



TURUN KAUPPAKORKEAKOULU
Turku School of Economics

**GLOBAL TACTICAL ASSET ALLOCATION
USING THE BLACK-LITTERMAN MODEL**

Master's Thesis
in Economics

Author
Lauri Hälikkä 10057

Supervisors
Ph.D. Mika Widgren
Ph.D. Hannu Vartiainen

1.5.2009
Turku

TABLE OF CONTENTS

1	INTRODUCTION	4
1.1	Motivation.....	4
1.2	Aim of the study.....	5
1.3	Structure of the thesis.....	5
2	PORTFOLIO OPTIMIZATION PROBLEMS.....	7
2.1	Portfolio optimization	7
2.1.1	Historical returns	11
2.1.2	Volatility and correlation.....	12
2.1.3	Constraints.....	13
2.1.4	Capital Asset Pricing Model.....	14
2.1.5	Arbitrage Pricing Theory.....	16
2.1.6	Limitations.....	17
2.2	Portfolio resampling.....	20
2.2.1	Measuring estimation error.....	20
2.2.2	Resampled efficient frontier	22
2.2.3	Resampled portfolio efficiency	25
2.3	Shrinkage estimation.....	29
2.3.1	Historical returns	30
2.3.2	Covariance matrix	30
2.3.3	Asset pricing models	32
3	ASSET ALLOCATION	33
3.1	Introduction to asset allocation decisions: SAA vs. TAA	33
3.2	Benchmark selection.....	34
3.2.1	Naive allocation: 1/n.....	34
3.3	Tactical asset allocation: The Black-Litterman approach.....	36
3.3.1	Derivation of the Black-Litterman model: the Bayesian theorem... ..	37
3.3.2	Reverse optimization and the equilibrium returns.....	42
3.3.3	Expressing views: the view vector and the matching matrix	43
3.3.4	Incorporating uncertainty into the views and scalar τ	47
3.4	Risk factors in the returns on stocks and bonds.....	47
3.5	Performance measures	54
3.5.1	Active return and active risk.....	54
3.5.2	Residual return and residual risk	54
3.5.3	Information ratio.....	56
3.5.4	Value added	57

4	EMPIRICAL STUDY	60
4.1	Methodology	60
4.2	Description of the data	60
4.3	Input estimation	62
4.4	Strategic weights	65
4.5	Vector autoregressive model for return forecasting.....	68
4.6	Overview on the result	72
5	CONCLUSION OF THE STUDY	75
6	REFERENCES	76

LIST OF FIGURES AND TABLES

Figure 1. Efficient frontier in a risk-return space.....	9
Figure 2. Convergence of the distance measure when short sales are allowed (T=30). ..	22
Figure 3. Estimation error visualized: the dispersion of efficient frontiers.....	23
Figure 4. Error inheritance (Scherer 2004). Omega stands for the covariance matrix... ..	23
Figure 5. Convergence of the distance measure when short sales are not allowed (T=30).....	26
Figure 6. Portfolio weights along the efficient frontier (short sales not allowed).....	27
Figure 7. Portfolio weights along the efficient frontier (short sales not allowed, 500 simulation runs per portfolio).....	28
Figure 8. Intuition behind the Black-Litterman process (Drobetz, 2002).....	37
Figure 9. The process behind the new combined return vector (Idzorek 2004).....	40
Figure 10. Stock price series calculated from monthly excess log-returns.....	62
Figure 11. Bond price series calculated from monthly excess log-returns.....	62
Figure 12. Portfolio weights when short sales are not allowed. Annual target excess return is 3%.	65
Figure 13. Portfolio weights when short sales are not allowed and covariance matrix shrinkage is used. Annual target excess return is 3%.....	66
Figure 14. Strategic in-sample performance. Time series are calculated on excess return basis.	67
Figure 15. Strategic out-of-sample portfolio performance. Time series are calculated on excess return basis.	68
Figure 16. State variable behavior over sample.....	71
Figure 17. TAA and SAA out-of-sample performance.....	73
Figure 18. Out-of-sample portfolio weights.....	74
Table 1. Descriptive in-sample statistics of the asset classes. Numbers are monthly results.....	61
Table 2. Expected annual excess returns using different return estimation methods. Time period in sample data 1/1988-12/1999.....	64
Table 3. Strategic monthly performance of different strategic weights.....	67
Table 4. Estimated VAR-model for return forecasting. Estimation period is in-sample 1/1988-12/1999.	71
Table 5. Monthly TAA out-of-sample results.....	73

1 INTRODUCTION

1.1 Motivation

For a long time, the demand for quantitative portfolio management has increased among the financial institutions, but practical tools available have been very poor. Although Harry Markowitz invented the solution for an optimal portfolio in 1952, the number of professional users has been very small. Among practitioners, the portfolio optimization between risk and return has been difficult to implement into practice for several reasons. Although the method is very straightforward in theory and easy to implement especially now in the computer age, the results are useless. Optimal portfolio weights given by Markowitz's solution are way too extreme to implement into practice and results where assets have great negative weights are frequent. Primarily portfolio optimization is a statistical problem, because the quality of the data available for calculating expected returns is too modest compared to powerful optimization algorithm. The key problem is that input estimates contain an estimation error, but algorithm treats them as true values. Fortunately there are methods that can be used to achieve more convenient optimal portfolios and this study covers some of the most common ones.

The Markowitz approach to portfolio selection is myopic and it is a one-period model. In practice, asset managers use longer horizon allocation and besides that, they do short-term changes in the underlying long-term portfolio in order to raise long term portfolio's risk adjusted return. That's why this study takes a few steps further from the Markowitz's approach and tries to offer a more sophisticated way to complete the task of today's active portfolio manager. The classical framework for portfolio optimization with possible improvements will be introduced and the solution for two-period investment problem called the Black-Litterman model is studied thoroughly. The main improvements on estimation error problem studied in the thesis are shrinkage estimation and portfolio resampling. Short-period investing decisions are always based on forecasts. Hence we will construct a vector autoregressive model for return forecasting and the whole process is tested empirically¹. The results of this study show, that model based portfolio management can raise risk adjusted return compared to the passive long-term portfolio (benchmark).

¹ I would like to thank Hannu Kahra & Juha Joenväärä from University of Oulu and Tommi Sjöblom from Turku School of Economics for insightful comments and R-code support.

1.2 Aim of the study

The main goal of this study is to show how the Black-Litterman model is implemented into practice and test empirically if the Black-Litterman model can beat the traditional static one period model. Besides that, we go through the conventional portfolio optimization procedure á la Markowitz and discuss the reasons why it is so difficult to implement it into practice. Moreover, we will go through some practical tools that can be used to enhance static portfolio optimization results to be more applicable. These more convenient results will be used as a benchmark.

As we know, the Markowitz solution is a one-period case, and it can be considered as strategic allocation i.e. the long-term asset allocation strategy. However, practitioners tend to consider investment decisions as a two-period problem where short-term (tactical) investment decisions differ from long-term (strategic) goals. The model is known as the Black-Litterman model and it's a method where an investor can combine, in a very flexible way, his strategic asset allocation decisions with short term views about market movements, while trying to achieve higher risk-return ratio.

As a whole, the global tactical asset allocation is a complex problem and it needs a lot of things to be solved before implementing it successfully into practice. This study is supposed to go through one practical way to solve investor's two-period investment problem and to cover a lot of important information about basic framework of global tactical asset allocation. This study tries to give a good starting point for a practitioner who tries to move from intuitive portfolio construction to model based asset allocation. The approach on the subject is kept practical in order to cover practical problems that a portfolio manager may encounter while implementing the Black-Litterman approach into practice.

1.3 Structure of the thesis

The second chapter covers some general topics on the modern portfolio theory. The basic solution for Markowitz portfolio optimization will be given and then its practical problems and limitations are discussed. The second chapter also covers different techniques to derive input parameters for portfolio optimization and give different solutions for tackling the input estimation error in the optimization process. The method called portfolio resampling is also introduced for measuring the estimation error impact on optimal portfolios. It can also be used to form significantly better behaving portfolios. We will also discuss the results of resampled efficient portfolios.

The third chapter introduces long-term and short-term asset allocation decisions and then concentrates on the Black-Litterman model. Tactical asset allocation is all about forecasting returns in some way or another, so in the third chapter we examine several economic factors that explain risk premiums of stocks and bonds. This chapter also introduces several performance measures to evaluate the results later depicted in chapter four.

Chapter four is empirical, and it will bind together the whole study in a practical way. It starts with description of the data. Then expected returns are estimated using the methods studied in chapter two. Using these results we derive optimal strategic portfolios for all the input estimates. These static portfolios are tested in- and out-of-sample and the rationality of the results is discussed. Then we estimate the vector autoregressive model in order to predict the short-term movements of the assets. Using these short-term forecasts and one of the long-term return estimates, we will construct the Black-Litterman model and run the empirical test for the model. The fifth chapter concludes the whole study.

2 PORTFOLIO OPTIMIZATION PROBLEMS

2.1 Portfolio optimization

Harry Markowitz was officially the first pioneer on the field of optimal portfolio selection. His view on investor maximizing return and minimizing risk is still today a widely accepted approach on portfolio optimization even though there are several problems when implementing it into practice. According to the deserving financial economist Mark Rubinstein (2002), Harry Markowitz's theory of optimal portfolio selection (1952) gave birth to a new era - the modern financial economics. Markowitz rejected the hypothesis that investors should always maximize discounted future returns. If doing so, investors would place all their assets in to a security that maximizes return. Because the future is unknown and returns are stochastic, Markowitz pointed out that it is important to talk about expected returns and take risk into consideration. Variance is a widely accepted measure for dispersion of a random variable and therefore it is also used to measure investment risk. As long as investors consider return as a desirable thing and variance as an undesirable one, investors should maximize their return while minimizing the risk (variance). Markowitz's most important insight was that due to imperfect correlation structure among different assets, portfolio variance can be reduced without lowering expected return - until a certain point. When the allocation is called "efficient", the variance can be reduced only by giving up expected return. Having these facts, Markowitz concentrated on *efficient portfolios* instead of single securities. He also pointed out in his groundbreaking article in 1952 that diversification does not imply just to diversifying funds into an adequate number of different assets, but the "right kind" of assets. Taking advantage of low correlation between different asset classes leads to much more efficient portfolios (Markowitz 1952).

The terms minimum variance portfolio and efficient portfolio should be distinguished. When investors diversify their funds among securities that give the highest possible expected return, this doesn't necessarily mean that it will also generate the minimum variance for the portfolio. The minimum variance portfolio is the saturation point on the efficient frontier that has the lowest variance. After this point, any efficient portfolio on the efficient frontier cannot be attained without increasing variance. The efficient portfolio in a mean-variance space can be described in two different ways. The portfolio is efficient if it has the lowest possible variance for a given level of expected return. The MV efficiency is equivalent with expression that the portfolio is efficient when it has the maximum expected return for a given level of risk.

Suppose that an investor can allocate his wealth to N risky assets that have mean vector μ of excess returns $r_i = R_i - r_f$, where R_i is the nominal return of the asset i ($i=1\dots N$) and r_f is a risk free asset with known return. We define covariance matrix of the N asset returns as Σ . Assume now that the means are *iid*² with constant moments. (Brandt, 2004)

In the absence of a risk-free asset, the mean variance problem is to choose portfolio weights w_i , which represent the investor's relative allocations of wealth to each of the N risky assets, to minimize the variance of the resulting portfolio return $R_{pfo} = w'R$, for pre-determined target expected return $r_f + \bar{\mu}$:

$$\min_w \text{var}[R_{pfo}] = w'\Sigma w, \quad (2.11)$$

$$E[R_{pfo}] = w'(r_f + \mu) = (r_f + \bar{\mu}) \text{ and } \sum_{i=1}^N w_i = 1 \quad (2.12)$$

The first constraint fixes the expected return of the portfolio to its target, and the second constraint ensures that all wealth is invested in the risky assets. Setting up Lagrangian and solving the corresponding first-order conditions, the optimal portfolio weights are:

$$w^* = \Lambda_1 + \Lambda_2 \bar{\mu} \quad (2.13)$$

where

$$\Lambda_1 = \frac{1}{D} [B(\Sigma^{-1}\mathbf{1}) - A(\Sigma^{-1}\mu)] \text{ and } \Lambda_2 = \frac{1}{D} [C(\Sigma^{-1}\mu) - A(\Sigma^{-1}\mathbf{1})].$$

$\mathbf{1}$ is an appropriately sized vector of ones, μ is a vector of expected returns and where $A = \mathbf{1}'\Sigma^{-1}\mu$, $B = \mu'\Sigma^{-1}\mu$, $C = \mathbf{1}'\Sigma^{-1}\mathbf{1}$ and $D = BC - A^2$. The minimized portfolio variance with a given return is equal to $w^{*\prime}\Sigma w^*$. (Brandt, 2004)

The Markowitz paradigm yields two important economic insights. First, it illustrates the effect of diversification. The variance of the portfolio return is always less than the weighted sum of the portfolio's components when the correlation³ between at least two components is less than 1. Imperfectly correlating assets can be combined into portfolios that have preferred expected return-risk characteristics. Second, the Markowitz paradigm shows that, once the portfolio is fully diversified, higher expected return can only be achieved by more extreme allocations (notice w^* is linear in $\bar{\mu}$ in the equation 2.13) and therefore by taking more risk (see Figure 1).(Brandt, 2004)

² Independent and identically distributed

³ The correlation coefficient $\rho_{i,j} = \frac{\sigma_{i,j}}{\sigma_i\sigma_j}$

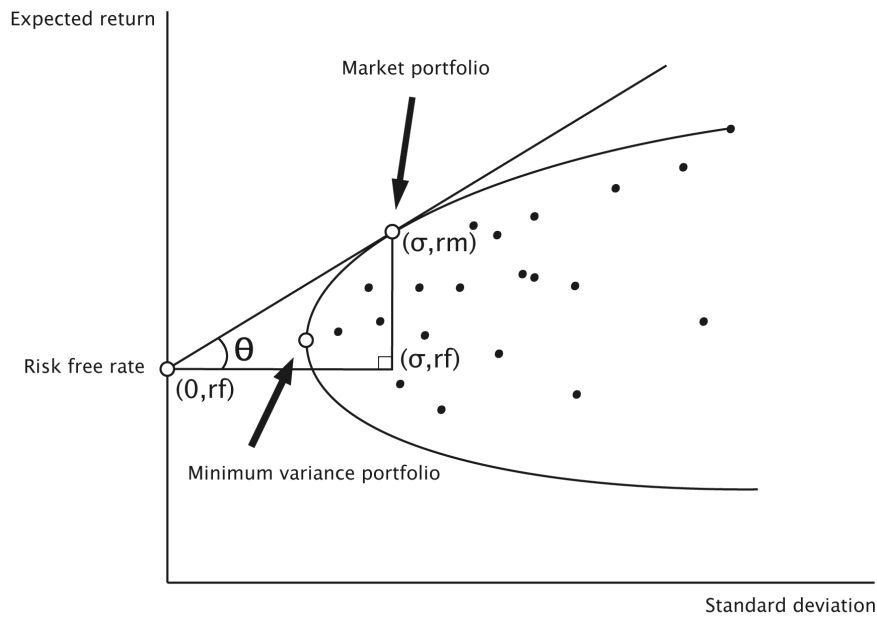


Figure 1. Efficient frontier in a risk-return space

Figure 1. illustrates these two important insights. The figure plots the mean-variance frontier as hyperbola generated by the historical annualized moments of asset returns. Each point on the frontier along the horizontal axis gives the minimized portfolio return volatility for a pre-determined expected portfolio return along the vertical axis. The dots inside the hyperbola represent the individual assets from which the frontier is constructed of. The fact that these dots lie inside the frontier is due to the covariance structure of the assets and illustrates the effect of diversification. The individual assets can be combined to generate higher expected returns with the same or a lower volatility. The figure also illustrates the fundamental trade-off between expected return and risk. Starting from the most diversified portfolio (left tip of the hyperbola), the minimum variance portfolio (Cochrane 2005, 83) is calculated by

$$w_{\min \text{ var}} = \frac{\Sigma^{-1} \mathbf{1}_N}{\mathbf{1}' \Sigma^{-1} \mathbf{1}_N}. \quad (2.14)$$

The higher expected return could be achieved only at the cost of risk (volatility).

