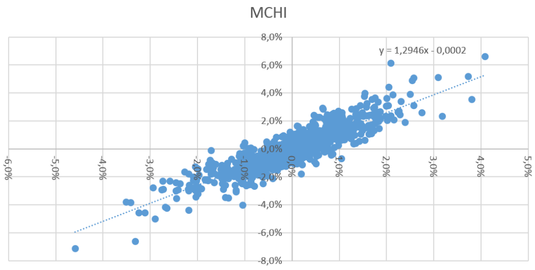
(c) Germany.



(d) France.



(e) Hong Kong.



(f) China.

Figure 12: Regression analysis on risk premiums using downloaded Yahoo Finance data (the x-axis shows risk premiums of the simple market portfolio).

- A general mathematical way (unbiased estimator) that assumes data is normally distributed:

$$\mathbf{Q_n} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})^T, \tag{58}$$

  where n is the number of elements in the random vector $\mathbf{x}$ and $\bar{x}$ is the average. This assumes the whole population was or is known. If that does not hold true, instead the sample covariance matrix can be calculated by dividing by $n-1$ instead of n. By experimentation it was found that the Excel "Analysis Toolpack" function for constructing a covariance matrix, assumes that the whole population is known. For large vectors or sample sizes, it does not matter.

- A single-index model which assumes that the stock returns i are $r_i = \alpha_i + \beta_i r_m + \varepsilon_i$

	EUSA	EWG	EWH	EWQ	EWU	MCHI	Market
EUSA	0,006773%	0,006351%	0,004882%	0,006459%	0,005871%	0,006742%	0,006336%
EWG	0,006351%	0,015139%	0,007643%	0,014453%	0,010325%	0,010680%	0,010541%
EWH	0,004882%	0,007643%	0,012168%	0,007825%	0,007073%	0,013678%	0,009124%
EWQ	0,006459%	0,014453%	0,007825%	0,015869%	0,010819%	0,011026%	0,010826%
EWU	0,005871%	0,010325%	0,007073%	0,010819%	0,010475%	0,009760%	0,008885%
MCHI	0,006742%	0,010680%	0,013678%	0,011026%	0,009760%	0,020985%	0,012869%
Market	0,006336%	0,010541%	0,009124%	0,010826%	0,008885%	0,012869%	0,009940%

(a) "Standard" or normal covariance matrix, for normal distributed data.

	EUSA	EWG	EWH	EWQ	EWU	MCHI	Market
EUSA	0,006774%	0,006712%	0,005810%	0,006894%	0,005658%	0,008195%	0,006330%
EWG	0,006712%	0,015136%	0,009666%	0,011468%	0,009412%	0,013633%	0,010530%
EWH	0,005810%	0,009666%	0,012167%	0,009927%	0,008147%	0,011801%	0,009115%
EWQ	0,006894%	0,011468%	0,009927%	0,015865%	0,009666%	0,014001%	0,010815%
EWU	0,005658%	0,009412%	0,008147%	0,009666%	0,010472%	0,011491%	0,008876%
MCHI	0,008195%	0,013633%	0,011801%	0,014001%	0,011491%	0,020977%	0,012856%
Market	0,006330%	0,010530%	0,009115%	0,010815%	0,008876%	0,012856%	0,009931%

(b) Covariance matrix constructed using the "index model".

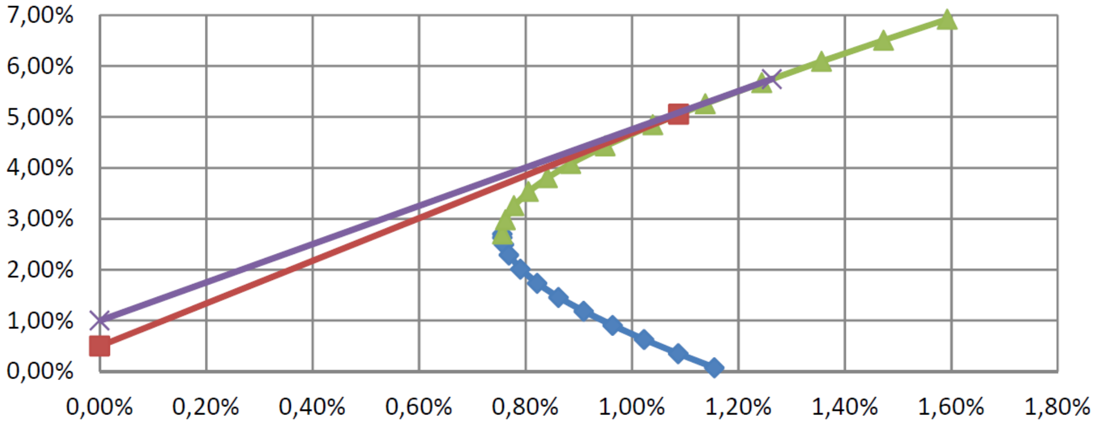
	EUSA	EWG	EWH	EWQ	EWU	MCHI	Market
EUSA	1,00	0,63	0,54	0,62	0,70	0,57	0,77
EWG	0,63	1,00	0,56	0,93	0,82	0,60	0,86
EWH	0,54	0,56	1,00	0,56	0,63	0,86	0,83
EWQ	0,62	0,93	0,56	1,00	0,84	0,60	0,86
EWU	0,70	0,82	0,63	0,84	1,00	0,66	0,87
MCHI	0,57	0,60	0,86	0,60	0,66	1,00	0,89
Market	0,77	0,86	0,83	0,86	0,87	0,89	1,00

(c) Corresponding correlation matrix for method (a).

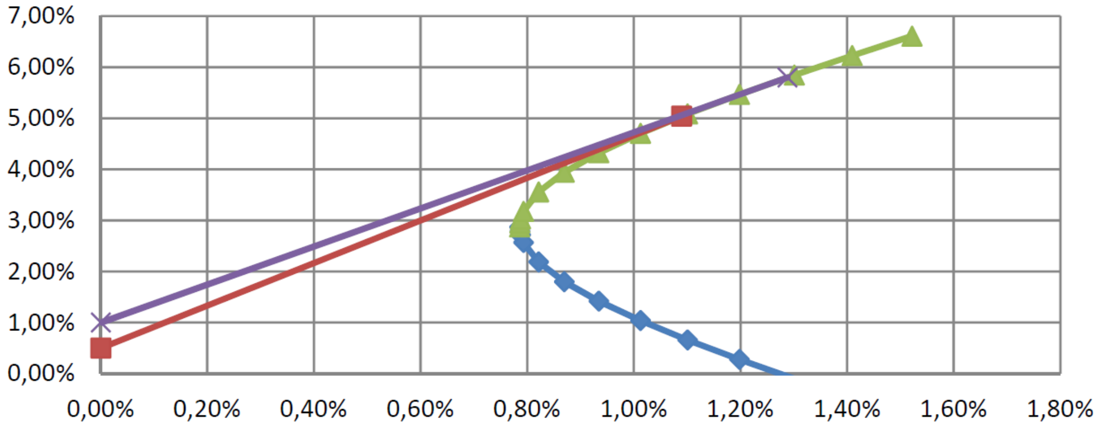
Figure 13: Two methods for creating the covariance matrix of centered risk premiums, together with the correlation matrix that corresponds to the first covariance matrix.

where α_i are the returns that exceeds the risk-free rate, $\beta_i r_m$ is a product that expresses the systematic movement with the market and ε is firm-specific unsystematic risk. The procedure is to follow the steps:

1. The variance of the market m is simply taken as the squared standard deviation from Excel (σ_m^2).
2. The variance of security or firm i is $\beta^2\sigma_m^2 + \varepsilon^2$ where ε is "standard error" from regression analysis in Excel (standard deviation of the residuals). In other words, this is the sum of a systematic and firm-specific component.
3. The covariance of market and security i is $\beta_i\sigma_m^2$.



(a) Efficient frontier calculated using the normal standard covariance matrix.



(b) Efficient frontier calculated using the index model covariance matrix.

Figure 14: Comparison of efficient frontier calculated using both covariance matrices.

4. Other covariances of security i and j is $\beta_i\beta_j\sigma_m^2$.

These methods produce the covariance matrices shown in Figure 13. Figure 13a and 13b are important for everything else and sometimes known as variance-covariance matrices (because the variance is in the diagonal). However, the covariances are not always intuitive to understand and for that reason people tend to prefer the normalized versions of these matrices such that correlation coefficients are in the interval $[-1 \leq \rho \leq 1]$. An example of that is made by constructing a diagonal matrix whose elements are the square root of the covariance matrix. This matrix is inverted and pre-multiplied by the matrix product of the covariance matrix and the matrix itself. It should be the same as dividing each element of the covariance matrix by $(\sigma_i\sigma_j)$. Figure 13c shows the result of converting the first covariance matrix into a correlation matrix. It can be seen that there are no negative correlation coefficients. This means that all regional markets to a certain degree are positively correlated.

A correlation coefficient of 0 means there is no linear dependence on two assets or securities while a negative correlation coefficient means there is a negative dependence. In other words if e.g. investors see a bull market in Germany (EWG) but a bear market in e.g. asian or Hong Kong (EWH), the correlation would be negative. Negative correlation coefficients are fine for reducing the risk by diversification of the portfolio.

Next step is to construct two portfolios with two different constants and the first constant was chosen to be $c_1 = 0,50\%$ while the second constant was chosen to be $c_2 = 1,00\%$. By applying (22) and normalizing asset weights (should sum to 100%), two normalized sets of optimal portfolios are found (one for each covariance matrix). Short-selling is allowed, thus negative weights can appear in the solution to the unconstrained optimization problem. The specific results and weights are not shown here but Figure 14 shows the straight portfolio lines that make up a tangent point on the resulting efficient frontier.

3.2.2 Moving away from traditional mean-variance analysis

It is generally accepted in the litterature, that the traditional mean-variance analysis has some weaknesses that portfolio managers should address. The most common weakness is that it produces extreme portfolio weights of both long and short positions. The problem can be reduced by including constraints on the optimization problem.

Another problem is that the solution in terms of weights is very sensitive to estimation errors in return vector and covariance matrices. This can be examined by slightly changing the vector of expected returns as investigated in Best and Grauer (1991). The authors showed that the sensitivity of the portfolio weights to mean return vector errors grows, as the ratio of the largest to the smallest eigenvalue of the covariance matrix grows. This is unfortunate as the ratio of largest to smallest eigenvalue typically grows when the number of assets increase, when the number of sample observations is kept the same. In other words, large errors are expected for large portfolios with few observations. The above analysis was performed on daily data, so the number of observations or samples for the covariance matrix was relatively high. If the only available data set was weekly or monthly end-of-day quotes, we should expect a high sensitivity of the covariance matrix and hence unrealistic portfolio weights.

There is also a discussion about, if we can assume data to follow a normal distribution which is a necessary assumption to consider. Empirical evidence has later led many people away from this assumption towards heavy-tailed distributions, see e.g. Rachev (2003).

Finally, investors typically might want to include subjective views into the equations which is difficult or impossible to do, without mixing objective market data with other subjective information. This was a conclusion in the article [Black and Litterman (1992):

> *... our approach allows us to generate optimal portfolios that start at a set of neutral weights and then tilt in the direction of the investor's views.*

3.3 Partial conclusion on the use of traditional portfolio theory

The previous paragraphs describes a method where real input data from a Python program is used to illustrate the solution to an unconstrained optimization problem. A few tickers have been selected, based on different geographic regions and a very simple market portfolio was constructed.

Issues concerning the risk free interest rate and premiums are discussed. Data have been post-processed in Excel and simple statistical information is used and presented. Demonstrations of regression analysis have been performed and two covariance matrices have been made. Finally, the efficient frontier using both covariance matrices have been graphically illustrated using background information from the previous section.

4 Practical use and validation of the Black-Litterman model

The following pages describe how the BL portfolio optimization model has been validated and there are examples and comments on the results.

4.1 Validation of the model in He and Litterman (2002)

Before using the model with our own data set(s) we need to verify or validate the model. The following input data with 2 views:

$$P = \begin{bmatrix} 0 & 0 & -0.295 & 1 & 0 & -0.705 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} \tag{59}$$

$$Q = \begin{bmatrix} 5\% \\ 3\% \end{bmatrix} \tag{60}$$

have been used to recreate the results in (He and Litterman, 2002, Table 5) as shown in Table 5. The first two columns show the countries and thereby the meaning of the P-vector. The following column μ is the posterior estimate of the means as calculated by either (29) or (32).

Country	P_1 (%)	P_2 (%)	μ (%)	w^* (%)	w_{eq} (%)	$\frac{w^*-w_{eq}}{1+\tau}$ (%)
Australia	0	0	4.422	1.524	1.6	0
Canada	0	100	8.73	41.86	2.2	39.77
France	-29.5	0	9.48	-3.409	5.2	-8.362
Germany	100	0	11.21	33.58	5.5	28.34
Japan	0	0	4.616	11.05	11.6	0
UK	-70.5	0	6.972	-8.174	12.4	-19.98
USA	0	-100	7.482	18.8	61.5	-39.77

Table 5: Results with 2 views that matches (He and Litterman, 2002, Table 5).

The w^*-column is the optimal portfolio or unconstrained mean variance weights as found using (28) but using the posterior covariance matrix as calculated by (34), i.e. $w_{eq} = (\lambda\Sigma_p)^{-1}\mu$. It is obvious that for the new returns, one cannot use the original covariance matrix for this calculation. There is a longer explanation for the reason for why the sum of w^* is not 100% in He and Litterman (2002). It can be derived that 1+τ is a scaling factor which is seen from:

$$w^* = \frac{1}{1+\tau}\left(w_{eq} + P^T \times \Lambda\right) \tag{61}$$

where Λ is a weight vector that is not shown here, because it does not bring that much new relevant information in here. The sum of $w^* = 0,95238$ is exactly equal to $1/(1+\tau)$, hence the calculation is correct.

The final column shows the change between w^* and the equilibriumn weights w_{eq}. The latter was initially defined and has been inserted as the second-last column, in order to easier interpret the results. It can be seen that the column has been scaled by $(1+\tau)$ and the sum is 0. Where w_{eq} sums to 1 it is maybe a little counter-intuitive that w^* only summates to around 95%. However it is very easy to see the influence of view 2 with a change of around +40% in favor of Canadian securities at the expense of around -40% US securities. Even though the second view only deals with an expected excess return of 3% the variance is lower for the second view ($\omega_1 = 0,001065$ and $\omega_2 = 0,000852$), hence the precision is larger and the view has a greater impact. From the first view, we see that the weights for Germany is now around 28% higher than when comparing to the market equilibrium situation.

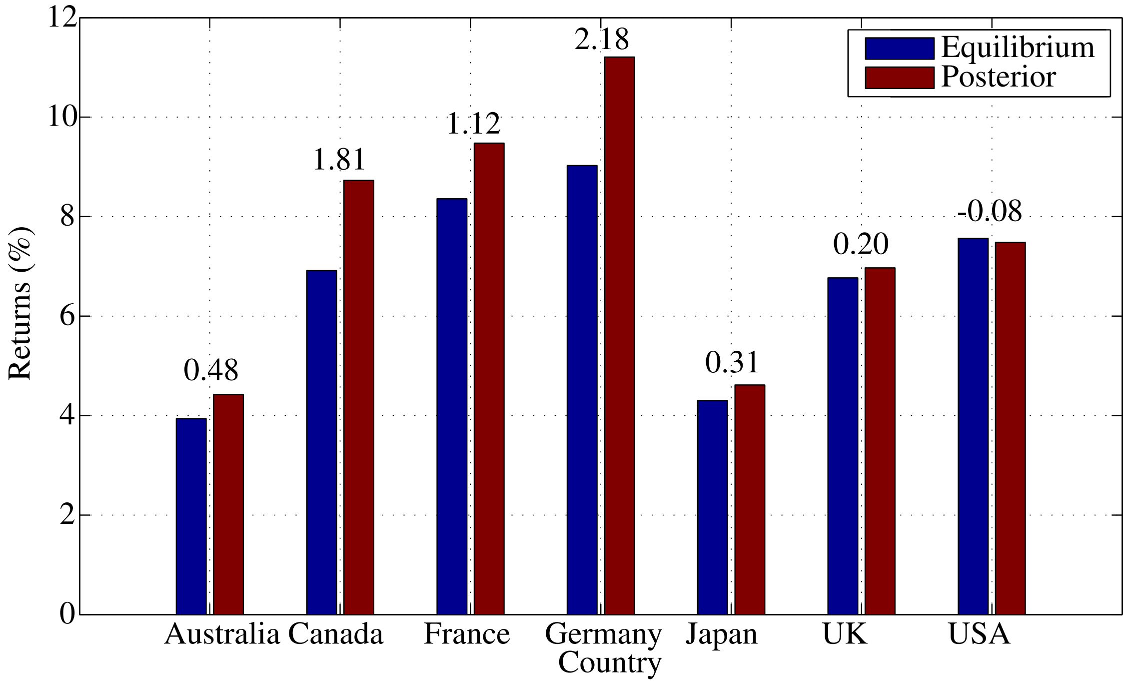


Figure 15: Bar plot of equilibrium and posterior returns. The numbers on the top denote the difference between posterior and equilibrium returns.

Likewise Figure 15 again shows the influence of the views Q and how it affects the assets through the P-vector. It is clear that view 1 has a strong positive impact on Germany which has returns that increases the most (from 9,03% to 11,21%). Also just by looking at the P-vector it is clear that Canadian securities should outperform American, hence bullish view on Canada and the returns for Canada increases from 6,92%

to 8,73%, which is second-best in class. The investor view on American securities is affected by a view factor -100% with $Q_2 = 3\%$ and the returns decrease from 7.56% to 7.48%. There are no negative returns. But the negative weights on France and UK should be understood as taking small short-sales positions. This figure does not tell anything about the risk taken. Again, the optimal portfolio is one that maximizes the utility function (49) with two opposite terms: A positive term consisting of $w^T\Pi$ which is the weighed returns subtracted from something that is proportional to the variance as $w^T\Sigma w = \sigma_m^2$. For this reason it is not straightforward only to look at the equilibrium and views and imagine what the posterior returns will be.

4.1.1 Reproduction of results of the open-source Akutan finance project

In Walters (2008) there is a reference to a open source finance project called Akutan, hosted at sourceforge.net[12]. It is a project written in Java by Jay Walters who contributed significantly by creating the `http://blacklitterman.org/`-website and also by publishing several freely available in-depth publications about quantitative finance algorithms. Last update was in April 2013 and it took some days to make it run on a recent, newer system. Some papers e.g. (Walters et al., 2014, p.27) illustrate the Bayesian methodology by 3 normal (probability) distributions with specified mean and variance (in the following the author will refer to “distribution/distributions” instead of maybe the more academic correct and longer “probability density function” or PDF). The reason for spending time doing implementing this, was that maybe it would be more intuitive or easy to interpret results, with a visual approach. It would in any case add an extra perspective to the “blending” of quantitative data and views.

	μ (%)			σ		
Country	Prior	Cond.	Post.	Prior	Cond.	Post.
Australia	3.938	3.938	4.422	0.036	0.146	0.164
Canada	6.915	6.915	8.730	0.045	0.146	0.207
France	8.358	6.883	9.480	0.055	0.146	0.254
Germany	9.027	14.03	11.21	0.061	0.146	0.276
Japan	4.303	4.303	4.616	0.047	0.146	0.215
UK	6.768	3.243	6.972	0.045	0.146	0.205
USA	7.560	7.560	7.482	0.042	0.146	0.191

Table 6: Results of the Akutan program (using only view number 1, because more than one dimension is more difficult to visualize). Yellow indicates where the mean distribution differs from the prior.

Table 6 shows the results of the program using 1 dimension, i.e. only the first view.

[12] Available May 2nd 2016 at: `https://sourceforge.net/projects/akutan/`

The prior and posterior mean values are also graphically shown in Figure 15. The mean value of the conditional distribution is the prior distribution added together with a correctional term which is a function of P and Q, hence the views. This makes sense, as if an investor or portfolio manager e.g. has a bullish view on US securities, that should definately affect the conditional mean value in upward direction. Likewise with a bearish view, then the mean value of the conditional distribution should decrease to below that of the prior distribution.

Equation (60) gives the view-vector Q and the P-vector from (59) couples individual assets to the views. The figure only visualizes the effect of first view, hence $Q_1 = 5\%$ and in the P-vector there's a factor (+)1 for Germany. The middle yellow cell in Table 6 exactly tells that the difference from 14,03% to 9,027% is 5% if we neglect that the table has been made using only 4 significant digits. Hence the correctional term is +5% for the conditional distribution for Germany. The P-vector additionally explains that the correctional term for France and UK is negative. Using the same methodology, the correctional term for France is -29,5%·5% $= -1,475\%$ and for UK it is even worser: -70,5%·5% $= 3,525\%$. These numbers are exactly the differences between the mean of the prior and conditional distributions in Table 6.

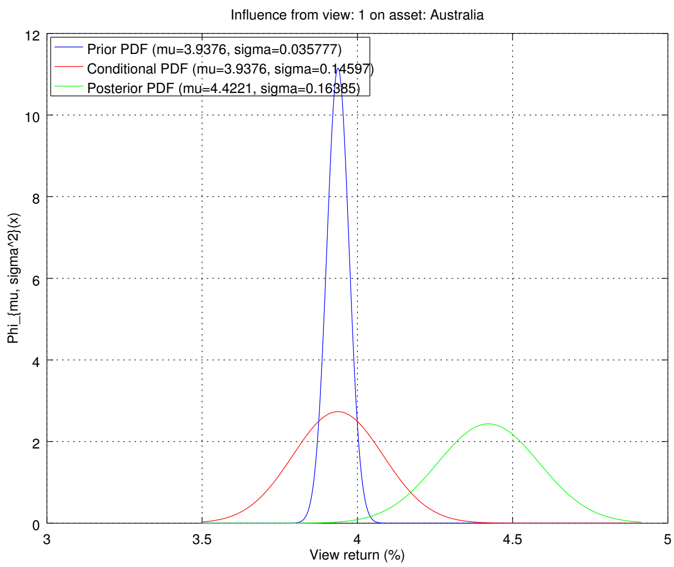


Figure 16: Influence of view 1, on the prior, conditional and posterior PDF for Australia as made in the Akutan project. The other 6 asset classes are shown in Figure 17.

Attention should also be on the variance or standard deviation of each of the distributions. With model input as in He and Litterman (2002), the standard deviation has been calculated as:

Prior PDF:

$$\sigma_i = \sqrt{\tau \, \text{diag}(\Sigma_i)}, \quad \text{where } \Sigma \text{ is the original covariance matrix.} \tag{62}$$

Conditional PDF:

$$\sigma_i = \sqrt{\frac{\omega_i}{\tau}}, \quad \text{where } \omega_i \text{ is diag}(\Omega) \text{ as seen in (35).} \tag{63}$$

Posterior PDF:

$$\sigma_i = \sqrt{\text{diag}(\Sigma_{p,i})}, \quad \text{where } \Sigma_p \text{ is the posterior covariance matrix..} \tag{64}$$

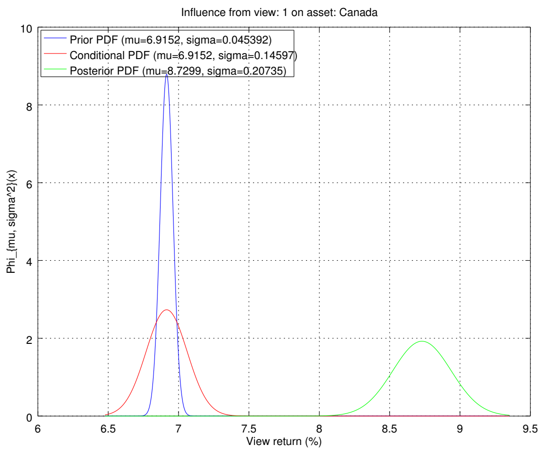
The posterior covariance matrix was given by (34).

Therefore in Table 6 the prior values for σ is merely a function of the diagonal in the original covariance matrix. Because $\omega_1 \approx 1,07 \cdot 10^{-3}$ and because $\tau = 5\%$ is a constant, the conditional standard deviation is 0,146 and the same for all assets. Finally, the posterior values for σ is the square root of the posterior covariance matrix Σ_p.

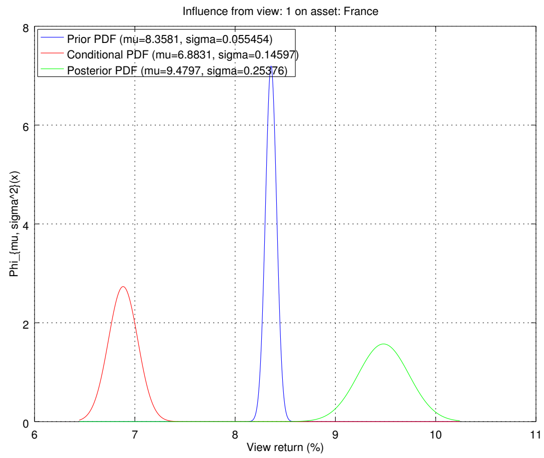
While it is clear that if views (the conditional PDF) are positive on certain assets or assets classes, one might expect that the posterior mean is always higher. Figures 16 and 17 however visually illustrate that this is not always the case. One should remember that the views are only a part of the equation system to be solved. If we remember the "master formula" given as (29) and reproduced below:

$$E(r) = \left[(\tau\Sigma)^{-1} + \underbrace{P^T \Omega^{-1} P}_{\text{view-dependent}} \right]^{-1} \left[(\tau\Sigma)^{-1}\Pi + \underbrace{P^T \Omega^{-1} Q}_{\text{view-dependent}} \right] \tag{65}$$

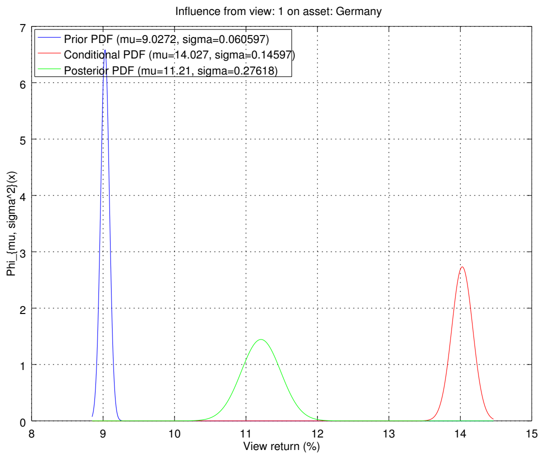
It is clear to see that Ω only partly affects the posterior results $E(r)$. And therefore the figures are not so easy to interpret, especially not in multiple dimension asset-space. Another topic that makes things a little hard to figure out is the effect of correlation on assets. If e.g. asset number 3 and number 4 are strongly correlated and there is a bullish view only on asset number 3. What will happen to the expected, posterior returns for asset number 4? While it is clear that the expected returns on asset number 3 increase, dependending on the covariance matrix it is assumed that asset number 4 will also have increased returns. What if there is strong correlation between asset number 3 and number 4 but the view on asset 3 is bearish and on asset 4 it is bullish. Maybe the effects of the views will then cancel each other out, even for high confidence in the views?



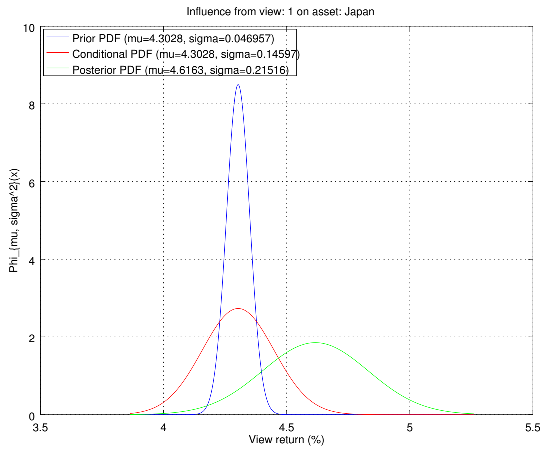
(a) Canada



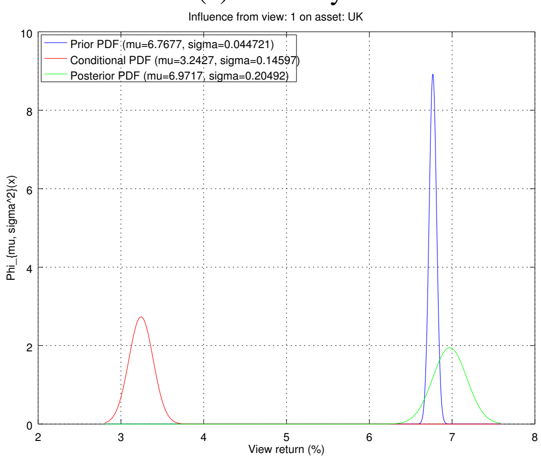
(b) France



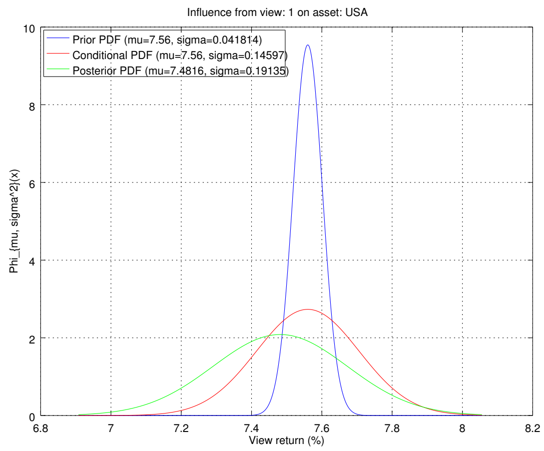
(c) Germany



(d) Japan



(e) UK



(f) USA

Figure 17: View 1 influence on assets (input data from He and Litterman (2002)).

4.2 Validation of the model in Idzorek (2005)

The theory of the method in Idzorek (2005) was described from page 2.5.4. Idzorek uses a risk-aversion coefficient (λ) of approximately 3.07 and uses an example with 8 assets. The implied equilibrium return vector is:

$$\Pi = \{0.08\% \quad 0.67\% \quad 6.41\% \quad 4.08\% \quad 7.43\% \quad 3.70\% \quad 4.80\% \quad 6.60\%\}^T \quad (66)$$

There are 3 views and the Q-vector is:

$$Q = \begin{bmatrix} 5,25\% \\ 0,25\% \\ 2\% \end{bmatrix} \quad (67)$$

The corresponding P-vector is shown as the first 3 columns in Table 7. The following column $E(r)$ is the new combined return vector while w^* is the vector of new weights. Equation (32) was used to calculate $E(r)$ and the posterior weights w are calculated using (28) where $\mu = E(r)$. The other weights w_{eq} and $w_{100\%}$ are plotted together and shown in Figure 7 on page 19. The results match (Idzorek, 2005, Table 2.6). The last column is the vector of aggregate or implied confidence level. It is connected to the diagonal of Ω because it expresses the confidence level. It has been calculated using (44) and results matches (Idzorek, 2005, Table 2.8) fully, when taken into account that the P-matrix does not have any links to the last asset class (int'l emerg. equity). While the last column expresses the implied confidence level in asset space (as a vector of size n), it is also possible to calculate the implied confidence level in view space (as a vector of size n).