If the investor can also allocate wealth to the risk-free asset, in the form of unlimited risk-free borrowing and lending at the risk-free rate r_f , any portfolio on the mean-variance frontier generated by the risky assets can be combined with the risk-free asset on the vertical axis to generate an expected return-risk profile that lies on a straight line from a risk-free rate (no risky assets) through the frontier portfolio (fully invested in risky assets) and beyond (leveraged risky investments). The optimal combination of the risky frontier portfolios with risk-free borrowing and lending is the one that maximizes

the Sharpe ratio⁴ of the overall portfolio. Sharpe ratio can also be presented graphically by the slope of the line from the risk-free asset through the risky frontier portfolio. In fact, maximizing the Sharpe ratio means that we maximize the left angle of the triangle $(0, r_f)$, (σ_{pfo}, r_f) and $(\sigma_{pfo}, \bar{\mu})$ (Figure 1.):

$$Max \tan \theta = \frac{w' \bar{\mu}}{\sqrt{w' \Sigma w}}. \quad (2.15)$$

The highest obtainable Sharpe ratio is achieved by upper tangency on the hyperbola (market portfolio). When the risk free asset is available, it leads investors to allocate part of their wealth w^* to the efficient combination of risky assets and the remainder $(1 - I'w^*)$ to the risk-free asset r_f . Any linear combination of the risk free asset and the risky portfolio (market portfolio) on the mean-variance frontier can be invested in. Every investor combines only these two alternatives according to their risk appetite. The portfolio return is therefore

$$E[r_{pfo}] = w'(r_f + \mu) + (1 - I'w)r_f \quad (2.16)$$

and the mean-variance problem can be expressed in terms of excess returns:

$$\min_w \text{var}[r_{pfo}] = w' \Sigma w,$$

subject to

$$E[r_{pfo}] = w' \mu = \bar{\mu}. \quad (2.17)$$

When there is a risk free asset available, the case is much simpler:

$$w^* = \frac{\bar{\mu}}{\underbrace{\mu' \Sigma^{-1} \mu}_{\lambda}} \times \Sigma^{-1} \mu \quad (2.18)$$

where λ is a constant that scales proportionately all elements of $\Sigma^{-1} \mu$ to the desired portfolio risk premium $\bar{\mu}$. Because we are dealing with tangency portfolio on the efficient frontier, weights in that portfolio must sum to one (zero weight on risk-free asset). The λ and $\bar{\mu}$ for the tangency portfolio are

$$\lambda_{tgc} = \frac{1}{I' \Sigma^{-1} \mu} \quad \text{and} \quad \bar{\mu}_{tgc} = \frac{\mu' \Sigma^{-1} \mu}{I' \Sigma^{-1} \mu}. \quad (\text{Brandt, 2004}) \quad (2.19)$$

⁴ Sharpe ratio $\frac{E(r) - r_f}{\sigma}$

Since we have gone through analytically the optimal portfolio choice and have shown that every investor combines the portfolio that has the largest Sharpe ratio and risk-free asset (the two fund separation theorem), it would be relevant to ask, what is the optimal allocation to certain investor. In order to get the answer, we need to know investor's tolerance for risk. To incorporate the risk aversion to the mean-variance problem, we can formulate the problem alternatively by maximizing investor's quadratic utility function:

$$\max U = w'\mu - \frac{\delta}{2}w'\Sigma w, \quad (2.20)$$

where U is the investor's utility (objective function) and

$$\lambda = \frac{1}{\delta} \quad (2.21)$$

measures the investor's level for relative risk aversion⁵. U is a concave function and thus it will have a single global maxima. Without any constraints, we find a closed form solution taking the first derivative of the function respect to weights and setting it to 0:

$$\frac{dU}{dw} = \mu - \delta\Sigma w = 0 \quad (2.22)$$

Solving this for w , we get the optimal weight vector:

$$w^* = (\delta\Sigma)^{-1}\mu, \quad (2.23)$$

which explicitly links the optimal portfolio to the tangency portfolio of the investor's tolerance for risk. (Brandt, 2004)

2.1.1 Historical returns

One of the most difficult tasks for a portfolio manager, and at the same time a highly important procedure, is expected return estimation. There are two methods used to calculate returns from series of historical prices. A very common way to calculate expected return is to use historical sample returns:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \times 100\%, \quad (2.28)$$

⁵ market risk aversion parameter is $\delta = \frac{E(r) - r_f}{\sigma^2}$

where p_t denotes the asset price at time t . Another option is to use continuously compounded returns. This approach is usually employed in the academic finance literature and it's also called log-return formulation:

$$r_t = 100\% \times \ln\left(\frac{p_t}{p_{t-1}}\right). \quad (2.29)$$

Taking logarithm constitutes non-linear transformation and that is why continuously compounded returns are time-additive. Due to the additivity of logarithmic returns, only the first and last observations are needed to calculate the expected return (Brooks, 2002). It will be shown in chapter 2.1.6, that expected returns derived from historical returns, include a relatively large estimation error and leads to unintuitive portfolio allocations. There are alternative ways to estimate expected returns that give more sensible allocations as a result in classical Markowitz optimization. Alternative methods will be discussed later.

2.1.2 Volatility and correlation

Since *standard deviation* (i.e. *volatility*) is a square root of variance, let's look at the concept of variance first. The variance is a simple measure of variation around the average. In other words, it measures the dispersion of a set of observations around the mean value. Mathematically, variance is defined as the average of the sum of squared differences between the returns and the mean of the returns:

$$\bar{r}_A = \frac{\sum_{t=1}^T r_{A,t}}{n}, \quad Var_A = \frac{1}{n-1} \sum_{t=1}^T (r_{A,t} - \bar{r}_A)^2 \quad (2.24)$$

Where \bar{r}_A is the mean of the variable $r_{A,t}$, when $t=1, 2, \dots, T$. Because the variance is a squared term, it is not directly comparable to the mean itself. So we need to take a square root of the variance in order to get standard deviation, i.e.

$$Std_A = \sqrt{\frac{1}{n-1} \sum_{t=1}^T (r_{A,t} - \bar{r}_A)^2}. \quad (2.25)$$

Financial risk in absolute terms is usually assessed by calculating the annualized standard deviation of the returns, also called as volatility. The volatility of the returns is thus derived by multiplying the standard deviation of returns with the square root of the sample frequency. The sample frequency is simply the number of observations per year, so that daily observations yield approximately 260, weekly observations yield 52 and

monthly observations yield 12. The volatility of the returns on an asset is thus defined as:

$$\sigma_A = \sqrt{\frac{1}{n-1} \sum_{t=1}^T (r_{A,t} - \bar{r}_A)^2} \cdot \sqrt{\hat{n}} \quad (2.26)$$

where \hat{n} is the sample frequency within one year. (Rasmussen 2003, 28)

The *covariance* measures the co-movement of two variables and it is defined quite similarly to the variance. However, instead of measuring the sum of the squared deviations around the mean of returns on a single asset, the covariance of an asset's return with other variable (e.g. market index or other asset's returns) measures the sum of the deviations in two variables from their respective means. The covariance of an asset with the market is thus defined as:

$$Cov(r_A, r_B) = \frac{1}{n-1} \sum_{t=1}^T (r_{A,t} - \bar{r}_A)(r_{B,t} - \bar{r}_B) \quad (2.27)$$

where $r_{A,t}$ is the return on the asset A at time t , and \bar{r}_A is the mean of asset return over T periods. $r_{B,t}$ is the return on the market at time t , and \bar{r}_B is the mean of the returns on the asset B over T periods. The closer the two returns are – simultaneously – to their respective means, the smaller are the deviations from those means, and the lower is the covariance. (Rasmussen 2003, 32)

2.1.3 Constraints

Maybe the most widely used constraint is a short-selling limitation. While restriction of short selling is very relevant in practice, there may also be a need for limiting maximum weight on one single asset (Brandt, 2004). Frost and Savarino (1988) scientifically show that portfolio constraints really truncate the extreme portfolio weights and thereby improve the performance of estimates. Their results suggest that extreme portfolio weights being truncated really are associated with estimation error. Jagannathan & Ma (2002) empirically show that constraining portfolio weights to be nonnegative is equivalent to using the sample covariance matrix after reducing its large elements (i.e. reducing the sample covariance of the corresponding asset with other assets by a certain amount) and then form the optimal portfolio without any restrictions on portfolio weights. Stocks that have high correlation with other stocks tend to receive negative portfolio weights. Hence, to the some extent, high estimated covariances are more likely to be caused by estimation error and so the short sales restriction can reduce the sampling error. The case is much alike in upper bounds. Upper bound constraint is equivalent to increase the sample covariances of the corresponding asset with other

assets by a certain amount. Since low covariances result in extreme portfolio weights and these low covariances are more likely to be caused by estimation error, imposing upper bound constraint is likely to reduce estimation error. However, their empirical research shows that when the short selling restriction is already in place, the upper bound constraint doesn't improve out-of-sample results significantly.

Constraints can be seen as some forms of shrinkage estimation (see chapter 2.3). Although constrained plug-in estimates are somewhat biased, they are still much less variable than unconstrained portfolios⁶ (Brandt, 2004).

2.1.4 Capital Asset Pricing Model

“Under the CAPM, an individual whose portfolio differs from the market is playing a zero-sum game. The player has additional risk and no additional expected return. This logic leads to passive investing; i.e., buy and hold the market portfolio.”(Grinold & Kahn, 1999)

An alternative way to estimate expected return is the *capital asset pricing model*. One of the valuable outcomes of the CAPM is a procedure for determining consensus expected returns. These consensus expected returns are important because they give us a standard of comparison. As we will later see, tactical asset allocation decisions will be driven by the difference between our expected returns and the consensus. While the CAPM provides a valuable source of consensus expectations, the active manager can succeed to the extent that his or her forecasts are superior to the CAPM consensus forecast. Grinold & Kahn (1999) states that among the other possible expected return estimating methods, the CAPM is arguable the best. They consider APT (see 2.1.5) being interesting tool for this job also, but it is not a source of consensus expected returns. If we use CAPM consensus expected returns to build optimal mean/variance portfolios, those portfolios will consist simply of positions in the market and risk free rate depending on investors risk tolerance. This means that optimal mean/variance portfolios will differ from the market portfolio only then if the forecasted excess returns differ from CAPM consensus excess returns. This is the fact called “consensus”. The market portfolio is the consensus portfolio and CAPM leads to the expected returns that make the market mean/variance optimal (Grinold & Kahn, 1999).

The CAPM relies on two constructs, first the idea of a market portfolio M and notion of the beta β , which links any stock or portfolio to the market. In theory, the market portfolio includes all possible assets such as stocks, bonds, real estate, art etc. In

⁶ see Jagannathan & Ma, 2002 for more information about effects of portfolio constraints

practice, the market portfolio is generally taken as some broad value-weighted index of traded domestic equities, such as the NYSE Composite in the United States, the FTA in the United Kingdom or the TOPIX in Japan (Grinold & Kahn, 1999).

Let's consider any portfolio P with excess returns r_p and the market portfolio M with excess returns r_M . The *beta* of the portfolio P is defined

$$\beta_p = \frac{\text{cov}(r_p, r_M)}{\text{var}(r_M)}. \quad (2.30)$$

Beta is proportional to the covariance between the portfolio's return and the market's return and it is a forecast for the future. Beta for the market portfolio is 1 and for risk free rate 0. Although beta is forward-looking concept, it comes from the simple linear regression of the portfolio excess returns $r_p(t)$ in periods $t=1,2,\dots,T$ on the market excess returns $r_M(t)$ in those same periods. The regression is

$$r_p(t) = \alpha_p + \beta_p r_M(t) + \varepsilon_p(t). \quad (2.31)$$

The estimates of β_p and α_p obtained from regression are historical or realized beta and alpha. These estimates show how the portfolios have interacted in the past. Historical beta is reasonable forecast of the betas that will be realized in the future, although it is possible to do better. There is tendency for betas to regress toward the mean. A stock with a high historical beta in one period will most likely have a lower (but still higher than 1.0) beta in the subsequent period. Similarly, a stock with a lower beta in one period will most likely have a higher (but less than 1.0) beta in the following period. In addition, forecasts of betas based on the fundamental attributes of the company, rather than it's own past, say 60 months, turn out to be much better forecasts of future betas (Grinold & Kahn, 1999).

Beta is a way of separating risk and return into two parts. If we know a portfolio's beta, we can break the excess return on that portfolio into a market component and a residual component:

$$r_p = \beta_p r_M + \theta_p. \quad (2.32)$$

In addition, the *residual return* θ_p will be uncorrelated with the market return r_M and so the variance of portfolio P is

$$\sigma_p^2 = \beta_p^2 \sigma_M^2 + \omega_p^2 \quad (2.33)$$

where ω_p^2 is the residual variance of the portfolio P , i.e. the variance of θ_p . Like this, beta allows us to separate the excess returns of any portfolio into two uncorrelated components – a market return and residual return (Grinold & Kahn, 1999).

The CAPM states that the expected residual return on all assets and any portfolio is equal to zero, i.e. that $E\{\theta_p\} = 0$. This means that the expected excess return on the portfolio, $E\{r_p\} = \mu_p$, is determined entirely by the expected excess return on the market, $E\{r_M\} = \mu_M$, and the portfolio's beta, β_p . The relationship is simple:

$$E\{r_p\} = \beta_p E\{r_M\} = \beta_p \mu_M. \quad (2.34)$$

As this equation clearly illustrates, the portfolio's expected return is proportional to the portfolio's beta. Notice that the CAPM result must hold for the market portfolio. If we sum the returns of all the stocks on a value-weighted basis, we get market return and so the value-weighted sum of the residual returns has to be exactly zero. In addition, if we input the expected returns from the CAPM into mean/variance optimizer, the result will be the market portfolio (Grinold & Kahn, 1999).

2.1.5 Arbitrage Pricing Theory

The arbitrage pricing theory (APT) can also be used to estimate expected returns. It states that the asset's excess return is determined by one or more risk factors and the asset's exposures on these factors. It assumes that there are K factors such that the excess returns can be expressed as

$$r_n = \sum_{k=1}^K \beta_{n,k} \cdot x_k + u_n \quad (2.35)$$

where $\beta_{n,k}$ is the factor loading of asset n to factor k , x_k is the factor return for factor k and u_n is the asset n 's specific return, that cannot be explained by factors (also called the idiosyncratic return on the asset). The main result is that we can express expected excess returns in terms of the model's factor exposures. The APT formula for expected excess return is

$$E\{r_n\} = \sum_{k=1}^K \beta_{n,k} \cdot m_k \quad (2.36)$$

where m_k is the factor forecast for factor k (Grinold & Kahn, 1999). Grinold and Kahn (1999) argue that for factor forecasts $m(k)$ can be used historical mean or trailing 12-months average of the factor returns. The factors can be macroeconomic factors as well as firm specific factors. Factors that influence on stock and bond returns are discussed more closely in chapter 4.3.

The APT result follows from an arbitrage argument. This means that if we find a case where asset's exposure for all the factors is zero (no risk) and there is positive excess expected return, we have possibility to arbitrage. Adding this asset to our portfolio, we gain higher expected return with no additional risk (Grinold & Kahn, 1999).

2.1.6 Limitations

Although Markowitz's solution for the portfolio choice problem gives a good intuition about diversification and risk-return trade off, it makes several defective assumptions. At first, it is clearly a myopic single-period problem (from the present to infinity) because an investor cares only about expected returns of the next period, whereas we think that investment decision involves a long horizon allocation decision (strategic allocation) and medium term portfolio rebalancing (tactical allocation). Second, because the investor cares only about the mean and variance of the portfolio return, it ignores the higher moments of the return distribution. Third, expected utility function is quadratic, which is a special case because it doesn't increase monotonically (Brandt, 2004).

Mean-variance analysis plays an important role among both practitioners and researchers in finance. However, practitioners have reported several difficulties in implementing MV analysis in practice. Essentially mean-variance problem is a statistical problem because optimization algorithm treats estimates as a truth and does not consider sampling errors in mean and covariance estimates. This is the case especially when estimating expected returns from historical data. In fact, the MV optimizer has been called as "error maximizer" (Michaud, 1998), because it tends to maximize sample errors. Small changes in input estimates result very large changes in portfolio allocation and result often extreme weights on assets. Especially in case of unconstrained portfolios, large negative weights are very common. This happens because optimization underweight considerable assets that have small expected return, strong positive correlation and large variance (and vice versa). When short sales are not allowed, resulting portfolios are so-called "corner solutions", when portfolios are highly concentrated on only a few assets. Results are very inconvenient and hard to implement in practice. (Brandt, 2004; Britten-Jones, 1999).