Asset class	P_1 (%)	P_2 (%)	P_3 (%)	$E(r)$ (%)	w^* (%)	A.C. (%)
US Bonds	0	-100	0	0.067	29.89	43.1
Intl Bonds	0	100	0	0.499	15.58	43.1
US Lg Grth	0	0	90	6.505	9.37	33.1
US Lg Value	0	0	-90	4.326	14.81	33.1
US Sm Grth	0	0	10	7.551	1.04	33.1
US Sm Value	0	0	-10	3.942	1.64	33.1
Intl Dev Eq	100	0	0	4.937	27.78	32.9
Intl Emg Eq	0	0	0	6.845	3.49	-

Table 7: Default (50%) confidence specified. Last column is the aggregate (implied) confidence level as calculated by (44). The results exactly match (Idzorek, 2005, Table 2.8).

The implied confidence level in view space is shown in Table 8 together with the diagonal elements ω/τ which is also the (unscaled) variance of the view ($P\Sigma P^T$). The

	View 1	View 2	View 3
Implied confidence	0.50	0.50	0.50
ω/τ (%)	2.84	0.56	3.46

Table 8: 50% confidence specified. The variances for the views are the same as shown in (Idzorek, 2005, Table 2.4).

implied confidence in view space is 50% for each view, which can be calculated using a modified version of (45):

$$C = \frac{1}{1+\alpha} \qquad \text{where} \quad \alpha = \text{diag}\left(\frac{\Omega}{P\Sigma P^T}\right) \tag{68}$$

Next, we want to change Ω to reflect that the user-specified confidence level has been used. We choose to have 25% confidence in view 1, 50% confidence in view 2 and 65% confidence in view 3. The results will now change from those shown in Table 7 to those shown in Table 9.

Asset class	P_1 (%)	P_2 (%)	P_3 (%)	$E(r)$ (%)	w* (%)	A.C. (%)
US Bonds	0	-100	0	0.069	29.63	42.1
Intl Bonds	0	100	0	0.497	15.84	42.1
US Lg Grth	0	0	90	6.279	8.94	38.3
US Lg Value	0	0	-90	4.217	15.24	38.3
US Sm Grth	0	0	10	7.280	0.99	38.3
US Sm Value	0	0	-10	3.830	1.69	38.3
Intl Dev Eq	100	0	0	4.767	26.03	16.9
Intl Emg Eq	0	0	0	6.627	3.49	-

Table 9: User-specified confidence levels are now specified to each view.

There is least confidence in view 1 and most confidence in view 3. The implied confidence level decreased slightly for the first two assets (from 43,1% to 42,1%). For all assets connected to view 3, the implied (aggregate) confidence level increased from 33,1% to 38,3%. Also for the first view, the implied confidence decreased from 32,9% to 16,9% (for Intl. Dev. Eq). This makes perfect sense, as now view 3 is the most confident and view 1 is the least confident. The new $E(r)$ and weights are included for reference but the most interesting is what happens to the aggregate implied confidence level when the confidence for each of the views is modified.

By comparing Table 8 and 10 it is easy to see the change in implied confidence for each view, also changed the diagonal of Ω. As the diagonal is the variance of the view, this increased for view 1 and decreased for view 3. It is the same for view 2, as the

	View 1	View 2	View 3
Implied confidence	0.25	0.50	0.65
ω/τ (%)	8.51	0.56	1.86

Table 10: User-specified confidence levels specified for each view.

implied confidence is still 50%. In other words, view 1 is more uncertain now and view 3 is more accurate. By playing with these numbers, it is possible to fine-tune the results in a more intuitive way and also to tune each view individually. This is deemed more intuitive and better, than by changing the uncertainty parameter τ which also affects all diagonal elements with the same factor at the same time, as described in He and Litterman (2002). We can also take the analysis a bit deeper, by investigating what happens for each view.

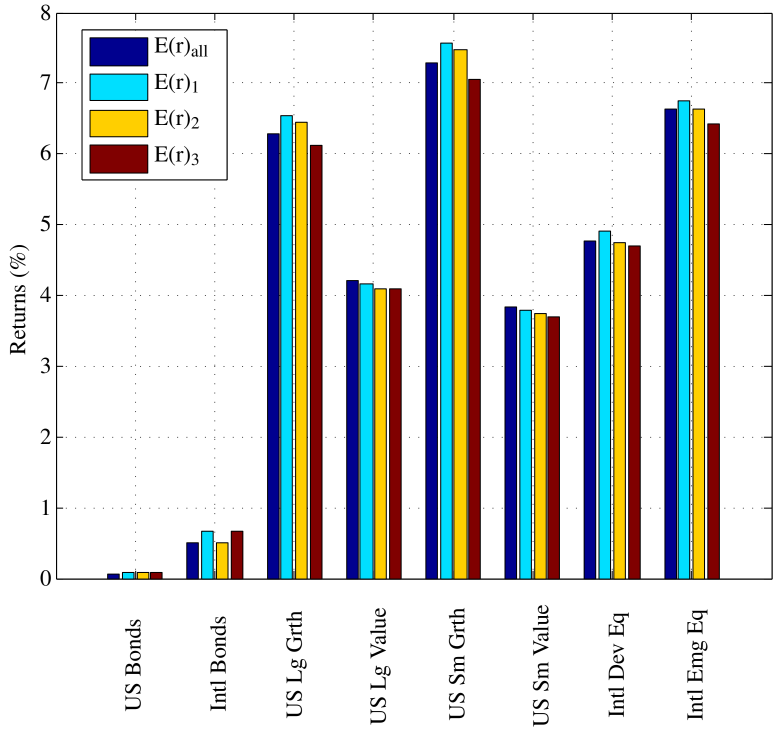
4.2.1 View details for confidence level 25%, 50% and 65%

Table 11 summarizes the effect on the posterior returns, based on each individual view. The table can be constructed by using a program where only a single view is passed into the Black Litterman model, at a time. This means that Ω is not a matrix anymore because with a single view, it becomes a scalar. The parameters that change is P, Q and Ω. The constant parameters are the risk-aversion coefficient (λ), equilibrium weights w_{eq} and the variance-covariance matrix.

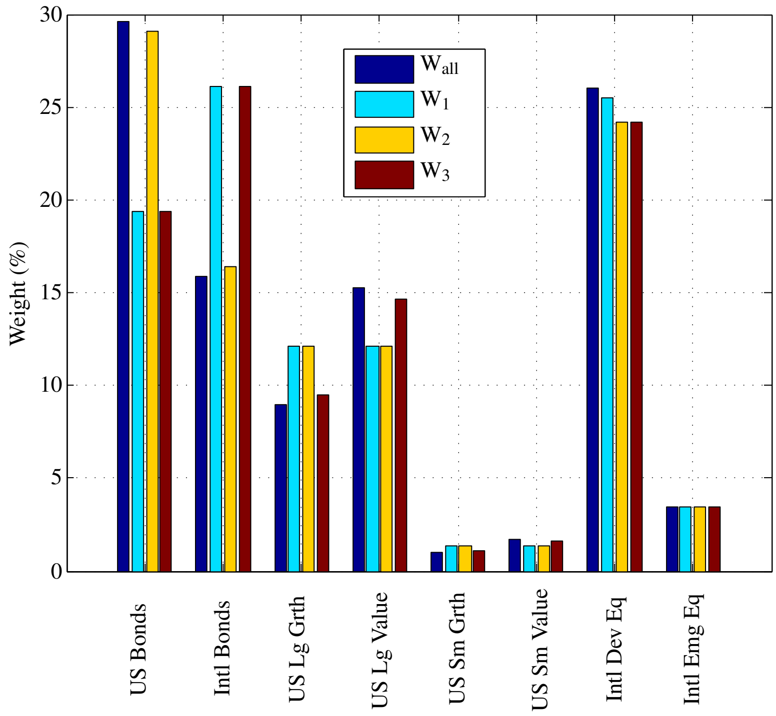
Asset class	P_1	$E(r)_1$	AC_1	P_2	$E(r)_2$	AC_2	P_3	$E(r)_3$	AC_3
US Bonds	0	0.08	0	-100	0.07	50	0	0.08	0
Intl Bonds	0	0.67	0	100	0.49	50	0	0.68	0
US Lg Grth	0	6.55	0	0	6.45	0	90	6.13	65
US Lg Value	0	4.17	0	0	4.09	0	-90	4.10	65
US Sm Grth	0	7.56	0	0	7.47	0	10	7.06	65
US Sm Value	0	3.79	0	0	3.74	0	-10	3.69	65
Intl Dev Eq	100	4.92	25	0	4.75	0	0	4.69	0
Intl Emg Eq	0	6.74	0	0	6.64	0	0	6.43	0

Table 11: Confidence levels specified. The P, $E(r)$ and AC-vector elements are in percent.

The P-vector from Table 11 is given and the aggregate confidence level from (68) is distributed to those assets involved in the view. The resulting posterior returns are also shown, however in this case it makes more sense to illustrate the influence of each view graphically, as shown in Figure 18. The figure contains both an illustration of $E(r)$ and the weights for all and each of the 3 views. The following can be seen from Figure 18a:



(a) Total $E(r)$ from all views together with the individual $E(r)$ from each view.



(b) Total weights from all views together with the individual ones from each view.

Figure 18: The influence of all and each of the 3 views, using Idzorek (2005).

1. View 1 with confidence 25% is bullish on “Intl Dev Eq” and it is an absolute view, unlike the two other views. $Q_1 = 5,25\%$ which is the highest of the 3 views. The equilibrium returns for this asset is only 4,8% but with this positiv view it increases to 4,92%. This is higher than with the other views including the total combined expected return for all views. It is expected that with the highest value of Q we should see an effect and we also do.

2. View 2 is relative with $Q_2 = 0,25\%$ which is the lowest of all. It indicates the view is insignificant in comparison with view 1. It is bearish on “US Bonds” and bullish on “Intl Bonds” with 50% confidence. If we look at Figure 18a it is difficult to not use Table 11 instead for the first asset, because the scale is very small. The equilibrium returns are 0,08% and 0,67% respectively. For view two, we see $E(r)$ decreases to 0,07% and 0,49% respectively. The expected returns for both assets decrease. The second value of Q is 0,25% so it is below the equilibrium return values. This likely explains the observed decrease.

3. View 3 with $Q_3 = 2\%$ is also relative. Strongly positive on “US Lg Grth” and strongly negative on “US Lg Value” which have equilibrium returns of 6,41% and 4,08% respectively. The view is lower than both equilibrium values. For the first asset, the expected returns decrease to 6,13% but for the second asset the expected returns increase to 4,10% (an increase, when comparing to the equilibrium returns). It is a maybe not always easy to make a conclusion on the optimization results, probably due to correlation between assets which is an effect of the covariance matrix. That the view is also weakly positive on “US Sm Grth” and weakly negative on “US Sm Value” is deemed not important because only 10% of the view is on these two assets and 10% of $Q_3 = 2\%$ is 0,2%. That is insignificant.

The following can be seen from Figure 18b which shows the final weights, after applying the BL-model:

1. The first view is absolute an should have an effect on “Intl Dev Eq”. Of all views it can be seen that view 1 has the highest weight (light-blue color under “Intl Dev Eq”). It turns out that the total weight is even higher than the effect of view 1 alone. This could indicate that there is a correlation effect between assets in the other views but it could also merely be a result of the optimization proces.

2. The second view should affect the first two assets relatively. The first asset view is negative with a P-factor of -1 and the second asset view is positive with a P-factor of +1. The figure shows that W_1 is significantly lower than W_2. The weight W_2 is also close to the sum of all views illustrated by W_{all}.

3. +90% of the P-factor in the third view deals with "US Lg Grth" and -90% deals with "US Lg Value". Both weights seem to be lower than W_{all} but the difference is not big.

Generally it seems that it is not easy to always directly see which impact a view has on the results in terms of e.g. expected returns and weights.

4.3 Partial conclusion on the use of the Black-Litterman model

The previous pages describes how the BL model has been implemented using:

- The model from He and Litterman (2002).
- The model from Idzorek (2005).

All results shown indicates that the programs work correct. Furthermore, time has been invested into trying to understand the connection between the views and the results (i.e. mostly the posterior returns $E(r)$ and the new weights, however implied confidence has also been investigated relatively detailed).

The mean and variances of the prior, conditional and posterior distributions has been plotted and shown because the results have been compared with the Akutan project in which we trust. It can be concluded that it is not always easy to predict the outcome in terms of $E(r)$ and the new weights, based only on the views. For this reason it has been decided to break up the model in even smaller pieces and to investigate very small parameter changes, when all (or most other) parameters are kept constant.

5 Investigation of the Black-Litterman model

In the following, the example data beginning from page 25 will be used together with parameter variation on P, Q, τ and ω_i.

5.1 A simple and very idealized BL model

The previously described output data was not always perfectly easy to interpret. For this reason a very simplified BL model has been made. There will be only 3 assets and the standard deviation of excess returns is $\sigma = 15\%$. Mean returns for the 3 assets are $\mu = \{-5\%;\ 0\%;\ 5\%\}^T$ p.a. respectively. First, 5000 years of fictive data was been generated in Excel using the command `=MEAN + NORMSINV(RAND())*STDDEV` where μ and σ was inserted. Then the 3×3 covariance matrix was studied. It is clear that all elements in the diagonal of the covariance matrix is the same, i.e. it is the variance $\sigma^2 = 0.15^2$. With 5000 years of artificial data, the offdiagonal elements are typically in the interval $\pm[1;5] \cdot 10^{-4}$ ($\sigma^2 = 0.0225$). We should remember that the covariance is (using N in the denominator for the whole population or $N-1$ for a sample):

$$\mathrm{cov}(X,Y) = \sum_{i=1}^{N} \frac{(x_i - \bar{x})(y_i - \bar{y})}{N} \tag{69}$$

From here it is clear that if there is no consistent development or change in variance, it is expected that the offdiagonal (co)-variance is 0. In other words, as there is no linear relationship between any of the assets the offdiagonal elements in the covariance matrix should not be in the order of 10^{-4}, it should approach 0 as $N \to \infty$. It is the same as saying that all assets move independently of each other which is exactly what we want to investigate, because the hypothesis is that this covariance matrix makes it much easier to interpret what happens in the BL model. Hence, we use a covariance matrix with 0 everywhere and $\sigma^2 = 0.0225$ in the diagonal. This also shows that now the BL-model cannot see this matrix started from excess returns with means -5%, 0% and +5% for assets 1-3, respectively.

We consider a market portfolio where $w_{eq} = \{1/3;\ 1/3;\ 1/3\}$ and the risk aversion coefficient $\delta = 3$. Using the variance $w_{eq}^T \Sigma w_{eq}$ the pro anno standard deviation of the market has been calculated to $\approx 2,9\%$ which is lower than compared with the standard deviation of each asset which was much more volatile ($\sigma = 15\%$ for each asset). The expected pro anno market returns is $\delta\sigma_{mkt}^2 = 0,25\%$ and the prior returns are calculated as $\Pi = \delta\Sigma w_{eq} = 0,25\%$ for each of the 3 assets. This corresponds with the theory of the CAPM, i.e. more volatility and higher risk should be rewarded. There is no α or β information left in the model[13], from the original problem where μ ranged from -5%

[13] From regression: $R_i - R_f = \alpha_i + \beta_i(R_M - R_f) + \varepsilon_i$, see also Table 4 on page 31 or `https://en.`

to +5%. This means, if the CAPM does not provide meaningful results, the BL-model probably will also not provide realistic results. Investors or fund managers should be aware of this.

5.1.1 Investigating variations on P and Q

To begin $\tau = 1\%$ implying that the uncertainty in the variance of the view (or views) are constant, i.e. $\omega = 2,5 \cdot 10^{-5}$. The problem can also be illustrated by the equation:

$$E(r) = \left[(\tau\Sigma)^{-1} + \underbrace{P}_{**}{}^{T}\Omega^{-1}\underbrace{P}_{**} \right]^{-1} \left[(\tau\Sigma)^{-1}\Pi + \underbrace{P}_{**}{}^{T}\Omega^{-1}\underbrace{Q}_{*} \right] \tag{70}$$

What happens when Q (*) is changed when using different P-vectors and when everything else is constant and the covariance matrix is only a diagonal matrix (so there is no cross-correlation between assets)? Figure 19 tries to explain this.

Figure 19a shows that while assets unaffected by the P-vector (asset number 2 and 3) does not change from equilibrium returns of 0,25%, the expected returns for asset number 1 changes linearly together with the view which is determined by the Q-scalar. Figure 19b is a relative view. When $Q = 0,01\%$ all assets have expected returns close to 0,25% (but not precisely). When $Q = 0,5\%$ $E(r)$ for asset number two decreased to only 0,125%. It makes sense because the view is negative on asset number 2, but positive on asset 1.

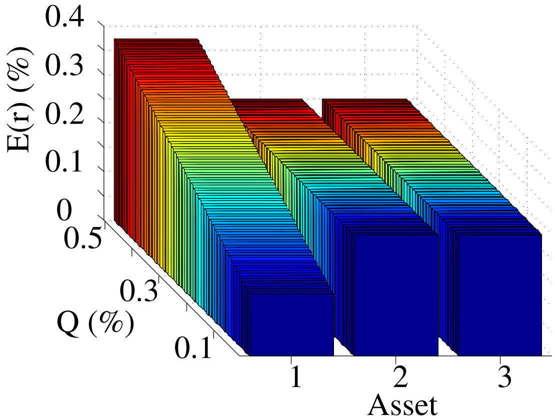
Figure 19c and 19d looks similar to the two previously described figures, however they are not the same. They show the weights calculated as $w = \lambda\Sigma_p^{-1}E(r)$. Had they been calculated as $w = \lambda\Sigma_p^{-1}\Pi$, the weight for each asset would not change as Q changes. At the same time, the weights would not change and continue to be approximately the same as when compared to the equilibrium situation where the market consisted of 1/3 of each asset.

One can almost guess what happens if an extra absolute view is added. An extra Q-value (the same value for both views) has been added with the following P-matrix:

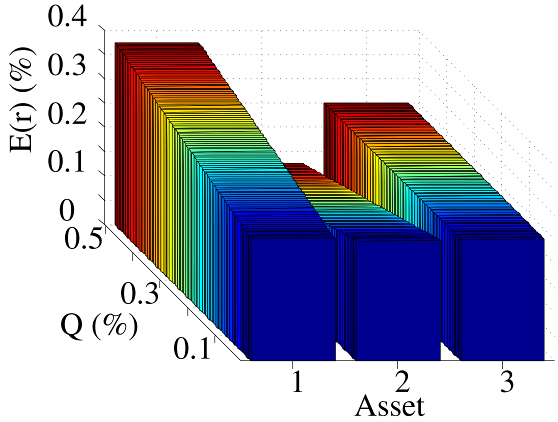
$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} \tag{71}$$

As seen in Figure 20, the only difference is whether the slope is positive or negative. The effect of -1 in view number 2 is the same as multiplying Q in view 2 by -1. The linear dependency is very clear. It means if Q becomes twice as big we also expect

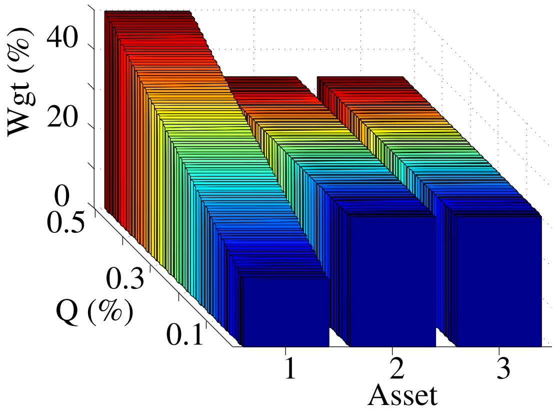
wikipedia.org/wiki/Alpha_(finance)



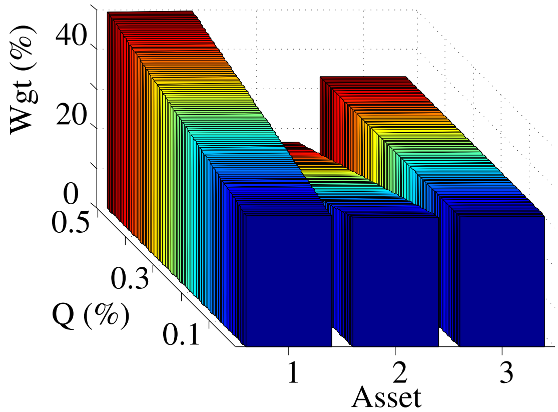
(a) $E(r)$: Absolute view: $P = \{1;\ 0;\ 0\}^T$.



(b) $E(r)$: Relative view: $P = \{1;\ -1;\ 0\}^T$.



(c) Wgt: Absolute view: $P = \{1;\ 0;\ 0\}^T$.



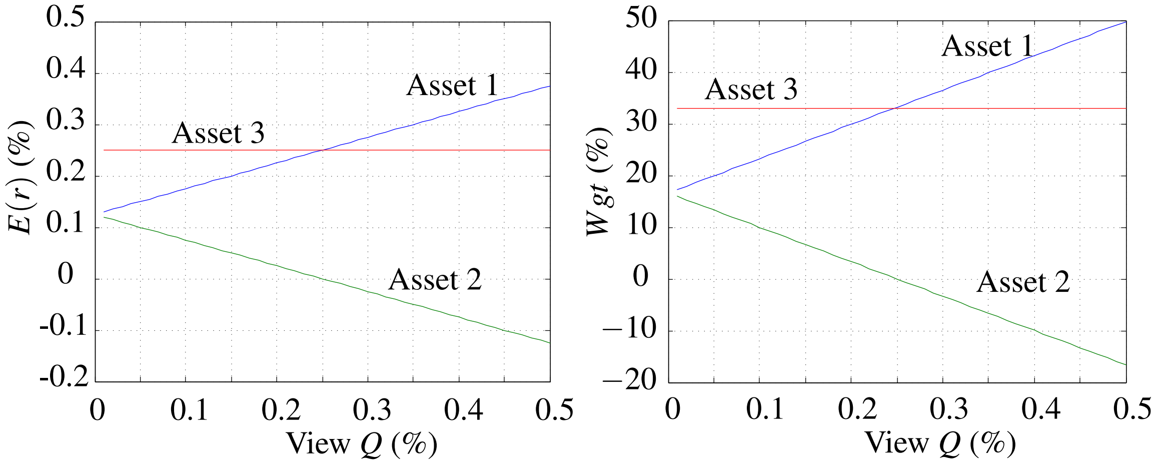
(d) Wgt: Relative view: $P = \{1;\ -1;\ 0\}^T$.

Figure 19: Difference between absolute and relative view in terms of $E(r)$ and weights calculated as $w = \lambda\Sigma_p^{-1}E(r)$. Minimum and maximum values of $E(r)$ are [0,125; 0,375]% while equilibrium is 0,25% as seen for asset 3. Weights are merely scaled where Π corresponds to 100% but when $E(r)$ changes, the sum can drastically differ from 100%.

the slope to become the double. This can directly be confirmed from (36) which for convenience is repeated below:

$$P \cdot E(r) = Q + \varepsilon = \begin{bmatrix} Q_1 \\ \vdots \\ Q_m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_m \end{bmatrix} \tag{72}$$

Additional effects of some P-matrices are shown in Figures 21 and 22 without the



(a) Posterior returns.

(b) Weights.

Figure 20: Two absolute views, with the same Q-values. The P-matrix is (71).

weights.

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix} \tag{73}$$

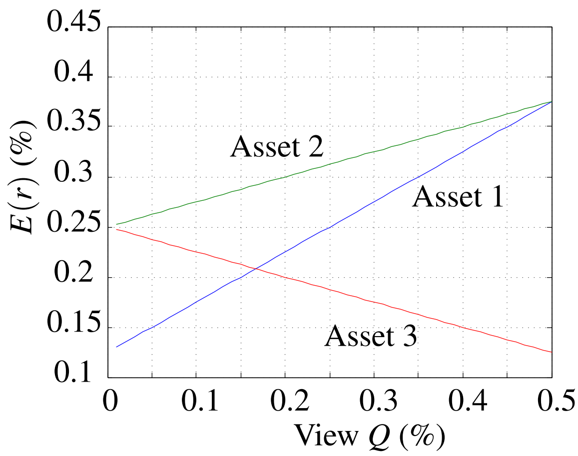
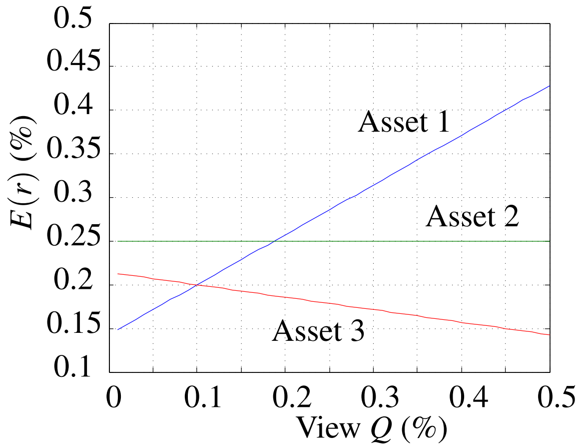


Figure 21: One absolute and one relative view, with the same Q-values.

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix} \tag{74}$$



$$P = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} \tag{75}$$

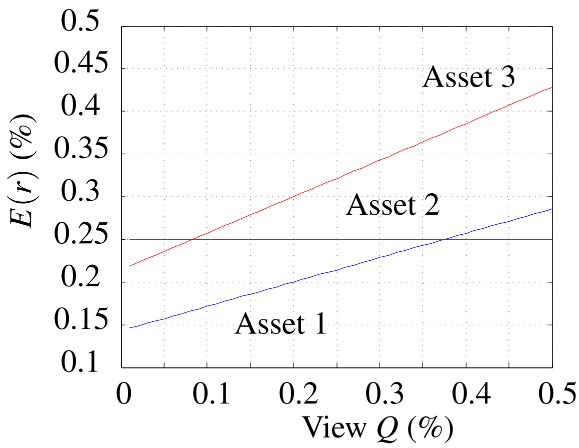
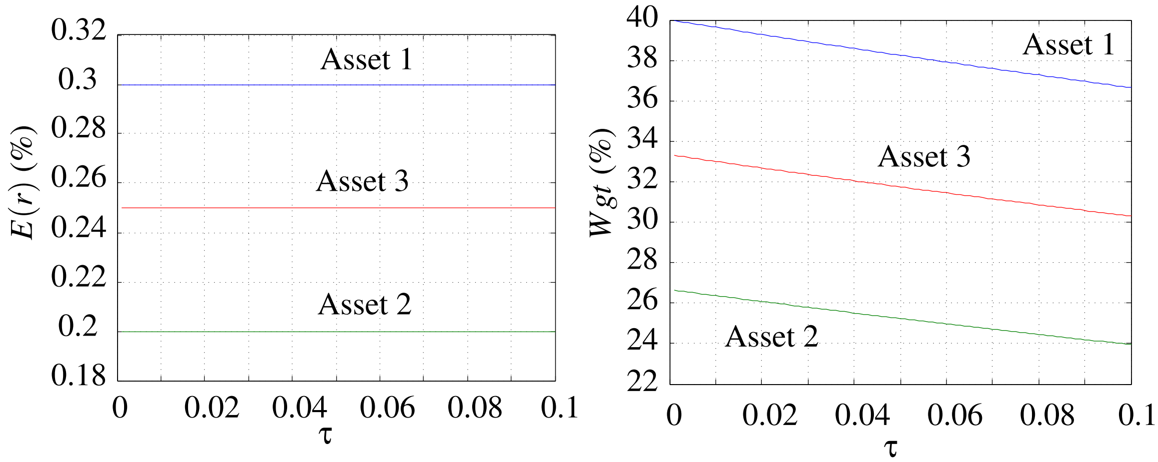


Figure 22: One absolute and one relative view, with the same Q-values.

5.1.2 About variance on the views, using He and Litterman (2002)

The parameter τ has an influence on the variance of the view described by Ω. The effect comes through the diagonal elements given by (35) from page (14). However that is not the only place it appears (*):

$$E(r) = \left[(\underbrace{\tau}_{*} \Sigma)^{-1} + P^T \underbrace{\Omega}_{*}{}^{-1} P \right]^{-1} \left[(\underbrace{\tau}_{*} \Sigma)^{-1} \Pi + P^T \underbrace{\Omega}_{*}{}^{-1} Q \right] \tag{76}$$



(a) Posterior returns.

(b) Weights.

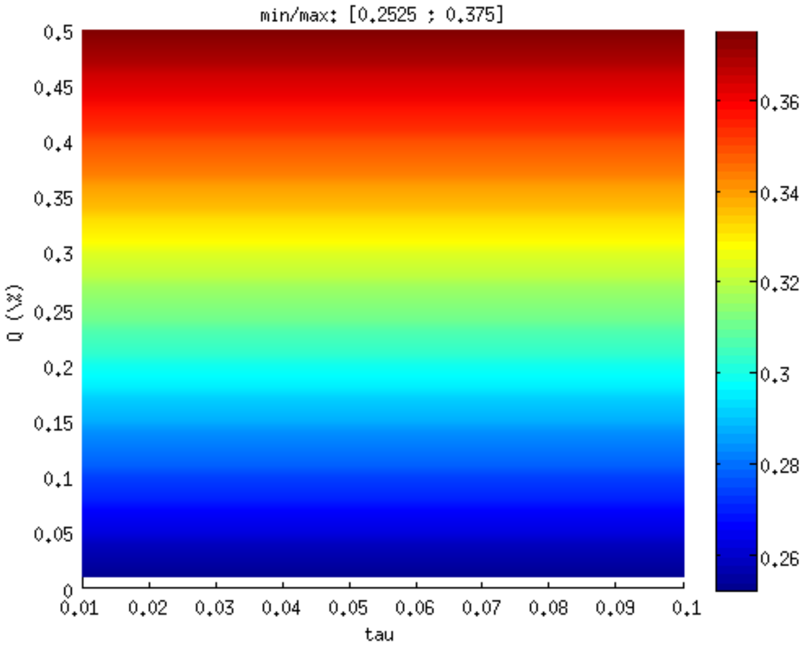
Figure 23: Influence of τ on $E(r)$ and weights, when $P = \{1;\ -1;\ 0\}^T$ and $Q = 0,20\%$.