Let's try to estimate the mean return of the sample period. We assume that the expected return \bar{r} and standard deviation σ of the asset are constant in each period. In addition we have to assume that individual returns are mutually uncorrelated (iid). Now suppose that we have n samples of period returns, when the best estimate of the mean return is the average of the sample. Hence,

$$\bar{r} = \frac{1}{n} \sum_{i=1}^n r_i. \quad (2.37)$$

The value \bar{r} itself is a random variable because it's value clearly depends on the sample even if distribution's first moment would remain constant. Now, if we calculate the standard deviation of the estimate \bar{r} , we are able to see how accurate our return estimates are. When expected value of the estimate is

$$E(\bar{r}) = E\left[\frac{1}{n}\sum_{i=1}^n r_i\right] = \bar{r}_{true} \quad (2.38)$$

we get the standard deviation for the estimate \bar{r} and see how accurate the estimate is likely to be

$$\sigma_{\bar{r}}^2 = E\left[(\bar{r} - \bar{r}_{true})^2\right] = E\left[\frac{1}{n}\sum_{i=1}^n (r_i - \bar{r}_{true})\right]^2 = \frac{1}{n}\sigma^2. \quad (2.39)$$

Hence,

$$\sigma_{\bar{r}} = \frac{\sigma}{\sqrt{n}} \quad (2.40)$$

Consider a period length of one month ($p=1/12$). We assume that monthly values for mean and standard deviations are $\bar{r}=1\%$ and $\sigma_{\bar{r}}=4\%$. If we use ten years of monthly data, we get $\sigma_{\bar{r}} = \frac{0,04}{\sqrt{120}} = 0,37\%$ which is 37% of the mean itself. If we use 20 years of monthly data, the standard deviation is still 0,26%. This mean that expected return is 1% +/- 0,26%, which is still very poor estimate given the fact that 4% monthly volatility is extremely low in general. According to Luenberger (1998), a standard deviation of one-tenth of the mean value would be considered as a good estimate. In our example, we would need 156 years of data to get this (it is even abnormal that company exists after this long period of time). In addition, mean values are not likely to be constant over time, so even long period of reliable data wouldn't help us much (Luenberger 1998, 215).

Luenberger (1998, 215) calls a measurement of \bar{r} as a "historical blur problem". It is basically impossible to measure \bar{r} to within workable accuracy using historical data. Furthermore, the problem cannot be improved much by changing the period length. If longer frequency is used, each sample is more reliable, but fewer independent samples are obtained in any year. Conversely, if smaller frequency is used, more samples are available, but each is worse in terms of the ratio of standard deviation to mean value (Luenberger 1998, 215)

Fortunately, the estimation fallacy is not that strong in second moment estimates. Since the sample variance is

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^n (r_i - \bar{r})^2. \quad (2.41)$$

The use of $n-1$ in the denominator instead of n compensates for the fact that \bar{r} is used instead of the true (but unknown) \bar{r}_{true} . It then follows that $E(\sigma^2) = \sigma_{true}^2$ and hence σ^2

is an unbiased estimate of variance. If the original samples are normally distributed, the accuracy of estimate σ^2 is given by its variance:

$$\text{var}(\sigma^2) = \frac{2\sigma^4}{n-1} \quad (2.42)$$

or equivalently

$$\text{std}(\sigma^2) = \frac{\sqrt{2}\sigma^2}{\sqrt{n-1}}. \quad (2.43)$$

If we assume same data length and variance as earlier, the standard deviation of the variance would be 0,024%. This shows that estimation error is not extremely large in volatility when n is reasonable large. (Luenberger 1998, 217)

In order to get realistic picture how these kinds of estimation errors affect on optimal portfolio weights, let's look at standard error of plug-in estimator \hat{w}^* . We have shown earlier that in the case of risk free asset, we get optimal portfolio weights by expression (Brandt, 2004)

$$\hat{w}^* = \frac{1}{\delta} \hat{\Sigma}^{-1} \hat{\mu}. \quad (2.44)$$

Under the assumption of normality, this estimator is unbiased:

$$E(\bar{w}^*) = \frac{1}{\delta} E(\hat{\Sigma}^{-1}) E(\hat{\mu}) = \frac{1}{\delta} \Sigma^{-1} \mu, \quad (2.45)$$

where the first equality follows from the standard independence of $\hat{\mu}$ and $\hat{\Sigma}$, and the second equality is due to the unbiasedness of $\hat{\mu}$ and $\hat{\Sigma}^{-1}$. Consider a single risky asset. Expanding $\hat{w}^* = (1/\delta)\hat{\mu}/\hat{\sigma}^2$ around both μ and σ^2 yields:

$$\hat{w}^* = \frac{1}{\delta} \frac{1}{\sigma^2} (\mu - \hat{\mu}) - \frac{1}{\delta} \frac{\mu}{\sigma^4} (\sigma^2 - \hat{\sigma}^2). \quad (2.46)$$

Take variance and rearrange:

$$\sigma^2(\hat{w}^*) = \frac{1}{\delta^2} \left(\frac{\mu}{\sigma^2} \right)^2 \left(\frac{\sigma^2(\hat{\mu})}{\mu^2} + \frac{\sigma^2(\hat{\sigma}^2)}{\sigma^4} \right). \quad (2.47)$$

This expression shows that imprecision of the plug-in estimator depends on both: risk premium and variance estimates, each scaled by their respective magnitudes. Suppose we have same input estimates as earlier: ten years of monthly data on a stock that have the same mean and standard deviation i.e. $\hat{\mu}=1\%$ and $\hat{\sigma}_\mu=4\%$. Given these estimates, we get standard error of the plug in estimator \hat{w}^* for a risk aversion $\delta = 5$

equal to 14%. This example illustrates that portfolio weights tend to be very poorly estimated because the inputs to estimator are difficult to specify precisely. (Brandt, 2004).

Fortunately all hope is not lost, because there is several ways to improve poor return estimates. One popular approach is the econometrical one, so called the plug-in estimation. In plug-in estimation we try to econometrically specify the parameters of the data generating process $f(y_t|y_{t-1})$ and then plug these parameter values into the analytical solution to the investor's optimization problem. As long as we treat parameters as estimates, resulting portfolio weights are estimations too and they inherit estimation error (Brandt, 2004). Brandt (2004) suggests three different ways to improve the plug-in estimations, which are discussed later more closely:

1. Imposing constraints
2. Factor models
3. Shrinkage estimation

2.2 Portfolio resampling

2.2.1 Measuring estimation error

As discussed above, inputs into the efficient portfolio algorithm are measured without error, and the optimizer tends to pick those assets with attractive features (high return and low risk and/or correlation with other assets) and tends to short or deselect those with the worst features. These are exactly the cases where estimation error is likely to be highest, hence maximizing the impact of estimation error on portfolio weights. An algorithm that takes point estimates as inputs and treats them as if they were known with certainty (which they are not) will react to small differences in returns that are well within measurement error. This is the case in Markowitz optimization and the problem gets worse as the number of assets rises because this increases the chance of outliers. Typically the estimated parameters used in asset allocation problems i.e. means, variances and covariances are point estimates and they are calculated using just one possible realization of a return history. Even if we assume stationarity i.e. constant mean, non-time-dependent covariances, we can only expect point estimates of the risk and return inputs to equal the true distribution parameters if our sample is very large.

Estimation error is the difference between estimated and true distribution parameters when samples are not sufficiently large. Portfolio resampling⁷ allows us to clearly distinguish the impact of the uncertainty due to estimation errors in means and variances by a Monte Carlo procedure. (Scherer 2002, 79)

Lets suppose that we have true estimates (estimated with error) of variance-covariance matrix $\bar{\Sigma}_0$ and mean vector $\bar{\mu}_0$ of excess returns of n assets using T observations. By portfolio resampling we are able to model the impact of estimation error in input estimates on efficient portfolios. In the portfolio resampling process we basically draw return data repeatedly for each of the n assets for m times from the return distribution $N(\bar{\mu}_0, \bar{\Sigma}_0)$ (which is assumed to be true), where $m=T$. We can use this newly created block of data ($m \times n$) matrix of asset returns to construct a new, but statistically equivalent mean vector and covariance matrix $\hat{\mu}_i$ and $\hat{\Sigma}_i$. Obviously the original and the resampled matrices will differ due to sampling error. The degree of confidence will depend on m . If m is small, then our estimates will fluctuate largely, while we will find much less difference for large m . By repeating this procedure n times we get n new sets of optimization inputs from $\hat{\Sigma}_1, \hat{\mu}_1$ to $\hat{\Sigma}_n, \hat{\mu}_n$. (Scherer, 2004)

How can we use these artificially-generated estimates in measuring the estimation error? As we have earlier shown that sharpe portfolio is the most important portfolio on the efficient set (while minimum variance trivially doesn't include any information on expected returns), we measure the distance between the centre of the weight distribution and the maximum Sharpe ratio portfolio (which is constructed without taking estimation error into account) as vector distance

$$\left(\bar{w} - w_{sharpe}^*\right)' \left(\bar{w} - w_{sharpe}^*\right), \text{ where } \bar{w} = \frac{1}{n} \sum_{i=1}^n w_i. \quad (2.48)$$

The distance between the centre of the resampled distribution and the maximum Sharpe ratio portfolio converges to zero with an increasing number of simulation runs. This implies a reduction in estimation error when number of simulation increases (figure 2). (Scherer, 2004)

⁷ This procedure is patented under title "Portfolio Optimization by Means of Resampled Efficient Frontiers" by Richard and Robert Michaud in December 1999.

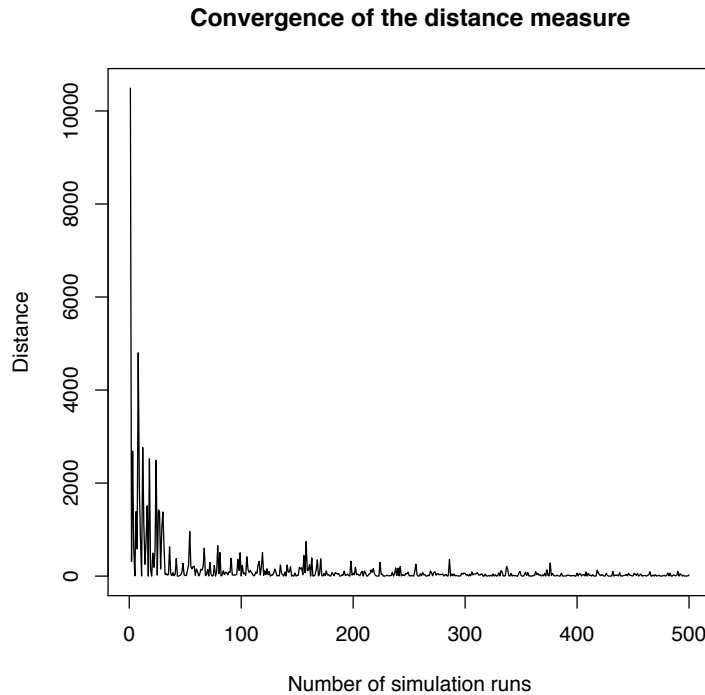


Figure 2. Convergence of the distance measure when short sales are allowed ($T=30$).

2.2.2 Resampled efficient frontier

The same effect can be shown by plotting the (assumed) true efficient frontier and resampled frontiers in the traditional risk-return space. For doing this, we calculate m portfolios along the frontier and save the corresponding allocation vectors w_{11}, \dots, w_{1m} to w_{n1}, \dots, w_{nm} . Evaluating all m frontier portfolios for each of the n runs, with the original optimization inputs $\bar{\Sigma}_0$ and $\bar{\mu}_0$, which will force all portfolios to plot below the original efficient frontier. This is because any weight vector optimal for $\hat{\Sigma}_i$ and $\hat{\mu}_i$ ($i=1, \dots, n$) cannot be optimal for $\bar{\Sigma}_0$ and $\bar{\mu}_0$. Therefore all portfolio weights result in portfolios plotting below the efficient frontier as the weights have been derived from data that *contain estimation error*. Hence, the result of the resampling procedure is that estimation error in the inputs is transformed into uncertainty about the optimal allocation vector. Increasing the number of draws, T , concentrates the data points closer to the original frontier as the dispersion of the inputs becomes smaller. This is equivalent to reducing sampling error. However, it is not clear from this resampling procedure where the “better” frontier lies (Scherer 2002, 81). Figure 3. illustrates how variation in resampled frontiers increases when the number of observations is decreased.

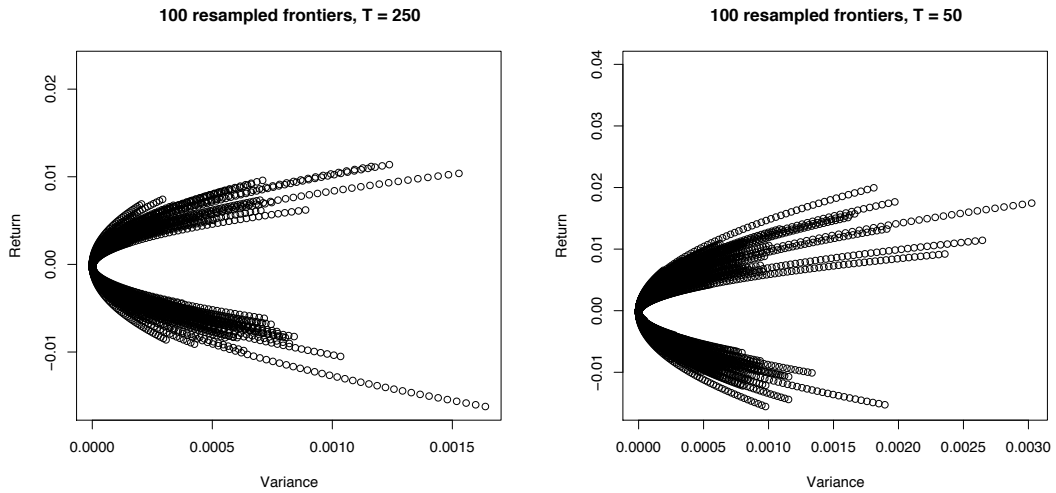


Figure 3. Estimation error visualized: the dispersion of efficient frontiers

All resampled data is derived from the same return vector and covariance matrix $\bar{\mu}_0$ and $\bar{\Sigma}_0$. However the true distribution is unknown. Resampling happens in the hope that $\bar{\mu}_0$ and $\bar{\Sigma}_0$ are reasonable close to μ_{true} and Σ_{true} . If this is not true, then estimation error in $\bar{\mu}_0$ and $\bar{\Sigma}_0$ will be passed on to $\hat{\mu}_1, \hat{\Sigma}_1, \hat{\mu}_2, \hat{\Sigma}_2 \dots \hat{\mu}_m, \hat{\Sigma}_m$ (see Figure 4.) (Scherer 2004)

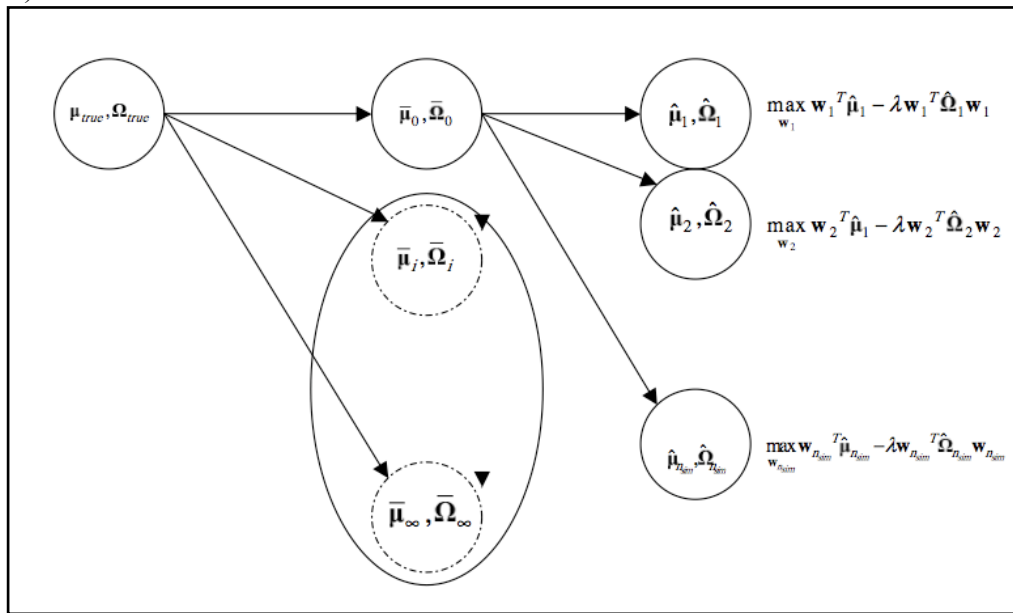


Figure 4. Error inheritance (Scherer 2004). Omega stands for the covariance matrix.

Resampled efficient frontier is defined as an “average of the rank-associated mean-variance-efficient portfolios”. This means that we average means over portfolios that have been given the same name (i.e. rank 1,2,...,m). Portfolios carrying rank 1 are the respective minimum-variance portfolios, while portfolios carrying rank m are the maximum-return portfolios. All other portfolios are ranked in between according to

their expected returns. The distance between the minimum-variance and maximum-return portfolios is divided into equal parts. This can also be done by adding up portfolios that show the same risk-return trade-off i.e. maximizing utility:

$$w^* = (\delta \Sigma)^{-1} \mu \quad (2.49)$$

for varying δ_m and then averaging the δ_m -associated portfolios. According to Scherer (2002, 87), the results of these alternative ways are very close to each other. Averaging maintains an important portfolio characteristic – that the weights sum to one – which is probably the main practical justification for the averaging procedure. However, the method is heuristic: it has no economic justification based on the optimizing behavior of rational agents. The resampled weight for a portfolio of rank m (portfolio number m along the frontier) is given by

$$\bar{w}_m^{resampled} = \frac{1}{n} \sum_{i=1}^n w_{im} \quad (2.50)$$

where w_{im} denotes the $k \times 1$ vector of the m th portfolio along the frontier for the i th resampling. Scherer (2002, 86) summarizes the procedure as follows:

1. Estimate the variance-covariance matrix and the mean vector of the historical inputs.
2. Resample, using the inputs created in step 1. taking T draws from the input distribution; the number draws, T , reflects the degree of uncertainty in the inputs. Calculate a new variance-covariance matrix from the sampled series. Estimation error will result in different matrices from those in step 1.
3. Calculate an efficient frontier for the inputs derived in step 2. Record the optimal portfolio weights for m equally distributed return points along the frontier.
4. Repeat steps 2 and 3 many times⁸. Calculate average portfolio weights for each return point. Evaluate a frontier of averaged portfolios with the variance-covariance matrix from step 1 to plot the resampled frontier.

⁸ As the number of sampling grows, statistical tests can be applied with greater confidence. For many applications 500 samplings will suffice (Scherer, 2002).

To measure the estimation errors in means only, we optimize using resampled means $\hat{\mu}_i, i=1, \dots, n$, and the original, truth variance-covariance matrix Σ_0 . Alternatively, we can reoptimize using resampled covariance matrices but treating means as known (Scherer 2002, 82). According to illustration of Scherer (2002, 82-83), the dispersion of risk-return points in mean-variance space is considerably reduced when estimation error is confined to variances.

So far, we have assumed that only source of estimation error is a sampling error, which is caused by insufficient data. If this would be the case, the problem would be temporary and the problem would vanish after 200 years. There is a second source of estimation error called “non-stationarity”. A time series is said to be non-stationary if its variance changes over time, its autocovariance is time- and also lag-dependent, or its mean changes over time. When there is non-stationarity the research might be well advised to use shorter data sets. Extending the length of data series might reduce the contribution of sampling error to estimation error, but at the same time it could increase that of non-stationarity (Scherer 2002, 83).

2.2.3 Resampled portfolio efficiency

As discussed earlier in chapter 2.1, there is a vast literature on the different methodologies that try to reduce estimation error and achieve more intuitive and better behaving portfolios. Resampling approach of Michaud (1998) on optimal portfolios does produce better behaving optimal portfolios even though it has no economic reasoning. The basic idea is to resample return data from assets’ historical time series to generate more samples and then calculate optimal portfolio for every resample and average the resulting portfolio weights. However, Scherer (2004) argues that this method, without short selling constraint, will only add noise into the portfolio construction process. In the case of long only constraints, on the top of the noise we also get biases into optimal allocations, as resampling tends to overweight the more volatile assets. Despite these deficiencies, by portfolio resampling we get valuable information about weight distributions and still, the resulting portfolios are very intuitive (see also Rasmussen 2003, 201: Quasi-Random Monte Carlo Simulated Asset Allocation).

If we look again the distance measure (equation 2.48):

$$\left(\bar{w} - w_{sharp}^*\right)' \left(\bar{w} - w_{sharp}^*\right), \text{ where } \bar{w} = \frac{1}{n} \sum_{i=1}^n w_i,$$

we concluded that this means effectively that the centre of the weight distribution recovers the (original) maximum Sharpe ratio portfolio. We can put this alternatively

$$\bar{w} = w_{sharp}^* + noise. \quad (2.51)$$

Large positive or negative weights can occur in single simulation runs, but will be averaged out eventually, which can be seen in figure 2. It becomes apparent that repeatedly drawing average returns and subsequently averaging across optimally constructed portfolio weights, yields the same result as averaging across returns in the first place and then use the averaged returns for portfolio optimization. We can clearly see this from following:

$$\begin{aligned} \bar{w} &= \frac{1}{n} \sum_{i=1}^n w_i \\ &= \frac{1}{n} \sum_{i=1}^n \delta^{-1} \Sigma^{-1} \mu_i = \delta^{-1} \Sigma^{-1} \frac{1}{n} \sum_{i=1}^n \mu_i \\ &= \delta^{-1} \Sigma^{-1} \bar{\mu} \end{aligned} \quad (2.52)$$

Neither the average portfolio, nor its risk changed. However, this should increase investment risk when investor is uncertain about the inputs. Resampling is unable to catch this effect. (Scherer, 2004)

Scherer (2004) illustrates, that when short sales are not allowed, the deviation from estimation error free solution is much smaller as long only constraint reduces the opportunities to leverage on information. Simulations do not converge to a distance measure of zero (figure 7).

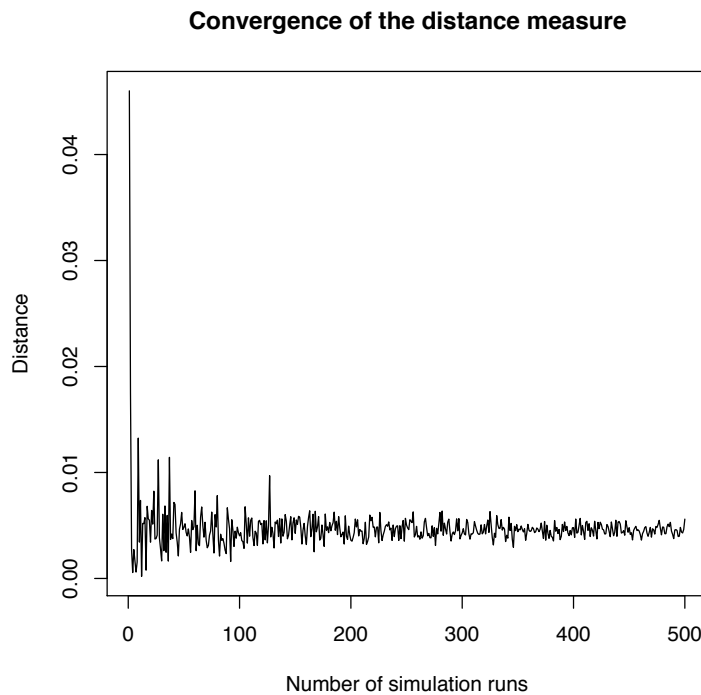


Figure 5. Convergence of the distance measure when short sales are not allowed (T=30).

Essentially this means that repeatedly sampling does not recover the Markowitz solution, as opposed to when short sales are allowed. Hence we get

$$\bar{w} = w_{sharp}^* + bias + noise. \quad (2.53)$$

The reason for this bias is that large portfolio weights are still the result of an upward biased asset mean (relative to the mean of other assets) within a single simulation run. However negative weights due to a downward biased asset mean in other simulation runs can no longer be implemented. This is the reason why long only constraint doesn't lead to Markowitz solution as individual assets are either in or out but never short. (Scherer 2004)

This actually explains why resampled portfolios are better diversified (see figures 6 and 7). Long only constrained resampled portfolios exclude negative weights and average all positive weights, which result higher than zero positive weights. Distance measure in figure 5. illustrates this also.

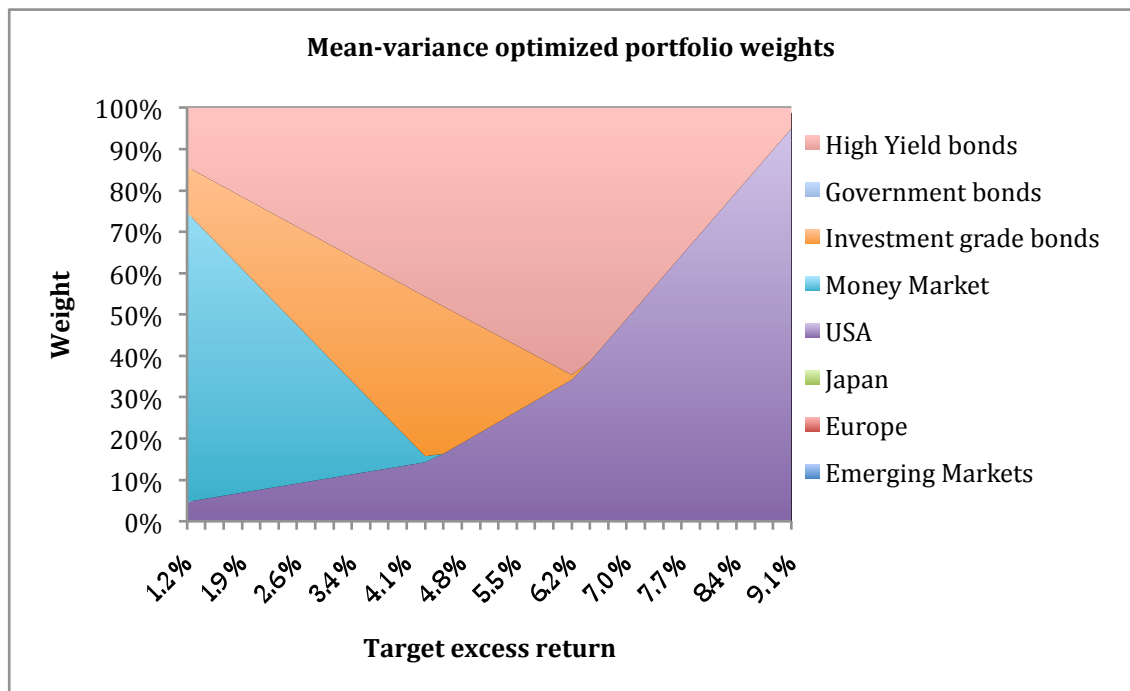


Figure 6. Portfolio weights along the efficient frontier (short sales not allowed)

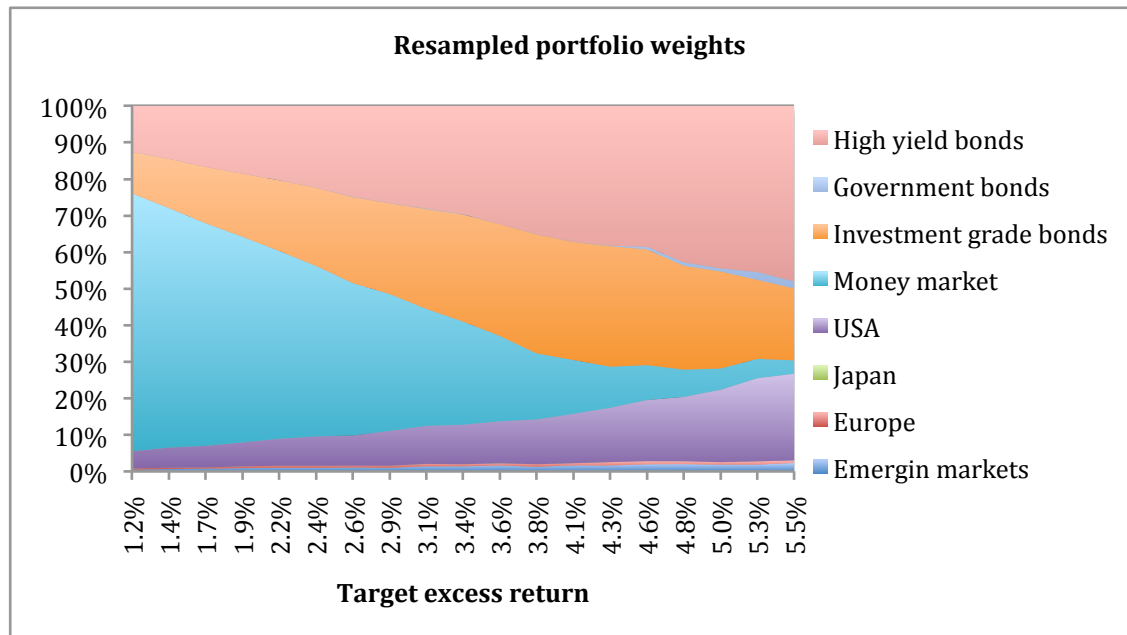


Figure 7. Portfolio weights along the efficient frontier (short sales not allowed, 500 simulation runs per portfolio)

How about volatility? Scherer (2004) demonstrates how volatility of the assets affects on resampled optimal portfolios by adding a lottery ticket with zero expected return, high volatility and zero correlation with other assets into a portfolio. By adding an asset with such characteristics should never push the efficient frontier outward and any sensible asset allocation model should never invest systematically in this kind of asset. In fact, as Scherer (2004) demonstrates, the traditionally derived Sharpe-ratio maximizing portfolio does not allocate to the lottery ticket. However, in resampled optimal portfolio, the lottery ticket does increase the distant measure because we can never short the lottery ticket and for some allocation runs result relatively high portfolio weight on the lottery ticket due to high volatility. This implies that it is the long only constraint that essentially transforms asset volatility into portfolio allocations. However, there are two effects at work when we have asset with unwanted characteristics: higher volatility induces an upward bias into the average resampled weight, but at the same time higher volatility makes the lottery ticket less attractive in the optimization process. The first effect comes predominant for maximum return portfolio and when increasing investors risk aversion, the second effect will become dominant. For reasonable high risk aversion, allocation in the lottery ticket will decrease due to higher risk. The volatility bias is still present, but at lower lever of risk aversion the risk effect dominates the upward bias. It should be clear by now that in portfolio resampling, the only parameter that address uncertainty is the number of draws per resampling which affects the drawing of returns for all assets in the same way. As the number of draws per resampling increases, *i.e.* confidence in estimated parameters increases, the allocation

into the lottery ticket -like assets decreases. It is the long only constraint that transforms asset volatility into asset allocation implicitly raising the return expectation for highly volatile assets. (Scherer 2004)

After all without long only constraint, resampled optimal portfolios will result Markowitz solution. Even though all inputs are measured with error, resampled efficiency will not pick that up. Asset risk remains unchanged even though the world became a much riskier in the presence of estimation error. In case of long only constraint, this changes dramatically. As assets can never be short, we will see that for some resampling runs the maximum return portfolio will be fully invested in cash. This leads to a sampling of cash into the maximum return portfolio. Another consequence is that we cannot derive portfolios that have high risk aversion parameters (without long only constraints we could have always shorted assets with a negative risk premium), which makes the similarity of rank and risk aversion based approaches questionable. Inclusion of cash into the maximum portfolio contrasts both with intuition and portfolio theory. (Scherer 2004). Despite these above-mentioned deficiencies, resampled optimal portfolios are better diversified even though the approach is totally heuristic.

2.3 Shrinkage estimation

Despite the seminal work of Markowitz (1952) about mean-variance optimization, we have shown that the model is pretty much useless when using historical mean returns. The application of Markowitz (1952) portfolio optimization traditionally proceeds in two steps. First, the moments of the return distribution are estimated from time-series of historical returns, and then the mean variance portfolio optimization is solved separately, as if the estimates were the true parameters (Jorion, 1986). The impact of parameter uncertainty on optimal portfolio selection as well as sensitivity of mean-variance-efficient portfolio to changes in asset means has been recognized by a number of authors e.g. Frankfurter et al. 1971 and Best & Grauer 1991. These studies have shown that practical application of portfolio analysis is seriously hampered by estimation error, especially in expected returns. Variances and covariances are also unknown, but are more stable over time (Merton, 1980). As we mentioned earlier, factor models and weight constraints can also be called for help. However, there is possibility to use historical returns without waiting for another couple of hundred years to have more accurate input estimates. We need to make the data “transformation” in order to reduce the sample error before we put Markowitz to work. This “transformation” is called shrinkage estimation and it is basically a statistical trick where we “shrink” sample estimates toward some grand value.