The parameter τ is included in all 4 terms, in the calculation for the posterior returns. In Figure 23 we consider the result of $P = \{1;\ -1;\ 0\}^T$ and $Q = 0,20\%$. In other words, $E(r)$ for asset 1 is 1/5 higher $(0,25 \cdot 1,2 = 0,3)$ and $E(r)$ for asset 2 is 20% lower $(0,25 \cdot 0,8 = 0,2)$. It can be seen that $E(r)$ is stable and apparently independent of τ. However, as the weights are $w = \lambda \Sigma_p^{-1} E(r)$ and since λ is a constant, Σ_p as given by 34 must increase, when $\tau \to \infty$:

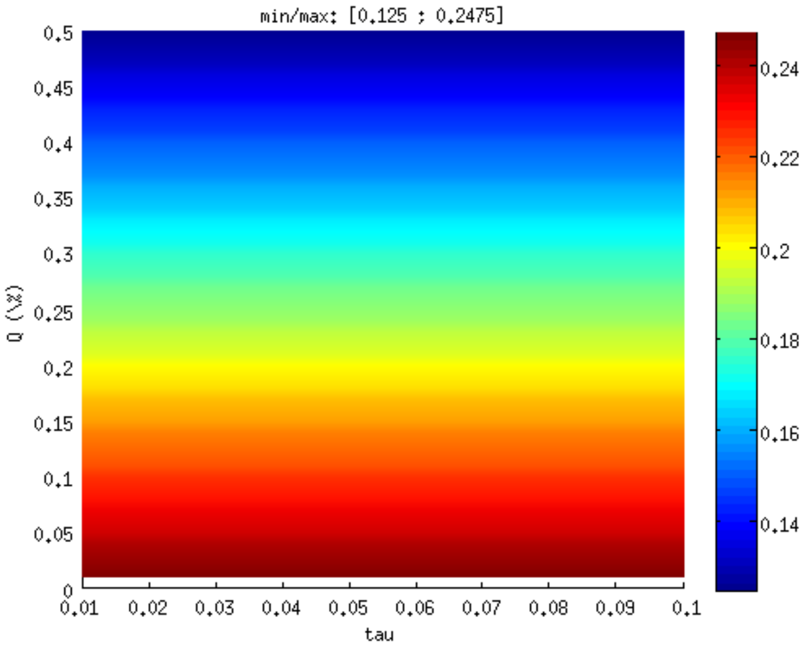
$$\Sigma_p = \Sigma + M = \Sigma + \left((\tau\Sigma)^{-1} \nearrow^{0} + P^T \underbrace{\Omega^{-1} \nearrow^{0}}_{\Omega = \mathrm{diag}(P(\tau\Sigma)P^T)} P \right)^{-1} \tag{77}$$

The last paranthesis should be inverted so the value of the last term becomes larger, when what is inside is small. The diagonal element of Ω is linearly proportional to τ. The posterior covariance matrix is always larger than the historical covariance matrix. An interpretation could sound: A larger covariance matrix means the variances and covariances are larger so the uncertainty did increase as τ increased. Larger values of τ increases the variance of the views and later, the posterior covariance matrix. Higher variance and covariance is the same as higher risk or volatility which again provides higher returns according to the CAPM relationship where higher risk provides higher returns. The view Q is fixed and we already calculated $E(r)$ for all 3 assets. When returns are fixed, the weights of risky assets in the portfolio must decrease, otherwise $E(r)$ will increase which they did not according to Figure 23a. Investors can buy the risk-free asset at a given rate r_f, while the uncertainty in the views determined by τ increases. That is an interpretation for why the weights are decreasing and why $E(r)$ is constant (it is bound by the view).

Figure 24 is a surface plot showing some of the same mechanisms as in the previous figures. The difference is that now τ and Q is changed at the same time. The weights are however not shown. The tendency as in the previous plots is the same, i.e. as τ increase, the weights decrease.



(a) $E(r)$ for asset 1.

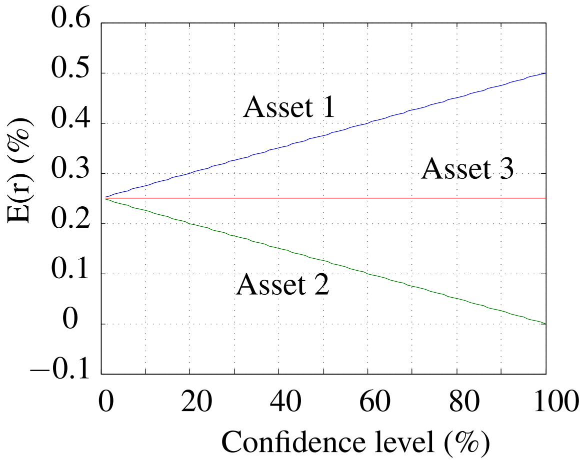


(b) $E(r)$ for asset 2 (opposite direction).

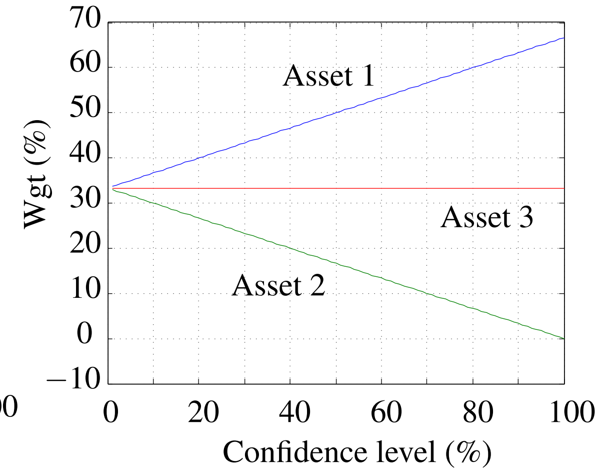
Figure 24: Influence of τ and Q on $E(r)$ for asset 1 and asset 2, when $P = \{1;\ -1;\ 0\}^T$. The figure for asset 3 is not shown. If it was shown, it would be completely green because $E(r) = 0,25\%$ and it is independent of τ and Q with this P-vector.

5.1.3 About variance on the views, using Idzorek (2005)

Instead of modifying Ω by τ, an alternative method has been described from page 19. The confidence is between 0% and 100%. We will consider a single view where $Q = 0,5\%$ and $P = \{1;\ -1;\ 0\}^T$ to check if everything behaves as expected. The results in terms of $E(r)$ and weights are shown in Figure 25.



(a) Posterior returns.



(b) Weights.

Figure 25: Influence of user-specified confidence levels as suggested by Idzorek (2005) on $E(r)$ and weights, when $P = \{1;\ -1;\ 0\}^T$ and $Q = 0,5\%$.

Figure 25 show that there is absolutely nothing surprising in the results. However it is interesting to see that 0% confidence levels does not change neither the expected returns, nor the weights from the equilibrium position. At the same time it is interesting to visually see, that 100% confidence raises $E(r)$ +0,5% for asset 1 and it decreases $E(r)$ 0,5% for asset 2. Everything in between 0 and 100% confidence is just a linear interpolation and the principle of superposition[14] is therefore expected to be in effect (there are no non-linear effects to be seen in any of the equations or results).

If we instead investigated the absolute view given by $P = \{1;\ 0;\ 0\}^T$, the linear relationship in terms of $E(r)$ for asset 1 would be unchanged, when compared to Figure 25a. But asset 2 and asset 3 would be unaffected by the change in confidence level, as the view does not involve does assets. The exact same holds also for Figure 25b.

5.2 Superposition effects from off-diagonal covariance matrix elements

The following requires an introduction to the theory of covariance matrices, before an example with offdiagonal covariance matrix elements will be given and the effect (or

[14] https://en.wikipedia.org/wiki/Superposition_principle

importance) will be illustrated.

5.2.1 Introduction

The problems shown on the previous pages are simple, because there are no off-diagonal elements in the variance-covariance matrix Σ. That will change now but first some background information about covariance matrices will be given. Figure 26 shows a 2×2 covariance matrix, because it is easier to illustrate key concepts in 2, rather than in 3 dimensions. Unlike from the situation described on the previous pages, the covariance matrix is not longer a diagonal matrix. 1000 random points, closely around the mean $\mu = (0,0)$ makes up a cloud of artificial returns. The figure can be interpreted as e.g. the annual returns for stock y given the annual returns of stock x. For the shown matrix Σ it can be seen that it is not easy or even not possible to construct a good fit for a regression line, that describes the variance in either x, y or the cross- (=co)-variance x, y for the artificial cloud ("of annual returns" or something else).

$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \approx \begin{bmatrix} 0.9759 & 0.0089 \\ 0.0089 & 1.0704 \end{bmatrix} \quad (78)$$

$$V_1 = \begin{bmatrix} -0.9956 \\ 0.0933 \end{bmatrix}, \quad V_2 = \begin{bmatrix} 0.0933 \\ 0.9956 \end{bmatrix} \quad (79)$$

$$\lambda = \begin{bmatrix} 0.9751 \\ 1.0713 \end{bmatrix} \quad (80)$$

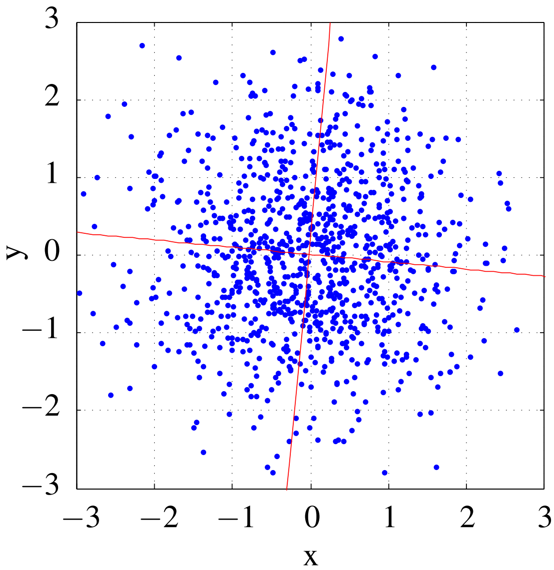


Figure 26: Cloud of e.g. annual returns variables (or assets) x and y, when the covariance matrix Σ is almost an identity matrix. The red lines illustrate the direction of the eigenvalues of the estimated covariance matrix.

Without going too much into details, mathematically it is possible to calculate the eigenvectors and eigenvalues[15] of the estimated covariance matrix Σ from (78). The result is shown in (79) and (80) respectively. Eigenvectors **v** and eigenvalues ˘ gives the solution to an eigenvalue problem of the form:

[15] https://en.wikipedia.org/wiki/Eigenvalues_and_eigenvectors

$$\mathbf{A}\mathbf{v} = \boldsymbol{\lambda}\mathbf{v} \tag{81}$$

where **A** is a real square matrix. The matrix **A** can also be seen as a linear transformation matrix (also called a linear mapping) between vector spaces, like Σ can be considered a linear transformation when multiplied by a vector. Equation (81) tells that the numbers or eigenvalues λ can be multiplied with a vector **v** and this gives the exact same result as when the transformation matrix **A** is multiplied by **v**. In other words, what we have shown is that:

$$\boldsymbol{\Sigma}\mathbf{v} = \boldsymbol{\lambda}\mathbf{v} \tag{82}$$

One might now ask the question "and so what?". At a first glance one might think that Figure 26 do not add any relevant new information and the talk about eigenvectors and eigenvalues is too complex. The red perpendicular lines in Figure 26 illustrate the directions of the eigenvectors of the covariance matrix while λ is a scaling factor that tells us which of the eigenvectors describes the most important direction of "cloud variance". It is easier to illustrate the effect of different covariances, when constructing a different cloud, such as e.g. the cloud shown in Figure 27.

$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 0.1 \end{bmatrix} \approx \begin{bmatrix} 1.0683 & 0.0194 \\ 0.0194 & 0.0967 \end{bmatrix} \tag{83}$$

$$V_1 = \begin{bmatrix} 0.0200 \\ -0.9998 \end{bmatrix}, \quad V_2 = \begin{bmatrix} -0.9998 \\ -0.0200 \end{bmatrix} \tag{84}$$

$$\lambda = \begin{bmatrix} 0.0963 \\ 1.0687 \end{bmatrix} \tag{85}$$

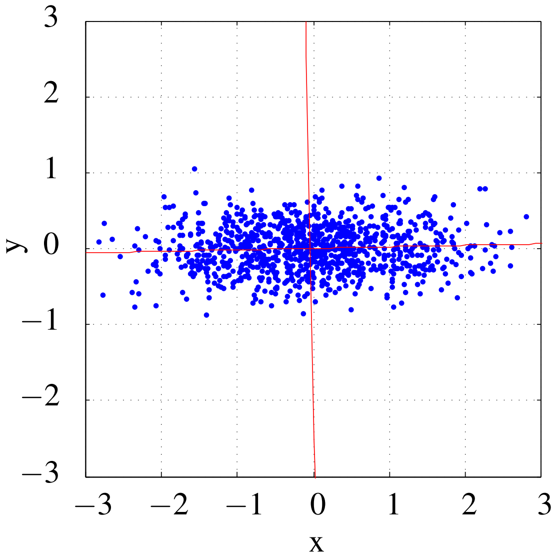


Figure 27: Another cloud of e.g. annual returns variables (or assets), but this time the variance in the y-direction is approximately 10 times smaller than the variance in the x-direction. This can be seen directly from the covariance matrix.

Figure 27 shows a new cloud, based on a new covariance matrix where $\sigma_{xx} \approx 1$ and $\sigma_{yy} \approx 0,1$. If someone asked us to make a linear regression line through the cloud, it would be much easier and the linear regression fit would be much better than it would for the cloud in Figure 26. The eigenvalues $\lambda_1 \approx 0,1$ and $\lambda_2 \approx 1$ tells that the direction which best describes the variance in the cloud data, is given by the maximum eigenvalue and its corresponding eigenvector v_2. The eigenvector v_2 is the red line which is almost horizontal and v_1 is by definition orthogonal to v_2. In Appendix B a few extra clouds based on new and different covariance matrices, are shown.

Not all matrices have the property that the eigenvalue problem (81) can be solved and also, sometimes the eigenvalues are complex. For practical purposes, we are only interested in real eigenvalues. The amount of litterature on linear algebra and the solution to these problems is huge. There are many rules which will not be described here. One of the rules is that if the matrix **A** is hermitian, i.e. in a way "symmetrical" then entries should be equal to their own conjugate transpose on the other side of the diagonal. The diagonal elements must be real so the following is an example of a hermitian matrix:

$$\begin{bmatrix} 2 & 2+\mathrm{i} & 4 \\ 2-\mathrm{i} & 3 & \mathrm{i} \\ 4 & \text{-i} & 1 \end{bmatrix} \tag{86}$$

where element (i,j) is equal to the complex conjugate of element (j,i). Covariance matrices are always not only hermetian and symmetric, but also positive semi-definite meaning that every eigenvalue is non-negative. The following is a summary of some general properties of covariance matrices, which can be found many places:

$$\boldsymbol{\Sigma}_x = E\{(\mathbf{x} - \bar{x})(\mathbf{x} - \bar{x})^T\} \tag{87}$$

where $\mathbf{x}$ is a random vector of real numbers with dimension or length n (which is written $\mathbf{X} \in \mathbf{R^n}$) and $\bar{x}$ is a mean vector. Alternatively, the elements $\sigma_{i,j}$ which is the (co)-variance between asset i and j, can be calculated using (13) or we could write

$$\sigma_{i,j} = E\{(\mathbf{x}_i - \bar{x}_i)(\mathbf{x_j} - \bar{x}_j)^T\} \tag{88}$$

In case of diagonal entries, the variance of $\mathbf{x}_i$ is:

$$\sigma_{i,i} = E\{(\mathbf{x}_i - \bar{x}_i)^2\} = \sigma_i^2 \tag{89}$$

As we only work with real numbers, the diagonal elements must always be positive. Also, the covariance matrix is symmetric, i.e. $\Sigma = \Sigma^T$ because $\sigma_{i,j} = \sigma j, i$. Furthermore, the covariance matrix is positive semi-definite which can be seen by the following operations using the vector $u \in \mathbf{R}^n$:

$$
\begin{aligned}
E\{[(\mathbf{x}-\bar{x})^T\mathbf{u}]^2\} = E\{[(\mathbf{x}-\bar{x})^T\mathbf{u}]^T[(\mathbf{x}-\bar{x})^T\mathbf{u}]\} &\geq 0 \\
E[\mathbf{u}^T(\mathbf{x}-\bar{x})(\mathbf{x}-\bar{x})^T\mathbf{u}] &\geq 0 \\
\mathbf{u}^T\mathbf{\Sigma}_x\mathbf{u} &\geq 0
\end{aligned}
\tag{90}
$$

A lot of mathematical litterature can be found on these topics. The intention was only to provide a quick overview and understanding, before changing the covariance matrix to something slightly more advanced than previously used, which only caused simple linear effects.