2.3.1 Historical returns

As discussed above, the plug-in estimates are extremely biased and they lead suboptimal portfolio choices in Markowitz portfolio optimization. Stein (1955) has shown that the classical mean is inadmissible. Motivated by the poor finite-sample property of plug-in estimates, one popular methodology among above-mentioned options (i.e. constraints and factor models) is *shrinkage estimation*. Jorion (1986) presents a simple empirical Bayes estimator that should outperform the sample mean in the context of a portfolio. The idea of shrinkage estimation is attributed to James and Stein (1961), who came up with the following estimator:

$$\hat{\mu}_{bs} = \delta^* \mu_0 + (1 - \delta^*) \bar{\mu} \text{ for } 0 < \delta^* < 1. \quad (2.54)$$

The James-Stein estimator forces the sample mean $\bar{\mu}$ toward a common constant μ_0 , which is often chosen to be a grand mean of all variables. The estimator μ_{bs} thereby reduces the extreme estimation error that may occur in the cross-section of individual means, resulting lower overall variance of the estimators and more stable results in portfolio optimization (Brandt, 2004). According to Jorion (1986) and DeMiguel et al (2007), the optimal shrinkage factor δ^* can be derived as follows:

$$\delta^* = \frac{N + 2}{(N + 2) + T(\bar{\mu} - \mu_0)' \Sigma^{-1} (\bar{\mu} - \mu_0)} \quad (2.55)$$

where Σ^{-1} is an inverse covariance matrix, N is a number of means and T number of observations. As it can be seen from the equation, the optimal shrinkage factor increases when number of means N increases, decreases when the sample size T increases and decreases in the dispersion of the sample means $\bar{\mu}$ from the shrinkage target μ_0 . Jorion (1986) uses the minimum variance portfolio return as a grand mean μ_0 .

According to Brandt (2004), improvements from using shrinkage estimation to expected returns are considerable. According to his example, when shrinkage to plug-in estimates was used, the sharpe ratio was 63% better than when using pure historical estimates.

2.3.2 Covariance matrix

The standard statistical method to estimate covariance matrix is to gather a history of past stock returns and compute their sample covariance matrix. Unfortunately this creates problems as well as use of historical means of expected return in return estimation. Ledoit and Wolf (2003) states that when the number of stocks under

consideration is large, especially relative to the number of historical return observation available, the sample covariance matrix is estimated with a lot of error. They also state that given the well-documented flaws of the sample covariance matrix, nobody should be using it anymore because the enhanced alternative is available.

Covariance matrix shrinkage can replace the sample covariance matrix in any mean-variance optimization application. The problem of the historical covariance matrix is that estimated large positive covariance estimates tend to contain a lot of positive error and therefore need to be pulled downwards to compensate for the error. Similarly the negative errors need to be compensated by pulling them upwards. This is called shrinkage of the extremes towards the center and properly implemented, this would clearly ease the problem of the sample covariance matrix described above (Ledoit & Wolf, 2003).

Jorion (1986) comes up with following Bayes-Stein shrinkage method on covariance matrix:

$$\Sigma_{bs} = \Sigma \left(1 + \frac{1}{T + \hat{\tau}} \right) + \frac{\hat{\tau}}{T(T + 1 + \hat{\tau})} \frac{1_N 1_N'}{1_N' \Sigma^{-1} 1_N}, \quad (2.56)$$

where

$$\hat{\tau} = T \frac{\delta^*}{1 - \delta^*} \quad (2.57)$$

and

$$\delta^* = \frac{N + 2}{(N + 2) + T(\bar{\mu} - \mu_0)' \Sigma^{-1} (\bar{\mu} - \mu_0)} \quad (2.58)$$

The portfolio obtained by using $\hat{\mu}_{bs}$ and Σ_{bs} is

$$\hat{w}_{bs} = \frac{\hat{\Sigma}_{bs}^{-1} \hat{\mu}_{bs}}{1_N' \hat{\Sigma}_{bs}^{-1} \hat{\mu}_{bs}}, \quad (2.59)$$

which combines both shrinkage approaches and is hence called as Bayes-Stein-portfolio (DeMiguel et al. 2007 and Jorion 1986). However, Ledoit & Wolf (2003) states that Jorion's (1986) approach to covariance matrix shrinkage breaks down when $N \geq T$ because their loss function involves inverse of the covariance matrix. Ledoit and Wolf (2003) propose a loss function that does not depend on this inverse covariance matrix.

2.3.3 *Asset pricing models*

In standard finance theory, investors optimally allocate their investment funds to assets using a given stochastic model of asset returns. It follows that the optimal asset allocation depends on the choice of the model. Uncertainty about correct choice of a stochastic model has recently become a research topic of interest. The use of asset pricing model such as CAPM reduces the dimension of the estimation problem. Although the use of asset-pricing models helps us to obtain portfolios that are more intuitive and easier to implement, we face uncertainty regarding models' pricing ability because all models are rejected in some empirical tests. Bayesian approach, which combines the prior beliefs in models and the information in data, is currently a popular approach in empirical studies of model uncertainty (Wang, 2005). Zhenyu Wang (2005) shows how shrinkage estimation can be used to reduce model uncertainty. He shows that Bayesian analysis of model uncertainty shrinks both the predictive mean and variance of asset returns from unrestricted sample moments to the estimates restricted by the asset-pricing model. The shrinkage approach reveals that a Bayesian investor, facing uncertainty about an asset-pricing model, implicitly assigns a weight between the unrestricted estimate and the estimate by the asset-pricing model. The weight is the shrinkage factor. For a given prior distribution, the weight on the estimate restricted by the asset-pricing model is larger if history lends stronger credibility to the factor-based pricing model. For instance the weight on the asset-pricing model is large if a long history of stationary data is not available. Investor who take the empirical Bayes approach, in which they use the observed data to estimate the prior, choose the shrinkage factor w to be $1/2$, which assigns equal weights to the restricted and unrestricted estimates (Wang, 2005). Thus, shrinkage estimated returns are

$$r_{shrinkage} = wr_{hist} + (1 - w)r_{MODEL} \quad (2.60)$$

where w is the shrinkage factor, r_{hist} the mean of historical returns and r_{MODEL} the return estimate from the asset pricing model. (Wang, 2005)

3 ASSET ALLOCATION

3.1 Introduction to asset allocation decisions: SAA vs. TAA

Strategic asset allocation (SAA) decision is a long-term investment decision and it is closely related to decision concerning the benchmark choice. Strategic asset allocation is also widely believed to be the single most important decision in the investment process. The decision essentially involves determining an appropriate long-term allocation across asset classes according to long term expectations about future risk and return of assets or asset classes. Also the expected correlation structure between the assets needs to be considered. It involves taking advantage of the correlation structure of the investable universe in order to achieve the desired return/risk profile over the long term. Strategic asset allocation decision is not – as opposed to *the tactical asset allocation* (TAA) decision - attempting to determine the return that one can expect from an asset class in short term. Strategic asset allocation decision can also be described as a broad framework or roadmap for determining an appropriate long-term asset allocation policy (Rasmussen, 2003). It should also be noted that according to Brinson et al (1986), well over 90% of the variation in investment performance of US pension plans could be explained by differences in asset allocation decisions. For example market timing and stock selection together explain only 10% of the performance.

With a given expected returns, risk, correlation structure and risk aversion, an asset manager makes a decision about allocation for next, lets say 100 years. However, the active managers have the opportunity to rotate between asset classes over time, based on superior information regarding market movements. These shorter-term decisions represent the manager's tactical decisions, designed to outperform the benchmark over short term. The asset allocation that results from these decisions to temporarily deviate from the long-term asset allocation is the tactical asset allocation (Rasmussen, 2003).

According to Lee (2000), the first tactical asset allocation product was introduced by Wells Fargo after severe market decline 1973-1974. Their attempt was to shift assets between stocks and bonds according to their forecasts about risk premiums. Philips, Rogers and Capaldi (1996) described tactical asset allocation as follows: "*A TAA manager's investment objective is to obtain better-than-benchmark returns with (possibly) lower-than-benchmark volatility by forecasting the returns of two or more asset classes, and varying asset class exposure accordingly, in a systematic manner*". In other words, tactical asset allocation is shifting portfolio weights away from the equilibrium allocation (strategic allocation) as a short time bets in order to achieve higher expected return with a given risk.

Sharpe (1987) states that as long as relative risk tolerance and investment opportunity set i.e. expected returns, variances and covariance remain constant the investor will choose a one-period static strategic portfolio. However, investor's risk aversion may change when he gets older or his wealth changes. Besides return distributions may be time varying when the investor tries to predict expected returns and covariances, whereat the investor will choose intertemporal tactical asset allocation (Sharpe 1987).

3.2 Benchmark selection

Selecting a benchmark implies selecting a market for observing consensus expected returns and this has implications for the active asset management. In practice, asset managers are most often evaluated relative to some benchmark. Benchmark can be for instance one of the MSCI's⁹ equity indices or an appropriate combination of equity and bond indices. This of course doesn't exclude the establishment of a very long-term asset allocation, from which managers can deviate tactically over shorter periods of time. However, in many cases the strategic asset allocation is the benchmark. In other words, without any short term market view, portfolio will follow the strategic (also called *neutral*) allocation – the benchmark.

In CAPM framework, index weights are sometimes used as a benchmark allocation. The market portfolio, that in theory includes all-embracing assets, gives us the performance of the market as a whole, but can rarely be used as a benchmark in practice. An active manager will be asked to outperform a benchmark portfolio that cannot reasonably be called “the market”. That is why a market index weights are sometimes used as a benchmark allocation while this allocation doesn't equal true “market portfolio”. Because managers as well as funds usually specialize, the benchmark is chosen accordingly to reflect the playing field as good as possible (Grinold & Kahn, 1999).

3.2.1 Naive allocation: 1/n

Naive allocation could reasonable be used as a benchmark. DeMiguel et al. (2007) evaluate the out-of-sample performance of the sample-based mean-variance model, and its extensions designed to reduce estimation error, relative to the naive 1/n portfolio.

⁹ Morgan Stanley Capital Index

They evaluated fourteen common portfolio optimization models across seven empirical datasets and found that none is consistently better than the $1/n$ portfolio in terms of Sharpe ratio, certainty-equivalent return (CEQ)¹⁰ or turnover. In general, they found that unconstrained policies that try to incorporate estimation error perform much worse than any of the strategies that exclude short sales. They also perform much worse than naive strategy. Imposing constraints on the sample-based mean-variance portfolio strategy leads to only a modest improvement in Sharpe ratio and CEQ returns, but substantial reduction in turnover. The minimum-variance portfolio with short-sale constraint performed best in terms of Sharpe ratio of all strategies studied. However, even this strategy was statistically superior to $1/n$ strategy in only one of the seven empirical datasets. A CEQ return wasn't superior in any dataset and turnover was always higher than the turnover of $1/n$ strategy. DeMiguel et al. (2007) also showed that in order to sample-based mean-variance strategy would achieve a higher CEQ return than the $1/n$ strategy estimation-window length should be 3000 months for a portfolio with only 25 assets. This critical length of estimation-window is a function of the number of assets because room for estimation error increases when amount of parameters increases. Overall conclusion from simulation results of DeMiguel et al. (2007) is that optimization models are expected to outperform $1/n$ naive allocation only if the estimation period is very long, the ex-ante (true) Sharpe ratio of the mean-variance efficient portfolio is substantially higher than the $1/n$ portfolio's and the number of assets is small.

What could be the reason for $1/n$ to perform so well? DeMiguel et al. (2007) argues that "allocation mistakes" caused by using naive allocation can turn out to be smaller than the error caused by optimizing model with inputs that have been estimated with error. A second reason for good performance of $1/n$ allocation is that DeMiguel et al. were using it to allocate wealth across portfolios of stocks rather than individual stocks. In this way the loss from naive allocation as opposed to optimal diversification is much smaller because well-diversified portfolios have lower idiosyncratic volatility than individual assets.

The naive portfolio as a benchmark has several advantages because it doesn't need any parameter estimation, optimization and it doesn't rely on any assumption about return distributions. $1/n$ strategy is also very easy to apply to a large number of assets, in contrast to optimizing models, which typically require additional parameter estimations as number of assets increases. As DeMiguel et al. has shown, $1/n$ portfolio has performed very well against several optimal portfolio strategies. According to DeMiguel et al (2007) research, as long as a portfolio manager is supposed to create

¹⁰ $CEQ_k = \hat{\mu}_k - \frac{\gamma}{2} \hat{\sigma}_k^2$, where $\hat{\mu}_k$ and $\hat{\sigma}_k^2$ are the mean and variance of out-of-sample excess returns for strategy k , and γ is a risk aversion coefficient.

value over the benchmark, portfolio manager have arguable done well, when she has beaten the $1/n$ portfolio.

3.3 Tactical asset allocation: The Black-Litterman approach

“Quantitative asset allocation models have not played the important role they should in global portfolio management. A good part of the problem is that such models are difficult to use and tend to result in portfolio that are badly behaved ” – Black & Litterman (1992)

Fisher Black and Robert Litterman in 1992 introduced a sophisticated alternative to overcome the major problems in quantitative portfolio management: corner solutions due to estimation error, input sensitivity and difficulty to incorporate subjective views effectively into optimization algorithm. When no constraints are imposed, standard portfolio optimization models give large negative positions for many assets. However, when constraints are used to rule out the short positions, the model suggests the “corner” solutions with zero weights and significantly large positions for few assets. This “estimation-error maximizer” problem in Markowitz optimization is discussed in chapter 2.1. In addition, investors typically have views about absolute or relative returns in a few markets, while standard optimization model requires investor to provide expected returns for all assets. As it is widely known, the portfolio weights are very sensitive to return estimates and poor input parameters result in very badly behaved portfolios. It is very clear that standard model has no way to distinguish strongly held views. As a result, the generated optimal portfolio appears very seldom to expose the views that investor wants to express. The solution for these problems, Black & Litterman combine Markowitz’ mean-variance optimization and capital asset pricing model of Sharpe and Litner. (Black & Litterman 1992)

The basic starting point in Black-Litterman model is equilibrium risk premiums, which are the excess returns that equate the supply and demand for global assets and currencies. These equilibrium risk premiums result market-capitalization-weighted portfolio, which provides neutral reference point for expected returns. Unlike expected returns estimated from historical data, these equilibrium returns generate better behaving optimal portfolio. The investor creates a new expected return vector by combining his views and confidence levels about the outlook for global equities and bonds with these equilibrium weights and tilts the asset weights in the direction of investor’s favor. However, the model does not assume that the world is always at CAPM equilibrium. The idea is that when expected returns move away from the

equilibrium returns, imbalances in markets will tend to push them back. According to Black & Litterman (1992), it is not reasonable to assume that expected returns are likely to deviate too far from the equilibrium returns. In fact, the expected returns used in BL optimization will deviate from equilibrium risk premiums in accordance with the investor's explicitly stated views and confident level. Actually this is the whole beauty behind the Black-Litterman model – the flexible way to combine equilibrium weights and investor's views with confidence levels, and get a portfolio that is possible to implement. The following figure illustrates the process behind the combined new return vector and the final tactical portfolio weights in the Black-Litterman approach:

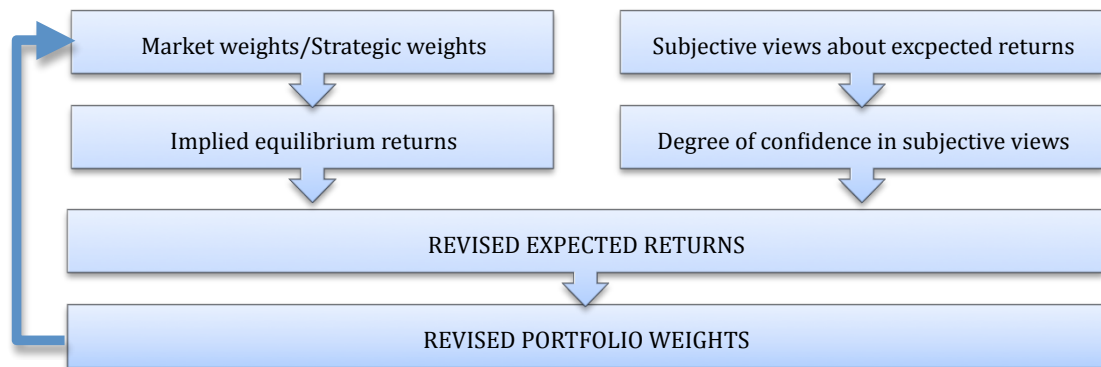


Figure 8. Intuition behind the Black-Litterman process (Drobetz, 2002)

3.3.1 Derivation of the Black-Litterman model: the Bayesian theorem

“The rational investor is a Bayesian” Harry Markowitz (1987, 57)

Previously, we have seen that information available doesn't suffice to get rid of parameter uncertainty on optimal portfolio choice. However, we don't only need more data, but it would also be irrational not to use other source of information e.g. the experience of insight (priors). The optimal combination of sample and non-sample information is found using Bayesian statistics. The main difference between traditional statistics and Bayesian statistics is that traditional approach creates point estimates for distribution parameters and treats them significant or insignificant. The Bayesian approach creates density function (posterior density) for parameters involved, given the observed data. It does so by combining sample information (likelihood functions) with

prior beliefs. Priors can be interpreted as the odds a research would be willing to accept if forced to bet on the true parameters before investigating the data. (Scherer 2002, 105)

Scherer (2004) distinguish different priors that can be used in input uncertainty modeling:

- *Uninformative priors* that only leverage the covariance matrix and leave expected returns unchanged. The efficient set remains unchanged
- *Statistical priors* that shrink return estimates towards the grand mean. The resulting portfolios converge towards the minimum variance portfolio if the available data history is short (e.g. James-Stein shrinkage)
- *Informative priors* that shrink return estimated towards equilibrium returns. Resulting allocations converge to the global market portfolio (e.g. The Black-Litterman model)

Suppose we have a history of risk premium for a single time series of a particular asset, and this history is summarized in a return vector $r = (r_1, r_2, r_3 \dots r_T)'$ where $r_i = R_i - r_f$. Suppose also that we are interested in estimates of mean and variance summarized in a parameter vector $\theta = (\mu \sigma^2)'$. Then the probability of obtaining the data (the return series) and the parameters can be written either as $p(r, \theta) = p(r|\theta)p(\theta)$ or, alternatively, as $p(r, \theta) = p(\theta|r)p(r)$. Equating both expressions we get Bayes' theorem:

$$\underbrace{p(\theta|r)}_{\text{Posterior density}} = \frac{\overbrace{p(r|\theta)}^{\text{Likelihood function}} \overbrace{p(\theta)}^{\text{prior}}}{p(r)}. \quad (3.1)$$

The posterior distribution $p(\theta|r)$ of the parameters describes our information about θ after we have observed the data, given our preknowledge captured via the likelihood function $p(r|\theta)$, which is the estimated probability of observing the data if the true parameter was θ . (Scherer 2002, 106)

This means that in portfolio optimization, the true parameters of the return distribution are still unknown, but investor replaces these parameters by subjective distribution given the historical data and subjective beliefs formed prior observing the data.

The Black-Litterman model is based on the Bayesian methodology, which effectively updates current market views with historical data to form new opinions. As discussed above, BL-model assumes that there is equilibrium between demand and

supply of asset and so there is also an equilibrium vector of excess returns implied by this market equilibrium. After having this equilibrium vector, the most essential problem arises: how do we combine this market equilibrium vector with portfolio manager's personal market views? Using the Bayesian rule, we get

$$p[E(r)|\Pi] = \frac{p[\Pi|E(r)]p[E(r)]}{p[\Pi]} \quad (3.11)$$

where

- $p[E(r)|\Pi]$ is the posterior return distribution that combines equilibrium information of market returns with forecasts by manager
- $p[\Pi|E(r)]$ is the conditional probability distribution of the equilibrium returns (data), given the forecasts of portfolio manager
- $p[E(r)]$ is the prior probability distribution that expresses the subjective (prior) views of the portfolio manager
- $p[\Pi]$ represents the marginal probability distribution of equilibrium returns

However, according to Satchell & Scowcroft (2000), Black and Litterman (1992) make following assumptions according to prior probability distribution (forecasts):

The manager has a set of k forecasts represented as linear relationships in the following way:

$$\gamma = PE(r), \quad \gamma \sim N(Q, \Omega) \quad (3.12)$$

where

- $\gamma = (k \times 1)$ vector of subjective forecasts of investor as linear relationships
- $P = (k \times n)$ matrix identifying and matching the assets involved in the forecasts
- $Q = (k \times 1)$ vector denoting the prior forecasts
- $\Omega = (k \times k)$ diagonal matrix with diagonal elements ω_{ii} representing uncertainty in each forecast γ

Concerning the return distribution of the conditional equilibrium returns, the following assumption can be made (Satchell & Scowcroft 2000):

$$P[\Pi|E(r)] \sim N(E(r), \tau\Sigma) \tag{3.13}$$

where

- $\Sigma = (n \times n)$ non-singular covariance matrix of excess returns on assets
- τ is a known scaling factor

This assumption means that the equilibrium excess returns conditional upon the investors' forecasts equals the investors forecast on average. According to Satchell and Scowcroft (2000) this may not hold in practice. Using above mentioned assumptions, Satchell and Scowcroft (2000) conclude the characteristics of the posterior return distribution as follows:

$$P[E(r)|\Pi] \sim N\left(\left[(\tau\Sigma)^{-1} + P'\Omega^{-1}P\right]^{-1}\left[(\tau\Sigma)^{-1}\Pi + P'\Omega^{-1}Q\right], \left[(\tau\Sigma)^{-1} + P'\Omega^{-1}P\right]^{-1}\right) \tag{3.14}$$

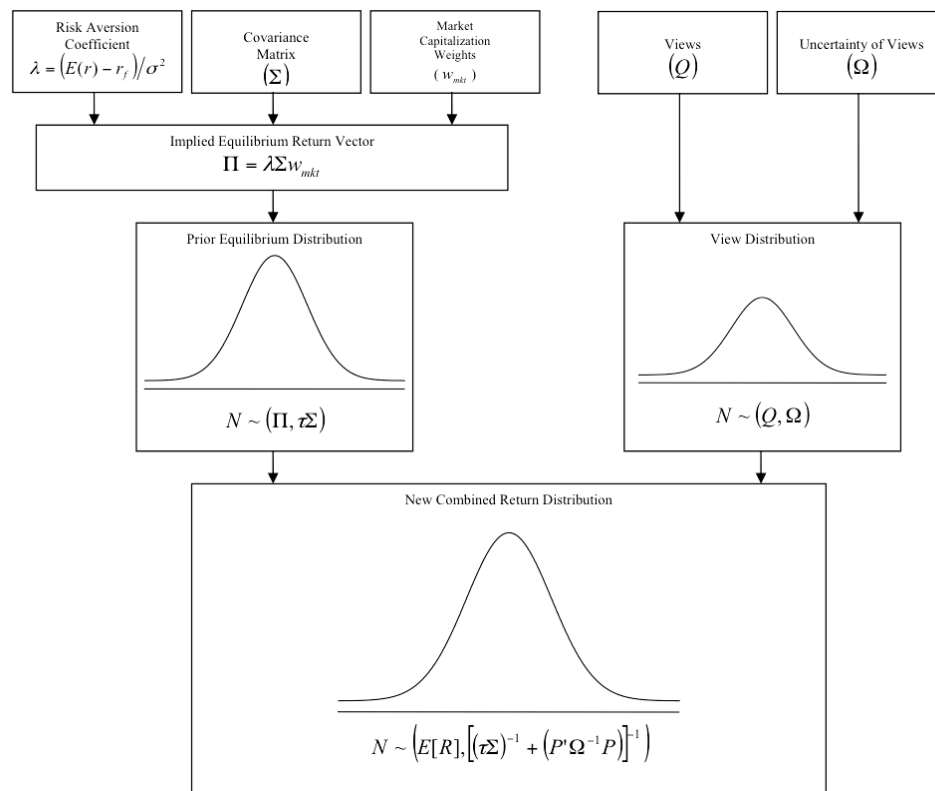


Figure 9. The process behind the new combined return vector (Idzorek 2004)

This mean representation of the posterior distribution is the so called master formula of the Black-Litterman model which gives mixed mean vector that is supposed to be used in portfolio optimization:

$$E(r) = \left[(\tau \Sigma)^{-1} + P' \Omega^{-1} P \right]^{-1} \left[(\tau \Sigma)^{-1} \Pi + P' \Omega^{-1} Q \right] \quad (3.15)$$

An alternative and maybe a more intuitive representation of the master formula is (Mankert, 2006):

$$E(r) = \Pi + \tau \Sigma P' (\Omega + \tau P' \Sigma^{-1} P)^{-1} (Q - P \Pi), \quad (3.16)$$

where

- τ is a scalar
- Σ is the covariance matrix of excess returns ($n \times n$ matrix)
- P is a matrix that identifies the assets involving the views ($k \times n$ matrix)
- Ω is a diagonal covariance matrix of error terms from the expressed views representing the uncertainty in each view ($k \times k$ matrix)
- Π is the implied equilibrium return vector ($n \times 1$ matrix)
- Q is the view vector ($k \times 1$ column vector)

This representation clearly shows that when the investor's view is the same as market (neutral) view, i.e. $(Q - P\Pi)$ equals zero, then our expected return equals the equilibrium return vector Π . In addition, when using quadratic utility maximization function

$$w^* = (\delta \Sigma)^{-1} E(r), \quad (3.17)$$

we get straight optimal non-constrained portfolio weights by

$$w^* = w^* + P' \left(\frac{\Omega}{\tau} + P \Sigma^{-1} P' \right)^{-1} \left(\frac{Q}{\delta} - P \Sigma w^* \right). \text{(Mankert, 2006)} \quad (3.18)$$

Conceptually, the BL model a weighted average of the implied equilibrium return vector Π and the view vector Q , in which the relative weightings are a function of the scalar τ and the uncertainty of the views Ω . However, the scalar and the uncertainty of the views are the most abstract and difficult parameters to specify in the model. The stronger the confidence of the views, the closer the new return vector will be to the views and relatively weak views will result in return vector that is closer to equilibrium returns. The scalar τ is more or less inversely proportional to the relative weight given to the implied equilibrium return vector Π .

3.3.2 Reverse optimization and the equilibrium returns

Black and Litterman (1992) explored three alternative ways to estimate expected returns to be used in their model: historical means, equal mean returns for all assets and risk adjusted returns. They demonstrated that all these alternative approaches lead to extreme portfolios. Therefore Black and Litterman came up with a different solution called reverse optimization, which is used in classical Black-Litterman model (Black & Litterman 1992).

For a starting point, the neutral reference point is needed. Black & Litterman (1992) argue that only sensible definition of “neutral” means is the set of expected returns that would clear the market if all investors had identical views. The neutral equilibrium vector is based on CAPM, which states that there is a common market portfolio that every rational investor holds. Algebraically, assuming the validity of the CAPM, the equilibrium vector is described as follows:

$$\Pi = \beta(r_m - r_f), \quad \beta = \frac{Cov(r, r'w_{mkt})}{\sigma_{mkt}^2}, \quad (3.19)$$

where Π is the $(n \times 1)$ market equilibrium return vector, r is the $(n \times 1)$ vector of asset returns, w_{mkt} is weights on the global market portfolio determined by market values and σ_{mkt}^2 is the variance of the global market portfolio return. If let Σ be the covariance matrix of the n asset space, then again maximizing quadratic utility function we can solve the *implied market neutral expected excess returns*:

$$\Pi = \delta \Sigma w_{mkt} \quad (3.20)$$

The risk aversion coefficient δ acts as a scaling factor for the excess returns estimates in the reverse optimization process. The bigger the excess return required for the unit of risk, the larger the estimated expected return. (Satchell & Scowcroft, 2000 ; Idzorek, 2004)

Idzorek (2004) points out a very important detail about excess returns used in the Black-Litterman model. Literature on the BL model can be confusing, since it often refers to the reverse-optimized implied equilibrium return vector Π as the CAPM returns. The regression based betas in CAPM returns can significantly differ from CAPM returns based on implied betas. Grinold and Kahn (1999) represent the following equation to derive implied betas:

$$\beta = \frac{\Sigma w_{mkt}}{w'_{mkt} \Sigma w_{mkt}} = \frac{\Sigma w_{mkt}}{\sigma^2}, \quad (3.21)$$

where β is the vector of implied betas, Σ is the covariance matrix, w_{mkt} is the market capitalization weights and

$$\sigma^2 = w'_{mkt} \Sigma w_{mkt} = 1/(\beta' \Sigma^{-1} \beta) \quad (3.22)$$

is the variance of the market (or benchmark) excess returns.

As the market portfolio is often conveniently proxied by a capitalization-weighted index of all publicly tracked assets, the implied returns would be the returns investors would need to hold the market portfolio. As capital market theory predicts that all investors hold the market portfolio, we could interpret these as equilibrium returns. This method can also be used to show investors (i.e. portfolio manager) whether their return expectations are consistent with market realities. When the benchmark portfolio has derived more or less using hand waving, it is reasonable check what kind of expected returns on assets the benchmark allocation implicates.

When constructing global tactical portfolio using Black-Litterman model, the portfolio manager may face a problem when trying to derive equilibrium returns following the basic literature on Black-Litterman model. The equilibrium returns are very straightforward to derive using reverse optimization when the asset space is some specific equity index, say MSCI World. However, when the asset space is equities and bonds world wide, it is not that simple to derive the market weights needed in reverse optimization. Of course portfolio manager can tackle this problem using some other method for calculating expected returns, for example shrinkage estimation introduced in chapter 2.5. However, the equilibrium weights can be seen as strategic weights chosen by portfolio manager and therefore benchmark weights could be used as market weights when calculating equilibrium returns in reverse optimization process.

3.3.3 *Expressing views: the view vector and the matching matrix*

The BL model uses a Bayesian approach to combine the investor's subjective views with the market equilibrium vector of expected returns. One of the brilliant characteristics of the model is that investor can have as many views as he wants or not at all. As long as an investor stays neutral, the equilibrium weights provide a neutral allocation. In addition, it is unrealistic to assume that investor is able to express exact expected returns on his views. Hence, when some views are stronger than other, an investor is able to express the difference.

The views can be relative or absolute. For example investor may feel that asset A will beat asset B by certain percent, or view can be such as market C will have specific absolute excess return with certain confidence level. Idzorek (2004) illustrates different views by following example.

Suppose we have eight asset classes in total:

1. US Bonds
2. International Bonds
3. US Large Growth
4. US Large Value
5. US Small Growth
6. US Small Value
7. International Developed Equity
8. International Emerging Equity

We have also following views on the certain markets:

View 1: International Developed Equity will have an absolute excess return of 5.25% with 25% confidence.

View 2 : International Bonds will outperform US Bonds by 25 basis points with 50% confidence.

View 3: US Large Growth and US Small Growth will outperform US Large Value and US Small Value by 2% with 65% confidence.

View 1. represents the absolute view, where excess return of 5.25% differ from equilibrium return of International Developed Equity. View 2. and 3. represent relative views. Relative views approximate the investor's feel about asset performance.

View 2. says that International Bonds will outperform the US Bond return by 0.25%, which does not tell us in which way the weights will be tilted to. We need to consider these returns relative to implied equilibrium returns of these markets. If implied equilibrium return for International Bonds is 0.67% and for US Bonds 0.08%, the implied difference is thus 0.59%. Since the view given to this difference is 0.25% i.e. lower than the implied difference, the model will tilt the portfolio away from better performing International Bonds weight in favor of US Bonds. In general, if the view is less than the difference in terms of implied equilibrium returns, the model tilts the portfolio toward the underperforming asset and vice versa.

View 3. illustrates view with multiple assets. The number of outperforming assets doesn't need to match the number of assets underperforming. However, the results of view involving multiple assets may be less intuitive. In this case, the assets of the view

form two separate mini-portfolios – a long portfolio and a short portfolio. Idzorek points out a very important detail: the relative weighting of each nominally outperforming asset is proportional to market capitalization weight of the asset divided by the sum of the market capitalization of the nominally outperforming assets of the particular view. This is naturally the case in underperforming assets. The net long positions equal the net short position. If the view is greater than the weighted average implied equilibrium return differential, the model will overweight the outperforming assets. The mini-portfolio that actually has the positive view may not be the nominally outperforming asset(s) from the expressed view.

Lets look at example view 3. Suppose that the weighted average of implied equilibrium return of the mini-portfolio formed from US Large Growth and US Small Growth is 6.25% and the weighted average of implied equilibrium return of another mini-portfolio formed from US Large Value and US Small Value is 4.04%. In this case the weighted average implied equilibrium return differential is 2.48%. Now, because the investor's view of the performance differential is 2% in favor of US Large Growth and US Small Growth relative to 2.48% measured in implied equilibrium return, it in fact represents reduction in the performance of US Large Growth and US Small Growth relative to US Large Value and US Small Value. (Idzorek, 2004)

The view vector Q needs to match with specific assets in matrix P . According to the three-view example presented earlier, lets assume that we have eight-asset case. In general, recall the prior return distribution. Hence, the investor's views can be represented as follows:

$$P \times E(r) = Q + \varepsilon, \quad \varepsilon \sim N(0, \Omega), \quad (3.23)$$

where uncertainty of the views is random, unknown, independent, normally-distributed error term vector ε with above mentioned return distribution.

In our three view example, it can be illustrate as follows:

$$Q + \varepsilon = \begin{bmatrix} 5,25 \\ 0,25 \\ 2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix} \quad (3.33)$$

If the investor is 100% sure about his view on the asset i , then error term ε_i equals zero. Otherwise error the term ε_i has either positive or negative value. The error term does not directly enter the Black-Litterman formula, but the *variance*, which is the absolute difference from the error term's expected value 0, does enter the formula. The variances of the error terms ω represent the uncertainty of the view. The omega matrix is discussed further in next subsection.