5.2.2 Tiny modification of the covariance matrix

With off-diagonal elements in the covariance matrix, the results from one view affecting one asset should affect other assets because a correlation effect or cross-variance effect will happen. To investigate this effect, the covariance matrix will be changed such that the covariance between assets 1 and 2 will be 50% of each of the variances:

$$
\underbrace{\Sigma = 10^{-3} \cdot \begin{bmatrix} 2.5 & 0 & 0 \\ 0 & 2.5 & 0 \\ 0 & 0 & 2.5 \end{bmatrix}}_{\text{Experiment A}} \quad \rightarrow \quad \underbrace{10^{-3} \cdot \begin{bmatrix} 2.5 & 1.25 & 0 \\ 1.25 & 2.5 & 0 \\ 0 & 0 & 2.5 \end{bmatrix}}_{\text{Experiment B}} \tag{91}
$$

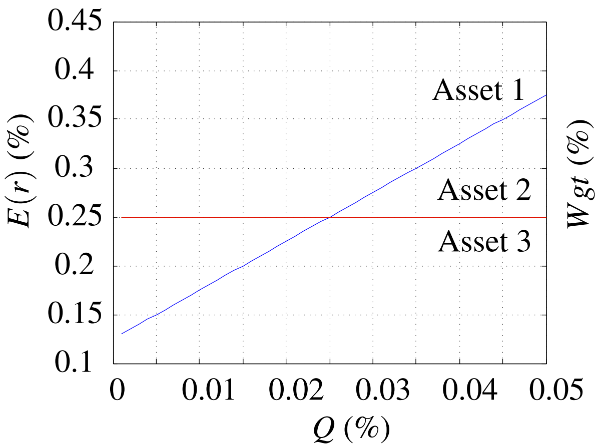
Corresponding correlation matrices can be calculated from:

$$
\mathrm{Cor}(x,y) = \frac{\mathrm{Cov}(x,y)}{\sigma_x \sigma_y} \Rightarrow \tag{92}
$$

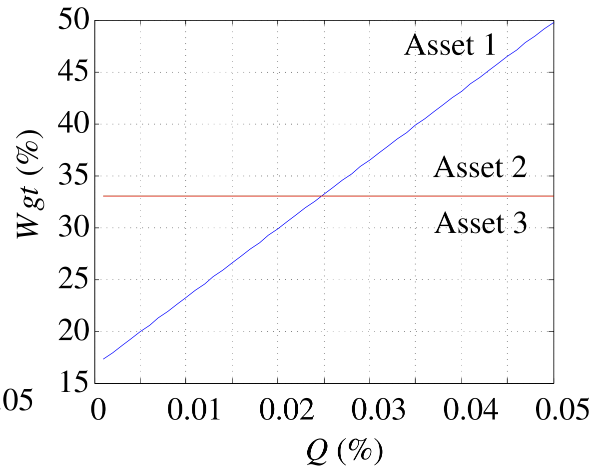
$$
\Sigma = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Experiment A}} \quad \rightarrow \quad \underbrace{\begin{bmatrix} 1 & 0.5 & 0 \\ 0.5 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Experiment B}} \tag{93}
$$

To begin, there will be 2 “experiments”, i.e. experiment A and experiment B which has the covariance matrices shown above. The first covariance matrix (A) is the one used in the examples starting from page 49 and it has been used until now. The new covariance matrix (B) is slightly modified. The modification has the effect, that there will be a relationship between assets 1 and 2 and we want to see this effect demonstrated, when everything else is kept constant in the BL-framework.

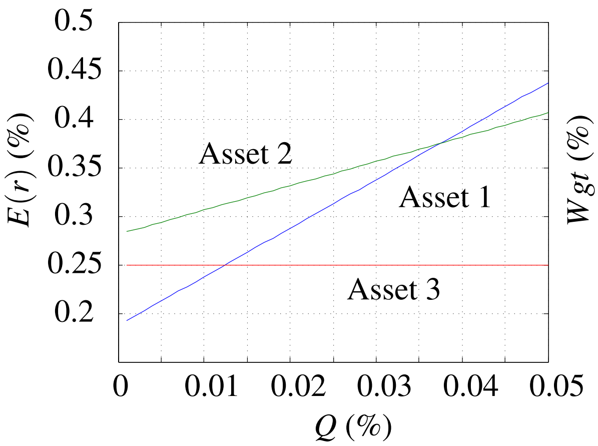
Figure 28 shows the results. Figure 28a and 28b did not change from earlier. Figures 28c and 28d are new. Even though there is no view on asset 2, now the posterior returns



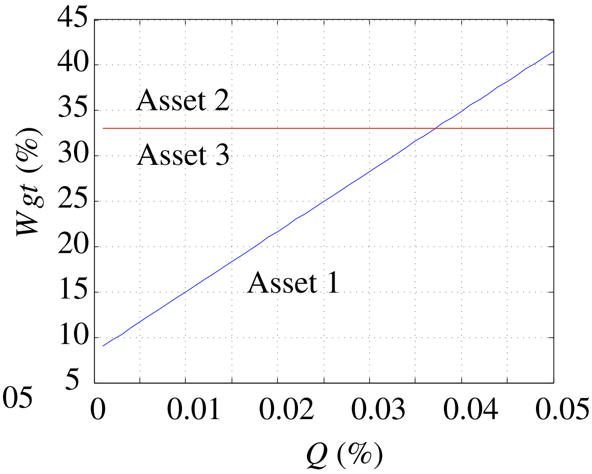
(a) $E(r)$: Experiment A (as Figure 19a).



(b) Wgt: Experiment A (as Figure 19c).



(c) $E(r)$: Experiment B.

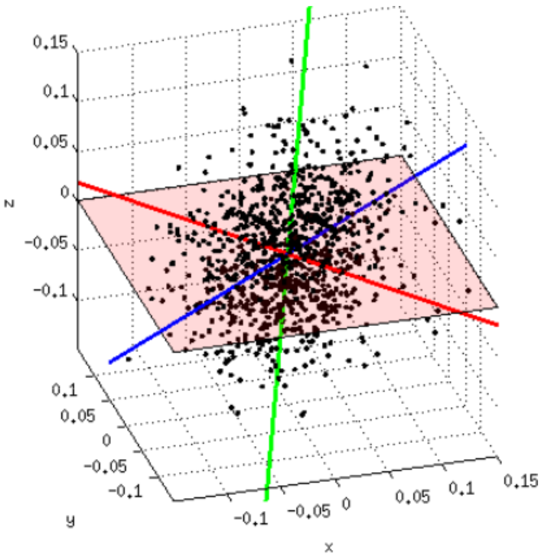


(d) Wgt: Experiment B.

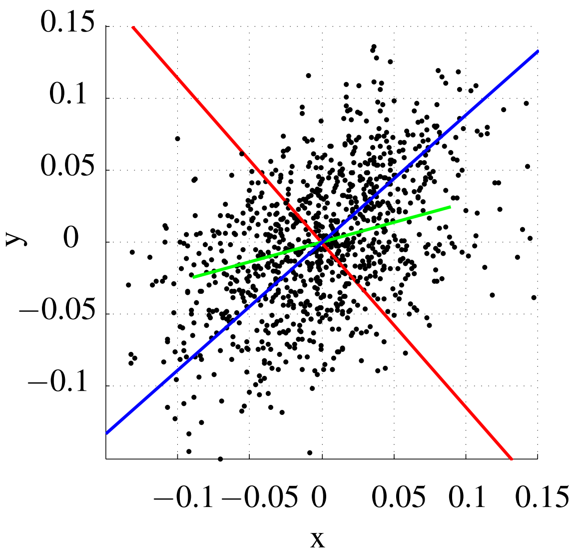
Figure 28: The effect of adding a covariance element outside the diagonal of Σ, which is 50% of the variance for the other assets as illustrated by (91). The same absolute view as previously have been used, i.e. $P = \{1;\ 0;\ 0\}^T$ (and $\tau = 1\%$ as earlier).

increase even though the view itself only affects view 1 and it was an absolute view. The explanation must be that now asset 1 and asset 2 are correlated positively. While the slope for asset 1 and 2 is not the same asset 3 is still completely unaffected. Asset 1 has the strongest dependency on the view Q. The slope for asset 1 in Figure 28c is 5 %/% while for asset 2 it is 2,5 %/%, hence not only is the relationship with $E(r)$ linear, the posterior returns seem to scale with the coefficients in Σ which also makes sense for linear problems. Also, the effect of superposition seems to be in place, because the with the unmodified covariance matrix, for asset 1 $E(r) = 0,375$ % but this number increases to $E(r) = 0,4375$ for the experiment B with modified covariance matrix.

At first, Figures 28b and 28d look similar. A closer look however reveals that the weights have changed. The un-modified covariance matrix ends up with 50% weight for asset 1. This is lower in experiment B, only around 41,5%. As the weights are calculated as $w = \lambda\Sigma_p^{-1}E(r)$ it makes sense that the weights should change but at the moment it does not seem easy to come up with an intuitive explanation for what we see. But we can see that the effect of increased returns from asset 2 (compared to asset 3) and the inverted covariance matrix seems to cancel each other out as asset 2 and asset 3 is independent of the view Q.



(a) 3D plot of cloud using new Σ.



(b) 2D plot of the same cloud.

Figure 29: New cloud including directions of eigenvectors in both 2 and 3 dimensions. The cloud and covariance parameters are shown in Equations (94)- (96). The red transparent surface illustrates the xy-plane at $z = 0$. The direction of the eigenvectors is now indicated by red, green and blue lines respectively.

If we use the same methodology used previously to try to understand the new covariance matrix given by (91) we can again take the new covariance matrix and create a cloud with 1000 points. This is shown in Figure 29 and the cloud has the following parameters (now in 3 dimensions):

$$\Sigma = \begin{bmatrix} 0.0025 & 0.0013 & 0 \\ 0.0013 & 0.0025 & 0 \\ 0 & 0 & 0.0025 \end{bmatrix} \approx \begin{bmatrix} 0.0026 & 0.0012 & -0.0001 \\ 0.0012 & 0.0023 & -0.0001 \\ -0.0001 & -0.0001 & 0.0024 \end{bmatrix} \tag{94}$$

$$V_1 = \begin{bmatrix} 0.6583 \\ -0.7516 \\ -0.0411 \end{bmatrix}, \quad V_2 = \begin{bmatrix} -0.0898 \\ -0.0242 \\ -0.9957 \end{bmatrix}, \quad V_3 = \begin{bmatrix} 0.7474 \\ 0.6592 \\ -0.0834 \end{bmatrix} \tag{95}$$

$$\lambda = \begin{bmatrix} 0.0013 \\ 0.0024 \\ 0.0037 \end{bmatrix} \tag{96}$$

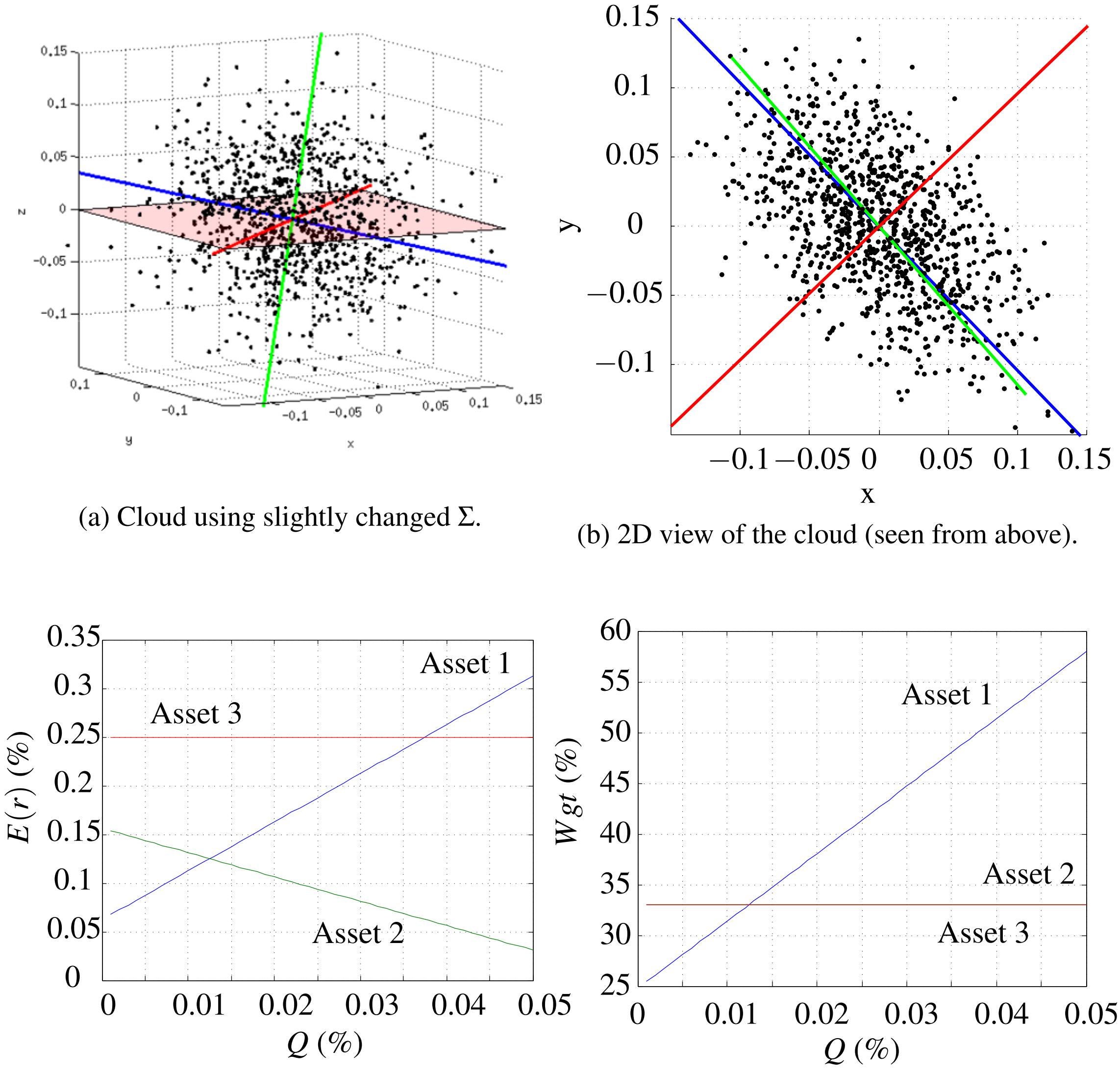
The direction with the highest risk is determined by the third eigenvector, which has been given a blue color. While that is difficult to see on Figure 29a it is much easier to verify and see in Figure 29b which is the same, but shown from the top. The second eigenvector is the second-most important director (green) while the first eigenvector (red) is the least important. In case of dimensionality reduction, we would throw away the direction the variance given by the red direction because it is the direction with the least variance.

It is reasonable to believe that if it was not $\sigma_{1,2}$ but instead $\sigma_{1,3}$ that was not zero anymore, then everything would be similar to here except that then we should just swap the y and z-directions of everything in all the results. If we substituted $\sigma_{1,2}$ by the same multiplied by minus 1, we get experiment C:

$$\underbrace{\Sigma = 10^{-3} \cdot \begin{bmatrix} 2.5 & -1.25 & 0 \\ -1.25 & 2.5 & 0 \\ 0 & 0 & 2.5 \end{bmatrix}}_{\text{Experiment C}} \quad \text{or:} \quad \text{Cor} = \underbrace{\begin{bmatrix} 1 & -0.5 & 0 \\ -0.5 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Experiment C}} \tag{97}$$

The results for experiment C is shown in Figure 30. By comparing Figure 30b and 29b we notice that direction of the most important eigenvector changed from around 45° to around -45°. Again the last eigenvalue (the blue eigenvector) is the largest. If we made a regression line between asset x and y we would also say the slope is now around -45° which is in contrast to experiment B.

Comparing Figure 30c with Figure 28c reveals that asset 3 is once again completely unaffected the the view. The slope of asset 1 is again +5 %/% while it is -2,5 %/% for asset 2. Again we see that the solution to the problem is simple, because it is linear.



(a) Cloud using slightly changed Σ.

(b) 2D view of the cloud (seen from above).

(c) $E(r)$: Experiment C.

(d) Wgt: Experiment C.

Figure 30: Covariance matrix given by (97), with negative correlation between assets 1 and 2. It is clear that $E(r)$ moves in opposite direction with increasing Q.

There are many ways to continue to make changes to the covariance matrix and explore the effects, after having processed the input data with the Black-Litterman model. However as everything is linear and the superposition principle applies, it is deemed to be more interesting to take a real example, using real input data instead of artificial input data.

5.3 Using real input data on more complex covariance matrices

We start with the covariance matrix from Markowitz portfolio optimization, i.e. see Figure 13a on page 33. There are several ways to find the market weights but in the end they have been chosen to be as in Table 12:

	EUSA	EWG	EWH	EWQ	EWU	MCHI
w_{eq}	-6.90	35.6	10.2	33.4	13.1	14.7
Π	4.56	9.90	6.74	10.2	7.65	9.39

Table 12: Market weights and prior or equilibrium returns (everything in percent).

The risk-aversion coefficient was chosen to be $\delta = 3$, but the data processed was based on daily returns and therefore the covariance matrix is also only based on daily returns. Using the variance $w_{eq}^T \Sigma w_{eq}$ the daily standard deviation of the market has been calculated to $11{,}5 \cdot 10^{-3}$. Multiplying this with the square root of 245 which was used as the number of trading days per year yields 17,95% which is also the pro anno standard deviation of the market. For calculating the prior returns, $\Pi = \delta \Sigma w_{eq}$ and using the rest of the BL model framework, either the covariance matrix had be modified from daily to yearly variances and covariances. Or something else had to be done. An easy way to make everything work out correctly, was to scale the risk-aversion coefficient δ by 245 times. The prior returns can now be calculated and are also shown in Table 12.

Having validated and done many experiments in the spirit of the BL framework now, we want to investigate something we believe is interesting to look at at something that we have not seen earlier. Before setting up our views, we've learned that it is not completely irrelevant how the covariance matrix looks or how the correlation between each individual assets are. Therefore, this is our starting point (we could also set up some arbitrary views, make up some strange-looking P-vector or P-matrix and setup some views in the Q-vector and plot the results but so what?). The covariance (the same as in Figure 13a) matrix for our "real input data"-based problem is:

$$\Sigma = 10^{-3} \cdot \begin{bmatrix} 0.0677 & 0.0635 & 0.0488 & 0.0646 & 0.0587 & 0.0674 \\ 0.0635 & 0.1514 & 0.0764 & 0.1445 & 0.1032 & 0.1068 \\ 0.0488 & 0.0764 & 0.1217 & 0.0782 & 0.0707 & 0.1368 \\ 0.0646 & 0.1445 & 0.0782 & 0.1587 & 0.1082 & 0.1103 \\ 0.0587 & 0.1032 & 0.0707 & 0.1082 & 0.1047 & 0.0976 \\ 0.0674 & 0.1068 & 0.1368 & 0.1103 & 0.0976 & 0.2098 \end{bmatrix} \tag{98}$$

where all (co)-variances above $0{,}15 \cdot 10^{-3}$ are red, everything above $0{,}13 \cdot 10^{-3}$ is yellow and the numbers above $0{,}11 \cdot 10^{-3}$ are green. The corresponding correlation matrix is

$$\text{Cor} = \begin{bmatrix} 1.0000 & 0.6272 & 0.5378 & 0.6231 & 0.6971 & 0.5656 \\ 0.6272 & 1.0000 & 0.5632 & 0.9325 & 0.8199 & 0.5992 \\ 0.5378 & 0.5632 & 1.0000 & 0.5631 & 0.6265 & 0.8560 \\ 0.6231 & 0.9325 & 0.5631 & 1.0000 & 0.8392 & 0.6042 \\ 0.6971 & 0.8199 & 0.6265 & 0.8392 & 1.0000 & 0.6583 \\ 0.5656 & 0.5992 & 0.8560 & 0.6042 & 0.6583 & 1.0000 \end{bmatrix} \tag{99}$$

where the diagonal has been made red, correlation coefficients above 0,9 are yellow and correlation coefficients above 0,8 are green. The colors indicates which assets correlates the most with each other. As an example, asset 2 (EWG) correlates a lot with asset 4 (EWQ) and slightly lesser with asset 5 (EWU). Asset 1 seems to be the most defensive of all assets. There are no negative correlation coefficients, which is bad seen from a diversification point of view. In other words, when the market rises everything rises and vice versa.

As we can not make 6-dimensional figures in this report, we have to think of the 6-asset space as something that is similar to that shown in 2 and 3 d dimensions. The eigenvectors are:

$$V_1 = \begin{bmatrix} -0.0666 \\ 0.6688 \\ -0.0355 \\ -0.7241 \\ 0.1501 \\ 0.0156 \end{bmatrix}, \quad V_2 = \begin{bmatrix} 0.3930 \\ 0.3290 \\ 0.1291 \\ 0.0857 \\ -0.8439 \\ -0.0335 \end{bmatrix}, \quad V_3 = \begin{bmatrix} 0.0768 \\ 0.0014 \\ -0.7949 \\ 0.0239 \\ -0.1063 \\ 0.5919 \end{bmatrix} \tag{100}$$

$$V_4 = \begin{bmatrix} -0.8782 \\ 0.2511 \\ 0.0612 \\ 0.2543 \\ -0.2813 \\ 0.1347 \end{bmatrix}, \quad V_5 = \begin{bmatrix} -0.0616 \\ -0.4275 \\ 0.4600 \\ -0.4387 \\ -0.1937 \\ 0.6098 \end{bmatrix}, \quad V_6 = \begin{bmatrix} 0.2452 \\ 0.4458 \\ 0.3672 \\ 0.4590 \\ 0.3706 \\ 0.5082 \end{bmatrix} \tag{101}$$

The way eigenvectors and eigenvalues are presented depends normally on code implementation. In this case it seems like they're always sorted from the least significant to the most significant eigenvalue:

$$\lambda = 10^{-3} \cdot \begin{bmatrix} 0.0102 \\ 0.0192 \\ 0.0221 \\ 0.0359 \\ 0.1210 \\ 0.6056 \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} 1.25\% \\ 2.36\% \\ 2.71\% \\ 4.41\% \\ 14.87\% \\ 74.40\% \end{bmatrix} \tag{102}$$

Eigenvector number 6 seems to very important as it is the largest. By taking the sum of all eigenvalues, we see that λ_6 accounts for 74,4% of the magnification from the linear transformation given by Σ. The fifth and sixth eigenvector spans almost 90% of the total variance.

We shall setup our views, based on the two extreme eigenvectors V_1 and V_6. We will setup absolute views that resembles the portfolios expressed by V_1 and V_6 by scaling the eigenvectors such that the sum of all elements is 100%, i.e.

$$V_{1,\%} = P_1 = \begin{bmatrix} -809\% \\ 8116\% \\ -431\% \\ -8787\% \\ 1821\% \\ 189\% \end{bmatrix}, \qquad V_{6,\%} = P_2 = \begin{bmatrix} 10.2\% \\ 18.6\% \\ 15.3\% \\ 19.2\% \\ 15.5\% \\ 21.2\% \end{bmatrix} \tag{103}$$

The $P_1 = V_{1,\%}$ vector seems to have some very large fluctuations, maybe caused by machine precision errors that are magnified or maybe not. It is always easier to work on the largest eigenvectors and eigenvalues and the relative scaling between 10% and 22% for $P_2 = V_{6,\%}$ looks much better. Equation (103) should be understood such that if we take e.g. $P_2 = V_{6,\%}$ then our view is +10,2% on EUSA, +18,6% on EWG and so on. In any case we clearly see that with negative views on individual assets such as with P_1 short-sales is allowed, thus we're looking at an unconstrained optimization problem. We know the maximum element in the vector of prior returns is 10,2% so let us define $Q = 20\%$ with an absolute view and test both P-vectors, one by one.

5.3.1 *P* with the weakest amplification of the linear transformation

Table 13 confirms the hypothesis about the importance of the eigenvalues and eigenvectors. Even though our view tells the BL-model that we "expect" an absolute view of $Q = 20\%$ exactly this P-vector combination gives a change in expected returns that is less than 1%.

5.3.2 *P* with the strongest amplification of the linear transformation

Table 14 confirms the hypothesis about the importance of the eigenvalues and eigenvectors. Even though our view told the BL-model that we "expect" an absolute view of $Q = 20\%$ exactly this P-vector combination gives a change in expected returns that ranges from an increase from around 60-75% above the equilibrium levels for all 6 assets. Absolute posterior returns are however never 20%, they range from around 8-17% but it is still much better than the equilibrium returns.

	EUSA	EWG	EWH	EWQ	EWU	MCHI
Π (%)	4.56	9.90	6.74	10.16	7.65	9.39
$E(r)$ (%) =	4.55	9.95	6.74	10.11	7.66	9.39
$\Delta_r = \frac{E(r)-\Pi}{\Pi}$ (%)	-0.11	0.51	-0.04	-0.54	0.15	0.01
w_{eq} (%)	-6.90	35.57	10.17	33.37	13.05	14.73
Wgt (%)	-7.50	41.92	9.71	25.79	14.43	14.74
$\Delta_w = \frac{Wgt-w_{eq}}{w_{eq}}$ (%)	8.69	17.84	-4.49	-22.72	10.53	0.07

Table 13: Test of P-vector using the vector associated with the minimum eigenvalue of Σ. Top: Posterior returns compared to equilibrium returns. Bottom: New weights compared to equilibrium weights.

	EUSA	EWG	EWH	EWQ	EWU	MCHI
Π (%)	4.56	9.90	6.74	10.16	7.65	9.39
$E(r)$ (%) =	7.95	16.06	11.81	16.51	12.77	16.41
$\Delta_r = \frac{E(r)-\Pi}{\Pi}$ (%)	74.34	62.25	75.31	62.44	67.01	74.85
w_{eq} (%)	-6.90	35.57	10.17	33.37	13.05	14.73
Wgt (%)	0.80	49.10	21.49	47.33	24.46	30.40
$\Delta_w = \frac{Wgt-w_{eq}}{w_{eq}}$ (%)	-111.6	38.01	111.4	41.81	87.39	106.4

Table 14: Test of P-vector using the vector associated with the maximum eigenvalue of Σ. Top: Posterior returns compared to equilibrium returns. Bottom: New weights compared to equilibrium weights.

5.3.3 Compression of large covariance matrices

The "real input" data in this project only consists of 6 assets. With modern CPU's much more data can be processed. However, the topic of decomposition into eigenvalues and eigenvectors is so close related to the topics and techniques that have already been illustrated. The same techniques can be applied to e.g. image or sound compression. The method described here is named PCA or Principal Component Analysis but it is also known as SVD or Singular Value Decomposition and the method has more names. From Equation (102) we already know that what we could interpret as almost 90% of the total variance is described by the fifth and sixth eigenvector. If we also include the fourth eigenvector, we describe 93,68% of the total variance in the covariance matrix assuming we use the corresponding eigenvectors to describe the data (which was previously illustrated by a cloud). The singular value decomposition theorem states that

$$\mathbf{M} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{104}$$

we can decompose our matrix $\mathbf{M}$ into 3 other matrices that when multiplied together approximates **M**, depending on how many eigenvectors and eigenvalue pairs are included. Using the implementation on my system, **U** is the matrix containing all eigenvectors which are described by (100) and (101) (but upside-down). It means $\mathbf{U} = \{\mathbf{V_1},\ \mathbf{V_2},\ \mathbf{V_3},\ \mathbf{V_4},\ \mathbf{V_5},\ \mathbf{V_6}\}$. The matrix $\boldsymbol{\Sigma}$ is an identity matrix whose elements are the λ-values shown to the left in Equation (102) (again upside-down). Finally the matrix $\mathbf{V} \approx \mathbf{U}$ down to machine precision, i.e. when subtracting the two matrices from each other, numbers in the order of 10^{-15} appear.

Now, we can choose to reconstruct the covariance matrix using e.g. a 4-asset space instead of 6 assets (but these 4 assets are not the samme as before):

$$\Sigma = 10^{-3} \cdot \begin{bmatrix} 1.0000 & 0.6352 & 0.5396 & 0.6373 \\ 0.6352 & 1.0000 & 0.5724 & 1.0000 \\ 0.5396 & 0.5724 & 1.0000 & 0.5704 \\ 0.6373 & 1.0000 & 0.5704 & 1.0000 \end{bmatrix} \quad \Rightarrow \tag{105}$$

$$\text{Cor} = \begin{bmatrix} 1.0000 & 0.6352 & 0.5396 & 0.6373 \\ 0.6352 & 1.0000 & 0.5724 & 1.0000 \\ 0.5396 & 0.5724 & 1.0000 & 0.5704 \\ 0.6373 & 1.0000 & 0.5704 & 1.0000 \end{bmatrix} \tag{106}$$

which for the top 4 rows and 4 columns, have many similarities with (98) and (99). In this case it is clear that Σ acts as a transformation matrix, because the new 4-asset space is different from the old 6-asset space. The effect of dimension reduction can best be understood by asking one-self, which is the best transformation to go from a higher asset space into a lower asset space? It can best be illustrated by previous figures going from a 3D cloud to a 2D representation of the cloud. It is completely acceptable to only look at the 3D cloud in 2D (in the $x-y$-plane), when the variance in the z-direction is low. The exact same thing happens with dimensionality reduction, when using SVD (or PCA, it is the same). The concept will however not be illustrated in greater detail than described here.

5.4 Partial conclusion on the use of the Black-Litterman model

The previous pages explains the most important conclusions, of the project. From the previous partial conclusion we knew that the Black Litterman model did not always give intuitive results when we only looked at the views Q and the corresponding P-vector (or matrix, for more views).

While the last partial conclusion maybe in a way opened up more questions (and problems) than it answered, by validating the BL model using the frameworks of He and Litterman (2002) and Idzorek (2005), the new experiments sheds light on some of the problems with e.g. correlation between assets which is not always intuitive to see by only looking at the Q and P-vectors or matrices.

The first pages starts out with a really simple 3×3 covariance matrix and then we investigate the questions:

1. What happens when everything is fixed and we only change the P-vector to switch between absolute and relative views? The answer is shown in Figure 19.

2. What happens when we add an extra absolute view and keep everything else constant? The answer is shown in Figure 20.

3. What happens when we keep the first absolute view, but add a relative view as number? The answer is shown in Figure 22. What if we change the views? See Figure 18.

4. Next, we investigate the effect of τ on both $E(r)$ and the new weights. What happens? See Figure 23.

5. What happens to $E(r)$ and the weights if we modify the user-specified confidence level as described by Idzorek (2005)? See Figure 25.

Furthermore, even for a fixed Q it does definately not matter how much "amplification" is given by considering the covariance matrix as a linear transformation operator. The effect of minimum and maximum eigenvalues and eigenvectors have been illustrated in not only 2 and 3 dimensions, but also in higher dimensions (although that has not been graphically visualized because that is not really possible).

Generally there are many ways to make "experiments". It is the authors belief that the best thing is to only change a few variables while fixing all other parameters, otherwise it is nearly impossible to try to understand what happens. Many experiments have been made on the previos pages. Maybe many other things could have been plotted or illustrated, but these things are the most important ideas that I found to be interesting to study.