The column vector Q has expressed views and they are matched to specific assets by matrix P . Each individual view results in a $(1 \times n)$ row vector. Thus, k views result in a $(k \times n)$ matrix. Thus, P is a (3×8) matrix:

$$P = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & .5 & -.5 & .5 & -.5 & 0 & 0 \end{bmatrix} \quad (3.34)$$

Matrix P is just a matching matrix and it illustrates whether the view is absolute or relative. It doesn't include any more information about confidence or views. The first row of the matrix P represents view 1., which is an absolute view on asset 7. (International Developed Equity). View 2. and view 3. are placed on their own rows; on row 2 and row 3 respectively. In the case of relative views, each row sums to zero. If the assets are nominally outperforming, they receive positive weighting, while nominally underperforming assets receive negative weightings.

However, there are several ways to specify the values in P matrix. Satchell and Scowcroft (2002) use an equally weighted scheme, which is also used in row 3. In above mentioned example. In this case, weights are calculated by the number of respective assets outperforming or underperforming (2 outperforming assets). (Idzorek, 2004)

However, this weighting scheme ignores the market capitalization weights, which can cause big deviations in assets having small market weights. Other possibility is to use market weighting scheme. Idzorek (2004) states that he prefer to use market capitalization weights: The relative weighting of each individual asset is proportional to asset's market capitalization divided by the total market capitalization of either the outperforming or underperforming assets of that particular view. For example, if market capitalization weights of the outperforming US Large Growth and US Small Growth are 90% and 10% respectively, while weights for the relative underperforming assets US Large Value and US Small Value are -90% and -10% respectively. This approach would result in following P matrix:

$$P = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & .9 & -.9 & .1 & -.1 & 0 & 0 \end{bmatrix} \text{.(Idzorek, 2004)} \quad (3.35)$$

3.3.4 Incorporating uncertainty into the views and scalar τ

Once the P matrix is defined, the variance of each individual view can be calculated. According to Idzorek (2004), the variance of an individual view can be calculated by $p_k \Sigma p_k'$, where p_k is a single ($1 \times n$) row vector from matrix P that corresponds to the k th view and the Σ is the covariance matrix of the excess returns. The respective variance of each individual view portfolio is an important source of information regarding the (un)certainty of the confidence that should be given to a view. This information about variance is used shortly to revisit the variances of the error terms ω which form the diagonal elements of Ω :

$$\Omega = \begin{bmatrix} \omega_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \omega_k \end{bmatrix}. \quad (3.36)$$

The reason why off-diagonal positions are zero is that the model assumes that the views are independent of each other. In fact, determining this matrix Ω and its elements ω_i is one of the most complicated aspects of the model.

As mentioned earlier, the BL-model is a weighted average of the implied equilibrium return vector Π and the view vector Q . The relative weightings are a function of the scalar τ and the uncertainty of the views Ω . However, the scalar and the uncertainty in the views are the most abstract and challenging to specify in the model. The stronger the views the more closer the new return vector will be to the views. If the confidence on the views are low, the new return vector will be closer to the implied equilibrium return vector. The scalar τ is more or less inversely proportional to the relative weight given to the implied equilibrium return vector Π .

However, guidance in the literature for setting the value for the scalar is scarce. Satchell and Scowcroft (2000) say that the value of the scalar is often set to 1. On the other hand Black and Litterman (1992) and Lee (2000) state that because the uncertainty in the mean is less than the uncertainty in the return, the scalar τ is close to zero. One would expect the equilibrium returns to be less volatile than the historical returns. (Idzorek, 2004)

3.4 Risk factors in the returns on stocks and bonds

There has been a lot of research about economic factors that are able to explain returns on stocks and bonds. In this study, we will lean on the essential studies by Chen, Roll & Ross (1986), Elton, Gruber & Blake (1995), Campbell & Ammer (1993) and Ferson & Harvey (1991, 1993).

Asset prices are commonly believed to react sensitively on economic news. Asset prices are usually considered to respond external forces, as well as equity prices may have feedback on the other variables as well. It is apparent that all economic variables are endogenous in some sense. Only unexpected events such as earthquakes can be considered as truly exogenous variables to the world economy. By the diversification argument that is implicit in capital market theory, only general economic state variables will influence the pricing of large stock market aggregates. Any systematic variables that affect the economy's pricing operator or that influence dividends would also influence equity market returns. Also indirect variables that don't necessary have influence on current cash flows but do describe changing investment opportunity set are also considered as systematic risk factors (Chen;Roll;& Ross, 1986).

Stock prices (or any investment asset prices) can be written as expected discounted cash flows:

$$p = \frac{E(c)}{k} \tag{3.37}$$

where c is the cash flow stream and k is the discount rate. It follows trivially that the systematic forces that influence returns are those that change discount factors, k and expected cash flows $E(c)$. The discount rates changes with both level of rates and the term-structure spreads across different maturities. Unexpected changes in risk-free interest rate will therefore influence pricing and through their influence on the time value of the future cash flows, they will influence returns. The discount rate k , also depends on the risk premium, when unexpected changes in risk premium also influence returns and prices. On the demand side, changes in the indirect marginal utility of real wealth, perhaps as measured by real consumption changes, will influence pricing, and such effect should also show up as unanticipated changes in risk premium (Chen, Roll and Ross, 1986).

Expected cash flows change because of both real and nominal forces. Changes in the expected rate of inflation would influence nominal expected cash flows as well as the nominal rate of interest. In addition, changes in the expected level of real production would affect the current real value of cash flows. Insofar as the risk premium measure does not take in to consideration industrial production uncertainty, innovations in the rate of productive activity should have an influence on stock returns through their impact on cash flows (Chen, Roll and Ross, 1986).

The asset pricing models imply that the expected returns of securities are related to their sensitivity to changes in the state of the economy. Sensitivity is measured by the securities' "beta" coefficients. For each of the relevant state variables, there is a marketwide price of beta measured in the form of an increment to the expected return – a risk premium per unit of beta. In such model, the predictable variation of returns can

be driven by changes in the betas and changes in the price of beta. A lot of studies have identified state variables that are “priced”, in the sense that the risk premiums are different from zero on average (Ferson & Campbell 1991).

Campbell & Ammer (1993) studied what moves the stock and bond markets by decomposing variance of long-term asset returns. They tried to account for the variance of stock returns jointly with the variance of long-term nominal bond returns and the covariance between stock and bond returns. They expressed the innovation to a long-term asset return as the sum of revisions in expectations of future real cash payments to investors and revisions in expectations of future real returns on the assets. They combined the asset-pricing framework with vector autoregression (VAR) in long-term asset returns, interest rates, inflation and other information that helps to forecast these variables.

The data used in the Campbell and Ammer’s (1993) study was post war data. The VAR system included six variables: the excess stock return, real interest rate, change in the nominal 1-month interest rate and yield spread between 10-year yield and 1-month yield, dividend-price ratio and relative bill. The relative bill rate is the current 1-month bill rate less a backwards 1-year moving average of bill rates. The relative bill rate helps to capture some of the long-run dynamics of changes in interest rates without introducing long lags, and hence a large number of parameters, into the VAR system. The number of variables in the VAR increases very rapidly with lag length, so there is some danger of overfitting when high-order VAR is used (Campbell & Ammer, 1993).

Campbell & Ammer’s (1993) main findings were that economic forces do create persistent changes in expected excess stock returns and these changes are not associated with important changes in long-horizon forecast of real interest rates. Even though their VAR model had quite modest forecasting power on excess returns at a monthly frequency, ranging from 6% to 13% depending on sample period, the key point was that changes in expected excess returns are highly persistent, so that modest movements in short-run expected returns are capitalized into large changes in stock prices. The persistence of expected returns arises largely from the persistence of dividend-price ratio, one of the main forecasting variables for excess stock returns (Campbell & Ammer, 1993).

The real interest rate component of the excess stock return has smaller variance than the other components. This suggests that theoretical models of stock market pricing should not rely heavily on changing real interest rates. A similar point can be made for bond returns. Even though there are a lot of theoretical pricing models for real bonds in financial literature, Campbell and Ammer (1993) find that variance of excess returns on long-term nominal bonds is accounted for primarily by news about future inflation rates, which would have no impact if bonds had real payoffs. Although long-horizon forecasts of real interest rates are not highly variable, there are short-run changes in the

ex ante real interest rate. Campbell and Ammer found that news about the one-month-ahead forecasts of inflation and real interest rate is an important component of the innovation variance of the 1-month bill rate. Since the real interest rate does help to move the short-term interest rate, but has little impact on the long-term bond yield, they found that real interest rate news is also a major factor accounting for the variability of the yield spread between 10-years bonds and 1-month bills. Campbell & Ammer (1993) also find evidens that excess bond returns can be forecast using the same variables that help to forecast excess stock returns. This consistent with results of Fama and French (1989) and Elton et al. (1995)

Chen, Roll and Ross (1986) examined prediction power of following economic variables on expected stock returns: industrial production, inflation, yield spread, term spread and market indices. They found that monthly growth rate in industrial production, unanticipated inflation, unanticipated change in term structure and unanticipated risk premium were significant factors explaining expected returns. However, they found that yearly growth rate in industrial production and stock market indices (value weighted New York stock exchange index) were not significant at all. They also examined influences on prices of exposure to innovations on real per capita consumption and oil price index changes but none of these two were significant.

Elton, Gruber & Blake (1995) developed relative pricing (APT) models that are successful in explaining expected returns in the bond market indices. They examined commonly used stock return risk factors and they employed a combination of returns on a set of portfolios that have been shown to affect bond return, but also employed unexpected changes in some economic variables to see if they add any additional explanatory power. They tried to examine the ability of their model (APT) not only to explain the time series of returns, but also cross sectional differences in average returns. In this cross-sectional study they were able to determine, whether the hypothesized influences are priced by the market.

The variables they studied were market returns (the excess return on the stock market over the risk free rate), default risk (the difference in returns between corporate bonds and government bonds), yield spread (the difference in returns between long-term and short-term government bonds), unexpected changes¹¹ in inflation and unexpected chances in a measure of economic performance.

The first variable, the excess return (return over riskless rate) on the stock market, can be considered as a measure of expectations about general economic conditions. The second variable, a measure of default risk should affect corporate bond returns. The third variable measures the expectations on future interest rates and inflation. The fourth

¹¹ Unlike others who estimated unexpected changes in economic variables from time-series analysis, Elton et al. used publicly available forecasts to measure unexpected changes in forecasts

and fifth variables are measures of macro-economic influences. Changes in both of these variables should affect the level of interest rates and hence bond returns. They could also affect the certainty of cash flows on different types of bonds (e.g. corporate bonds and mortgage debt).

They also studied two additional variables. The first one is an index of aggregate bond returns. They pointed out that if one were looking for the best single factor that best explains individual bond returns, the best single variable for this job would be a market index of bond returns. This is analogous to a market index for stock returns being the single index that best explains the performance of individual stocks. The aggregate index should also be included because it serves as the benchmark against which to measure the importance of other indices. The second index, a measure of the return on mortgage securities relative to the return on government bonds. This index is added because options are an important element in bond returns (Elton, Gruber & Blake 1995).

Traditionally studies such as Chen, Roll and Ross (1986) have used changes in realizations of economic variable. In another context, Elton, Gruber & Blake (1995) have shown that expectations determine prices and that changes in expectations represent unexpected influences. Treating chances as unexpected is consistent with a rational expectations view of economic decision-making and is consistent with large body of empirical evidence¹². According to Elton, Gruber & Blake (1995), the use of survey data has the advantage of focusing directly on expectations rather than depending on an unspecified link between the measure used and expectations as justification for the measure. The data they used to measure unanticipated change in inflation was obtained by surveying consumers on inflation over the next year (Survey Research Center of the University of Michigan). For the other fundamental expectational variable – the general economic condition measuring factor they used unexpected change in the forecast of real Gross National Product (nominal GNP with inflation removed). The variable was derived from monthly data forecasts of nominal GNP growth rates provided by Eggert Enterprises. The data is on a survey of professional forecasters from financial intermediaries, companies and brokerage firms, as well as on output from major econometric models (Elton, Gruber & Blake 1995).

Term spread was rejected and a possible explanation for the term variable not having explanatory power is that its influence main in fact be captured by other variables. The aggregate bond index minus the risk free-rate is in part a measure of the term premium. All the other six variables were included in the best model and models that didn't include fundamental variables (unexpected changes in inflation and GNP) were rejected

¹² See Elton *et al.* (1981)

in 5% level in favor of models that do contain those variables. However, the aggregate bond index (the weighted average of the Lehman Brothers aggregate bond index and the Blume/Keim high-yield bond index) explains the most of the expected return. When other fundamental variables are added, the ability of a model to explain time-series behavior is relatively small. This should not be surprising because for example interest rate changes are the dominant cause of changes in returns on bond portfolios. Thus, any bond index will explain a large percentage of the time series of returns. However, the addition of the fundamental variables leads to improvement in the explaining expected returns. The percentages of different factors explaining expected returns were as follows: S&P 3%, default 2%, option 3%, aggregate bond index 73%, GNP 5% and inflation 14% (Elton, Gruber & Blake 1995)

Ferson & Harvey (1991) studied the behavior of economic risk premiums over time. They showed that a rational asset-pricing model that focuses on risk can explain most of the predictability of returns on stocks and bonds. They found that premium associated with stock market risk is the most important for purposes of capturing predictable variation of the common stock portfolios, while premiums associated with interest rate risk capture predictability of the bond portfolio returns. They also found that time variation in the expected risk premiums – not the betas – is the primary source of predictability at the portfolio level. They studied number of proxies for the economic risks that influence security returns. Ferson & Harvey (1991) argues that of course there is no claim that the variables uniquely capture the relevant economic risks, but they could jointly proxy for set of latent variables that determine security returns.

The variables they studied were market portfolio (value-weighted NYSE index less 1-month Treasury bill return), monthly real per capita growth of personal consumption expenditures for non-durable goods, default spread (BAA corporate bond yield less long term U.S. government bond return), change in yield spread (change in the difference between average monthly yield of a 10-year Treasury bond and 3-month Treasury bill), unexpected inflation¹³ (the difference between the actual and forecasted inflation rate, formed from a time-series model) and real interest rate (1-month Treasury bill return less the monthly rate of inflation). The CAPM provides some motivation for the state variable proxies, especially the use of “market portfolio” (Ferson & Harvey 1991).

Using multivariate model where all six variables were estimates at the same time, default spread, change in the yield spread and personal consumption growth had t-statistics over 2.0. The market portfolio had t-statistics 1.07 but as Chen et al. (1986) argue, other economic variables largely subsume the market premium. This can be seen also from Ferson & Harvey’s (1991) bivariate model testing for size and bond portfolios, where market portfolio got t-statistics 2.23. In bivariate model testing for

¹³ See Fama and Gibbons 1984

industry, size and bond portfolios, all the variables were statistically significant except market portfolio (t-statistics 1.61) and unexpected inflation (t-statistics -0.93).

They also studied a set of information variables in order to measure predictable variation in above mentioned risk premiums. The examined information variables that investors use to set prices in the market were lagged return of the equal-weighted NYSE index less 1-month Treasury bill rate, 1-month return of a 3-month Treasury bill less the 1-month return of a 1-month bill, yield spread between corporate Baa yield less Aaa corporate bond yield, monthly dividend yield on the SP500 stock index and nominal 1-month Treasury bill rate. Ferson & Harvey (1991) concluded that the risk premiums associated with a stock market index captures the largest component of the predictable variation in the stock returns. The premiums associated with term structure shifts and default spreads are the most important for the fixed-income securities.

The stocks that they studied were firms that were listed on the New York Stock Exchange (NYSE). The bonds were long-term government bonds, a long-term corporate bonds and Treasury bill that is closest to 6 months to maturity. The data were provided by the Center for Research in Security Prices (CRSP) at the University of Chicago. Ten common stock portfolios are formed according to size deciles on the basis of the market value of equity outstanding at the beginning of each year and 12 portfolios of NYSE firms grouped by two-digit standard industrial classification (SIC).