6 Conclusions

The problem definition or statement described that the main research question was:

> *How to use and construct portfolios (how to mix assets) with traditional portfolio optimization theory and also by using the Black-Litterman model? What is the impact or effect from individual components of the Black-Litterman equation?*

This report describes using a step-by-step approach how to go from traditional portfolio optimization into using the Black-Litterman model. A lot of illustrative examples are shown, that hopefully should make this report easy to read for everyone who have not worked with the model before. The effect or impact of many individual components in the Black-Litterman equations have been described, studied, analyzed and many figures and illustrations have been made to visually better understand the model (partly on purpose, because illustrations makes everything much easier to read and understand).

The report describes how to calculate historical covariance matrices using both a traditional method and using the index-model method. Unconstrained reverse portfolio optimization have been carried out and a lot of not only theoretical but also historical background information has been included, where it was deemed appropriate to give credit.

One of the interesting things to study have been the effect of changing the views using both a very artificially simple covariance matrix, a not so simple covariance matrix and a more “normal” covariance matrix, based on real historical stock quotes. It can be concluded that the complexity increases to higher levels, as more and more correlations affects the views. The model uncertainty parameters have been described from a practical as well as from a theoritical point of view. Furthermore, there is a discussion about the influence of creating views using large or small confidence. The title and topic of the thesis “An investigation into the BL model” is maybe quite appropriate, as it seems to take a while before getting perfectly comfortable and familiar with the framework. It is my hope, that the report together with all the illustrations and background information can help other people getting more familiar with the model.

Appendices

A Extract from from the `BayesianDialog.java` file

Results from figures 16 and 17 starting from page 40 have been verified using the Akutan open source finance project[16]. It is written in Java and compiles using e.g. `ant build` (compile and build a jarfile) or `ant javac` (compile). It was a requirement to add the following to the `CLASSPATH` environment:

- akutan/lib/colt.jar
- akutan/lib/jfreechart-1.0.5.jar
- akutan/lib/jcommon-1.0.9.jar
- akutan/lib/lpsolve55j.jar
- akutan/lib/junit-4.4.jar

Having done this, it was easy to insert print-statements as shown below.

```
public XYDataset createDataset() {
  final int NUM_POINTS = 80;
  System.out.println("createDataset, ret/sigma=");
  NormalDistributionFunction2D norm = new
      NormalDistributionFunction2D(0, 1);
  XYSeriesCollection tColl = (XYSeriesCollection)
      DatasetUtilities.sampleFunction2D(norm, -3,
      +3, NUM_POINTS, "");
  XYSeries tSeries = tColl.getSeries(0);
  XYSeriesCollection sColl = new XYSeriesCollection();
  BLView v = _views.get(0);
  double ret[] = new double[3];
  //System.out.println("v.getEr()=" + v.getEr() + "
      v.getWeights().get(_assetId)=" + v.getWeights().get(_assetId));
  ret[0] = v.getEr() * v.getWeights().get(_assetId)
      + (((!v.isAbsolute()) ? _eRet.get(_assetId) : 0));
  ret[1] = _eRet.get(_assetId);
  ret[2] = _pRet.get(_assetId);
  double sigma[] = new double[3];
  sigma[0] = Math.sqrt(v.getOmega() / _tau);
```

[16] http://www.akutan.org/

```
  sigma[1] = Math.sqrt(_tau * _eVar.get(_assetId, _assetId));
  sigma[2] = Math.sqrt(_pVar.get(_assetId, _assetId)
      - (((_adjustVar) ? _eVar.get(_assetId, _assetId) : 0)));
  final String[] names = new String[] { "Views", "Equilibrium",
      "Posterior" };
  System.out.println(" ret[0]: " + ret[0] + " ret[1]: " + ret[1] + "
      ret[2]: " + ret[2]);
  System.out.println(" sigma[0]: " + sigma[0] + " sigma[1]: " + sigma[1]
      + " sigma[2]: " + sigma[2]);
  System.out.println(" sigma[2]=> _pVar.get(_assetId, _assetId): " +
      _pVar.get(_assetId, _assetId) + " _adjustVar: " +
      _eVar.get(_assetId, _assetId) );
  System.out.println("createDataset...");
  // Don't sort, allow duplicates
  XYSeries[] s = new XYSeries[3];
  for (int j = 0; j < NUM_POINTS; ++j) {
    double x = tSeries.getX(j).doubleValue();
    double y = tSeries.getY(j).doubleValue();
    for (int i = 0; i < 3; ++i) {
      if (s[i] == null) s[i] = new XYSeries(names[i], false, true);
      double adjX = ret[i] + x * sigma[i];
      s[i].add(adjX, y);
    }
  }
  for (int i = 0; i < 3; ++i) {
    sColl.addSeries(s[i]);
  }
  return sColl;
}
```

B Eigendecomposition of additional tiny 2×2 covariance matrices

For illustrative purposes, in addition to Figures 26 and 27, the following clouds of artificial returns and relevant parameters from covariance matrices have been included.

$$\Sigma = \begin{bmatrix} 0.6 & 0.3 \\ 0.3 & 0.6 \end{bmatrix} \approx \begin{bmatrix} 0.6624 & 0.3180 \\ 0.3180 & 0.5866 \end{bmatrix} \tag{107}$$

$$V_1 = \begin{bmatrix} 0.6640 \\ -0.7478 \end{bmatrix}, \quad V_2 = \begin{bmatrix} -0.7478 \\ -0.6640 \end{bmatrix} \tag{108}$$

$$\lambda = \begin{bmatrix} 0.3043 \\ 0.9447 \end{bmatrix} \tag{109}$$

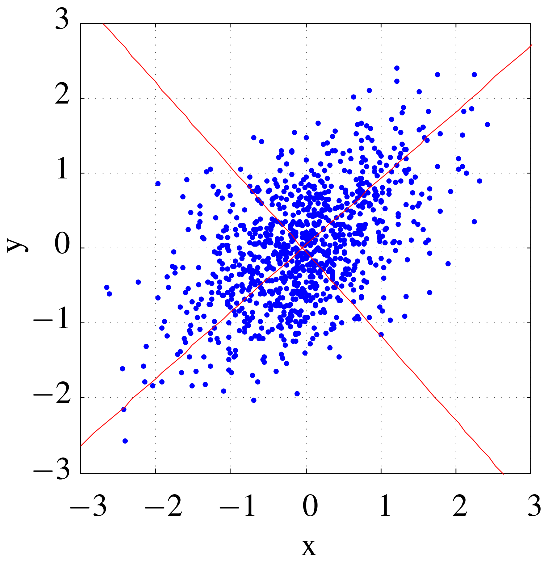


Figure 31: The slope of the most important direction is $\tan^{-1}(-0,664/-0.7478) \approx 42°$.

$$\Sigma = \begin{bmatrix} 1 & 0.3 \\ 0.3 & 0.3 \end{bmatrix} \approx \begin{bmatrix} 1.0148 & 0.3208 \\ 0.3208 & 0.3067 \end{bmatrix} \tag{110}$$

$$V_1 = \begin{bmatrix} 0.3598 \\ -0.9330 \end{bmatrix}, \quad V_2 = \begin{bmatrix} -0.9330 \\ -0.3598 \end{bmatrix} \tag{111}$$

$$\lambda = \begin{bmatrix} 0.1830 \\ 1.1385 \end{bmatrix} \tag{112}$$

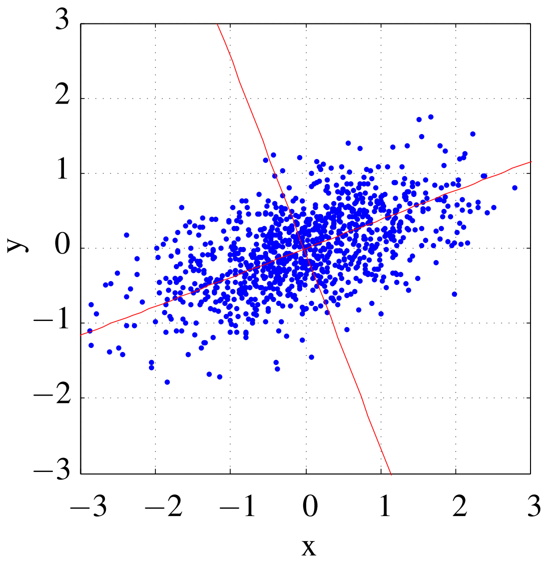


Figure 32: Here Σ corresponds to the direction: $\tan^{-1}(-0,3598/-0.9330) \approx 21°$. Hence the risk/volatility or variance for asset x is much greater than the risk for asset y.

$$\Sigma = \begin{bmatrix} 0.3 & 0.3 \\ 0.3 & 1.0 \end{bmatrix} \approx \begin{bmatrix} 0.2910 & 0.3046 \\ 0.3046 & 1.0048 \end{bmatrix} \tag{113}$$

$$V_1 = \begin{bmatrix} -0.9383 \\ 0.3459 \end{bmatrix}, \quad V_2 = \begin{bmatrix} 0.3459 \\ 0.9383 \end{bmatrix} \tag{114}$$

$$\lambda = \begin{bmatrix} 0.1787 \\ 1.1171 \end{bmatrix} \tag{115}$$

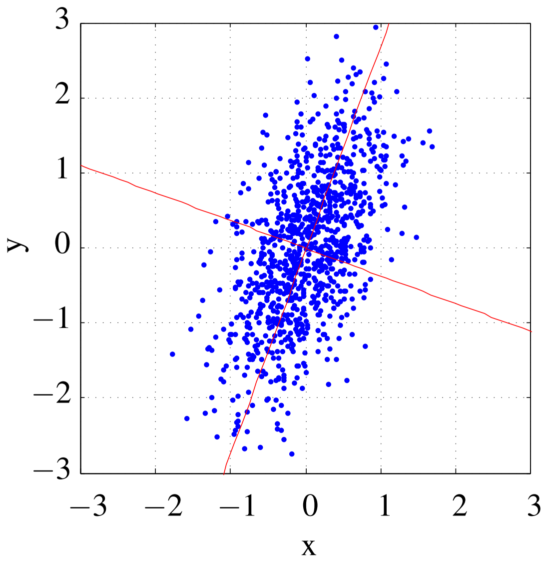


Figure 33: Here Σ corresponds to the direction 70°. Hence the risk/volatility or variance for asset y is much greater than the risk for asset x. With CAPM β would be greater for asset y than for asset x, which is the opposite result when comparing to the previous figure.

$$\Sigma = \begin{bmatrix} 0.3 & -0.2 \\ -0.2 & 0.3 \end{bmatrix} \approx \begin{bmatrix} 0.3016 & -0.1986 \\ -0.1986 & 0.3077 \end{bmatrix} \tag{116}$$

$$V_1 = \begin{bmatrix} -0.7125 \\ -0.7017 \end{bmatrix}, \quad V_2 = \begin{bmatrix} -0.7017 \\ 0.7125 \end{bmatrix} \tag{117}$$

$$\lambda = \begin{bmatrix} 0.1060 \\ 0.5033 \end{bmatrix} \tag{118}$$

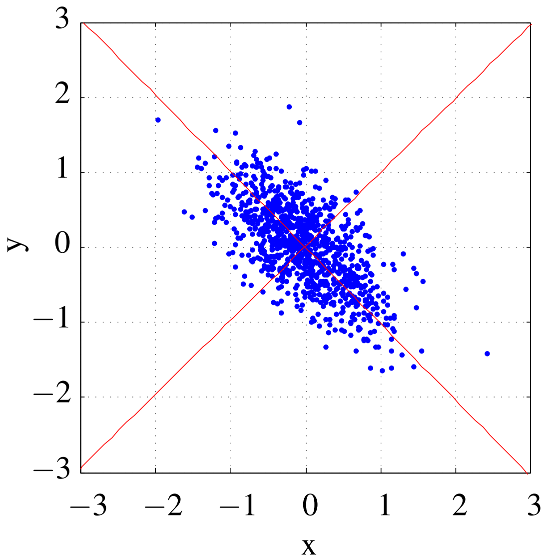


Figure 34: In this case it is clear that the negative offdiagonal elements of Σ makes assets x and y negatively correlated (angle of most important eigenvector is -45°). Such a covariance matrix is perfect for risk diversification of investment portfolios.

C References

Steven L Beach and Alexei G Orlov. An application of the black–litterman model with egarch-m-derived views for international portfolio management. *Financial Markets and Portfolio Management*, 21(2):147–166, 2007. 18

Simon Benninga et al. *Financial modeling*. The MIT Press, 2008. 9, 10, 11

Michael J Best and Robert R Grauer. Sensitivity analysis for mean-variance portfolio problems. *Management Science*, 37(8):980–989, 1991. 35

Andrew Bevan and Kurt Winkelmann. Using the black-litterman global asset allocation model: three years of practical experience. *Fixed Income Research*, 1998. 11

Fischer Black. Capital market equilibrium with restricted borrowing. *The Journal of Business*, 45(3):444–455, 1972. 10, 11

Fischer Black and Robert Litterman. Global portfolio optimization. *Financial Analysts Journal*, 48(5):28–43, 1992. 3, 11, 18, 36

Daniel Blamont and N Firoozy. Asset allocation model. *Global Markets Research: Fixed Income Research*, 2003. 18

Zvi Bodie, Alex Kane, and Alan J. Marcus. *Investments (10th ed, global edition)*. Mc Graw Hill Education. Berkshire, 2014. 1, 2, 8, 9

Guangliang He and Robert Litterman. The intuition behind black-litterman model portfolios. *Available at SSRN 334304*, 2002. ii, 16, 17, 37, 40, 42, 45, 48, 54, 70

Thomas M Idzorek. A step-by-step guide through the black-litterman model, incorporating user specified confidence levels. *Chicago: Ibbotson Associates*, pages 1–32, 2005. ii, 13, 14, 15, 18, 19, 20, 43, 44, 46, 48, 57, 70, 71

Richard A Johnson. *Probability and statistics for engineers*. 2000. 22

Wai Lee. *Theory and Methodology of Tactical Asset Allocation*, volume 65. John Wiley & Sons, 2000. 18

John Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The review of economics and statistics*, pages 13–37, 1965. 3

Bob Litterman et al. *Modern investment management: an equilibrium approach*. Wiley, 2003. 19

Harry Markowitz. Portfolio selection. *The journal of finance*, 7(1):77–91, 1952. 3

Attilio Meucci. *Risk and Asset Allocation*. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg, 2005. 16, 17

Attilio Meucci. The black-litterman approach: Original model and extensions. the encyclopedia of quantitative finance, 2010. 16

Jan Mossin. Equilibrium in a capital asset market. *Econometrica: Journal of the econometric society*, pages 768–783, 1966. 3

Svetlozar Todorov Rachev. *Handbook of Heavy Tailed Distributions in Finance: Handbooks in Finance*, volume 1. Elsevier, 2003. 35

Richard Roll. A critique of the asset pricing theory's tests part i: On past and potential testability of the theory. *Journal of financial economics*, 4(2):129–176, 1977. 7

Stephen Satchell. *Forecasting expected returns in the financial markets*. Academic Press, 2011. 12

Stephen Satchell and Alan Scowcroft. A demystification of the black–litterman model: Managing quantitative and traditional portfolio construction. *Journal of Asset Management*, 1(2):138–150, 2000. 7, 12, 13, 15, 16, 21

William F Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, 19(3):425–442, 1964. 3

James Tobin. Liquidity preference as behavior towards risk. *The review of economic studies*, 25(2):65–86, 1958. 3

CFA Walters et al. The factor tau in the black-litterman model. *The Factor Tau in the Black-Litterman Model (October 9, 2013)*, 2013. 13

CFA Walters et al. The black-litterman model in detail. *The Black-Litterman Model in Detail (June 20, 2014)*, 2014. 11, 14, 17, 18, 21, 39

J. Walters. The black-litterman model: A detailed exploration. *The Black-Litterman Model: A Detailed Exploration*, 2008. 39

Yahoo.com. Yahoo!finance. `http://finance.yahoo.com`, 2016. Accessed: 2016-02-29. 25