Ferson & Harvey (1993) studied international risk factors affecting national equity market returns. They studied risk factors including market portfolio, exchange rate fluctuations, global inflation, world interest rates, international default risk and world industrial production. They formulated an empirical beta-pricing model, where country-specific conditional betas measure sensitivity to the global risk factors. Both the betas and the expected risk premium were able to vary over time (60-month rolling regression). Their approach is unique in a way that they assumed national equity markets to be perfectly integrated in a global economy, with no barriers to extranational equity investments, no transactions of information costs or no taxes. As they assume globally integrated capital markets, it implies that the risk premium should not be country-specific. Such extreme assumptions are unlikely to provide a good approximation to the actual complexity of international investments. However, the results are encouraging. Their study proved that market portfolio is by far the most important factor. It alone explained 5 to 71 percent of the ex post variance, depending on the country (arithmetic return on the MSCI world equity total return index minus Ibbotson Associates one-month bill rate). Other variables were the log change in the G-10 foreign exchange rate (based on the trade-weighted dollar per foreign exchange rate of 10 industrialized countries), the change in the long-term expected G-7 inflation (result of projecting the four-year moving average of G-7 inflation on the set of lagged instrumental variables), oil price change and G-7 real interest rate (calculated by

aggregating individual countries' short term interest rates minus inflation using varying weights on quarterly shares in G-7 GDP). They found Eurodollar-Treasury yield spread, the unexpected inflation and industrial production to be insignificant.

3.5 Performance measures

3.5.1 Active return and active risk

In addition to absolute risk measure such as volatility of the portfolio return, active asset managers also require some measure of relative risk in order to demonstrate their performance relative to benchmark. Especially the clients would be interested to know that how much additional risk the manager has taken in order to achieve one additional return percent. In other words, we need some kind of measure of the difference between portfolios performance and the benchmark performance.

As managers are most often evaluated against a benchmark, the most common tool applied in order to quantify how the given portfolio tracks the chosen benchmark is so-called *tracking error* or *active risk*. Tracking error is defined as the volatility of the difference between the return on a portfolio and the return on the benchmark. This excess return on the benchmark performance, where the tracking error is calculated, is called portfolio's *active return* r_{PA} , which is defined as the difference between the return on the manager's portfolio r_p and the return of the benchmark portfolio r_B , i.e. $r_{PA} = r_p - r_B$. Active risk (tracking error) ψ_p , is expressed formally as follows:

$$\psi_p = Std\{r_{PA}\} = Std\{r_p - r_B\} = \sqrt{\frac{1}{n-1} \sum_{t=1}^T (r_{at} - \bar{r}_a)^2}, \quad (3.38)$$

where r_a is an active return at time t . (Grinold & Kahn 1999 and Rasmussen 2003).

3.5.2 Residual return and residual risk

Another useful measure of relative risk is called the *residual risk*, which is also known as *diversifiable*, *unsystematic* or *non-benchmark-related* risk. This is the portfolio risk orthogonal to the systematic risk of the portfolio. The concept can be illustrated by looking at the relationship between portfolio (excess) return and benchmark (excess) return:

$$r_p = \beta_p \cdot r_B + \theta_p. \quad (3.39)$$

Portfolio return can thus be divided into two components: a benchmark-related (systematic) component and a component that is independent (residual) of the benchmark return. The strength of the relation between the portfolio return and the benchmark return is determined by

$$\beta_p = \frac{\text{Cov}\{r_p, r_B\}}{\text{Var}\{r_B\}} \quad (3.40)$$

This is the factor that quantifies *systematic risk* (also called *un-diversifiable or benchmark-related risk*) relative to the benchmark. The first term in the equation (3.40) expresses the dependency of the portfolio return on the benchmark return. Note that this is a linear relationship, where any given level of beta, the return deviations from that linear relationship will show up as the residual θ_p . The term *residual* refers to the fact that in a linear relationship between portfolio returns and benchmark returns, the part of the portfolio returns that is unexplained by the benchmark (orthogonal) is called the *residual term*. The residual risk of the portfolio relative to the benchmark is thus a measure of variance of that *residual return*. The residual return is also called as *alpha*. Alpha is the return above (or below) that which would be expected by the relationship between the portfolio and benchmark return. The residual returns for portfolio P can also be separated as follows

$$\theta_p(t) = \alpha_p + \varepsilon_p(t) \quad (3.41)$$

where α_p is the average residual return and the $\varepsilon_p(t)$ is the mean zero random component of residual return. If $r_p(t)$ are the portfolio excess returns over those same periods, then equation (regression) is thus

$$r_p(t) = \alpha_p + \beta_p \cdot r_B(t) + \varepsilon_p(t) \quad (3.42)$$

The estimates of β_p and α_p obtained from the regression are the historical beta and alpha. Looking forward (ex ante), alpha is a forecast of residual return. Let θ_n be the residual return on stock n . We have

$$\alpha_n = E\{\theta_n\}. \quad (3.43)$$

Alpha has a portfolio property, since both residual returns and expectations have the portfolio property. Looking backward (ex post), alpha is the average of the realized residual returns. Realized alphas are for performance analysis and keeping score. Because the job of the manager is to score, we need good forecasts of alphas. The

benchmark portfolio will always have residual return equal to 0, i.e. $\theta_b = 0$. Therefore, the alpha of the benchmark portfolio must be 0.

The standard deviation of the residuals is the residual risk $Std[\theta_p] = \omega$. This means that the portfolio's risk can be expressed as

$$\sigma_p^2 = \beta_p^2 \cdot \sigma_B^2 + \omega_p^2 \quad (3.44)$$

where σ_p and σ_B are the volatilities of the portfolio and the benchmark respectively. Equation above states that portfolio volatility is a product of the benchmark-related risk and non-benchmark-related residual risk caused by the deviations of the portfolio from the benchmark. If the portfolio replicated the benchmark completely, the residual risk term will be equal to zero, and portfolio volatility will be equal to benchmark volatility. This residual volatility can be expressed as:

$$\omega_p = \sqrt{\sigma_p^2 - \beta_p^2 \cdot \sigma_B^2}. \quad (\text{Grinold \& Kahn 1999 and Rasmussen 2003}). \quad (3.45)$$

3.5.3 Information ratio

The *information ratio* (IR) of a portfolio measures the performance of the active manager by relating return above that of the benchmark to the risk above the benchmark. According to Rasmussen (2003), there seems to be some confusion in the industry as to the proper definition of the information ratio. He argues that the following definition is the most appropriate for assessing active manager performance relative to a benchmark. As we previously demonstrated, portfolio (excess) return can be described by

$$r_p = \beta_p \cdot r_B + \theta_p \quad (3.46)$$

where $E[\theta_p] = \alpha_p$ and $Std[\theta_p] = \omega$. The information ratio is defined as the annualized residual return to its residual risk. i.e.

$$IR_1 = \frac{\alpha_p}{\omega_p} \quad (\text{Rasmussen, 2003}). \quad (3.47)$$

If we look to the future, the information ratio is the expected level of annual residual return per unit of annual residual risk (Grinold & Kahn, 1999). As previously defined, the residual risk and residual return are unexplained by the return on the benchmark. In regression analysis of benchmark return on portfolio return, α is the intercept on the regression or simply the mean of residual returns:

$$\alpha = r_p - \beta_p \cdot r_B \quad (3.48)$$

and ω is the standard deviation of alpha. So, running a regression of benchmark returns on portfolio returns yields a series of residual returns equal to

$$\alpha = r_{pi} - \beta_P \cdot r_{Bi}, \quad (3.49)$$

where i denotes the observation number. The mean of these residual returns is thus the expectation of return, based on benchmark return. Rasmussen (2003) points out that one could argue that when evaluating active management, using active return in the numerator and active risk (tracking error) in the denominator of the information ratio would be more appropriate since we would then have a measure of actually obtained active return relative to actually incurred active risk. This would result in an information ratio defined as

$$IR_2 = \frac{r_P - r_B}{\psi_P}. \quad (3.50)$$

This would actually be quite intuitive and also akin to the well-known Sharpe ratio, which of course measures total portfolio return relative to the total portfolio risk. The reason why information ratio should be calculated using residual return and residual risk is that we are interested in isolating the performance of the manager, not benchmark. Thus, using residual return and residual risk instead of active return and active risk to calculate the information ratio means that the manager's performance relative to the benchmark is measured only in terms of the portion of active return and active risk that is not due to benchmark behavior – that is, the active performance that is due to the manager's skill and/or luck. In other words, the information ratio measures the quality of active manager's activeness relative to the benchmark, but independent of the return on the benchmark itself. (Rasmussen, 2003)

3.5.4 Value added

The active manager's first objective is to try to add value to client by deviating from the benchmark. Thereby creating tracking error that should result in a higher return than that of the benchmark. If the active manager does not hold superior information, he should not be active and thus be passive and replicate the benchmark. Whether an active manager is capable of producing superior returns under strict risk control is exactly what the information ratio attempts to discover. So, is the manager able to add value to the client by deviating from benchmark?

When examining the definition of information ratio, it is clear that the manager can only add value for the client by increasing residual return relative to residual risk.

$$IR = \frac{\alpha}{\omega} \Leftrightarrow \alpha = IR \cdot \omega \quad (3.51)$$

However, this also means that the expected residual return or alpha can be increased by rising the residual risk, but this does not necessarily mean increased value added for the client. The value added can only be increased by boosting the information ratio. This will increase the expected residual return for a given level of residual risk (Rasmussen, 2003). Grinold and Kahn (1999) defines the value added by the manager as follows:

$$VA = \alpha - \lambda \cdot \omega^2 \quad (3.52)$$

where λ is a measure of the client's risk aversion. Using the above-mentioned definition of alpha we can rewrite this

$$VA = IR \cdot \omega - \lambda \cdot \omega^2. \quad (3.53)$$

The value added by the manager is thus a concave function of residual risk, which means that the job of the active manager is to choose the level of residual risk that maximizes value added for the client depending on the client's risk aversion. Thus, we need to take partial derivative of VA with respect to residual risk and setting it equal to zero:

$$\frac{\partial VA}{\partial \omega} = IR - 2\lambda \cdot \omega = 0. \quad (3.54)$$

This gives us an optimal level of residual risk:

$$\omega^* = \frac{IR}{2\lambda}. \quad (3.55)$$

Using this result we can easily calculate the optimal value added which is

$$VA^* = \frac{(IR)^2}{4\lambda}. \quad (3.56)$$

However, identifying the client's risk aversion coefficient λ is not an exact science and that makes the identification of optimal level of residual risk quite difficult. Although the client's exact risk aversion may be very difficult to measure, there is still very important point to be made here. Regardless of the client's risk aversion coefficient, we can concentrate on the numerator and see that the measure of value added for the client is higher for higher levels of the information ratio. So even if the client doesn't have identifiable risk aversion, the goal of the manager is to seek to boost

the information ratio. This is because per unit of residual risk the manager delivers higher residual return for higher information ratios (Rasmussen, 2003).

As conclusion, there is a fundamental law of active management as Grinold and Kahn (1999) point it out. The information ratio is expressed in terms of two components. The first is the *information coefficient* (IC), which measures the manager's forecasting ability or skill. It is determined by the correlation between forecast and realized alphas or residual returns. The second is *breadth* (BR), defined as the number of independent bets per year:

$$IR = IC \cdot \overline{BR}. \quad (3.57)$$

The information coefficient is thus a measure of the forecasting ability of the manager. More precisely it is the correlation between forecast alphas and realized alphas. Breadth is defined as the number of independent bets. "Independent" in this context means separate bets that are based on perceived superior information regarding an asset class, which does not pertain to other asset classes (Grinold & Kahn 1999 and Rasmussen 2003).

4 EMPIRICAL STUDY

4.1 Methodology

The empirical part of the study is constructed as follows: first we describe the data, which is divided into in-sample and out-of-sample parts. Then we derive expected excess returns using different estimation methods discussed in chapter two and we also discuss the sensibility of the results. After that, the strategic portfolios are optimized based on the excess return estimates and also their in-sample and out-of-sample performances are investigated. These results are long-term strategic allocation decisions, which can be used as a neutral allocation in the tactical asset allocation. In chapter 4.5 we introduce risk factors in the returns on stocks and bonds that will be used to forecast expected returns. Finally the vector autoregressive model is estimated to create forecasts for the Black-Litterman model and an empirical test for the Black-Litterman model is run. Results can be found in chapter 4.6. All empirical statistics are performed using R¹⁴.

4.2 Description of the data

The following equity and bond time series are used in the empirical study¹⁵:

- MSCI Europe (Em)
- MSCI USA (US)
- MSCI Japan (Japan)
- MSCI Emerging markets (Em)
- Merrill Lynch Unsubordinated 7-10 year US Treasury bond total return index (Gvt)
- Citigroup investment grade total return index (IG)
- Merrill Lynch High Yield total return index (HY)
- Merrill Lynch US 3-month bill total return index (MM)

¹⁴ www.r-project.net

¹⁵ Bond price data is from Bloomberg and stock price data from www.msibarra.com

- US Treasury 1-month yield index as a risk free rate¹⁶

The four first indices are country equity indices provided by Morgan Stanley. The following four indices are debt indices from the United States. The first debt index represents a long government bond investment, the investment grade index represents an investment in high rated corporate bonds, the high yield is for low-grade corporate bonds and the 3-month bill index acts as a short government bond investment. Short names in parenthesis will be used later on in this study.

We are interested in excess returns only, so the 1-month US Treasury return is subtracted from all asset-return series. Asset price indices are depicted in figures 10. and 11. below. Prices are calculated from logarithmic excess returns. The data is of monthly observations and is collected between 1/1988-11/2008. The in-sample period is 1/1988-12/1999 and the out-of-sample period is 1/2000-11/2008.

Table 1. Descriptive in-sample statistics of the asset classes. Numbers are monthly results.

	Em	Europe	Japan	US	MM	IG	Gvt	HY
Mean	0.0060	0.0049	-0.0038	0.0077	-0.0003	0.0018	0.0019	0.0038
Std	0.0700	0.0433	0.0716	0.0385	0.0006	0.0117	0.0163	0.0179
Sharpe	0.0863	0.1122	-0.0525	0.2001	-0.4905	0.1501	0.1184	0.2117

Correlation matrix								
	Em	Europe	Japan	US	MM	IG	Gvt	HY
Em	1.0							
Europe	0.4564	1.0						
Japan	0.3634	0.5183	1.0					
US	0.5108	0.6168	0.3230	1.0				
MM	-0.0136	0.0577	0.0507	0.1046	1.0			
IG	-0.0125	0.2092	0.0999	0.4107	0.4612	1.0		
Gvt	-0.0521	0.1795	0.0948	0.3825	0.4643	0.9847	1.0	
HY	0.3437	0.3030	0.1491	0.4614	0.2039	0.3381	0.2606	1.0

¹⁶ Source: <http://www.federalreserve.gov/datadownload>

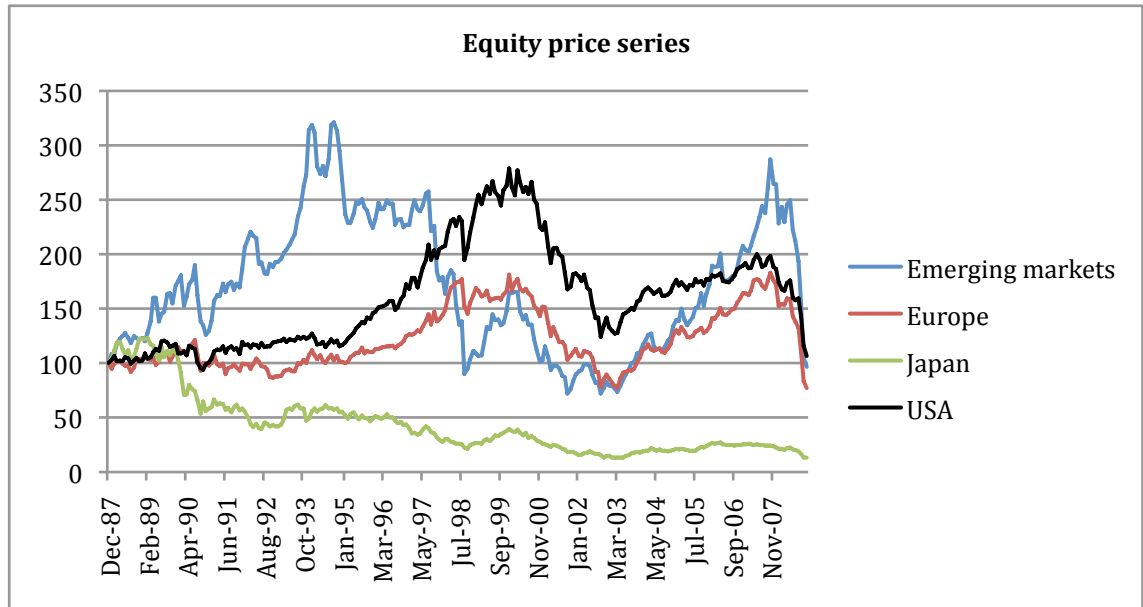


Figure 10. Stock price series calculated from monthly excess log-returns

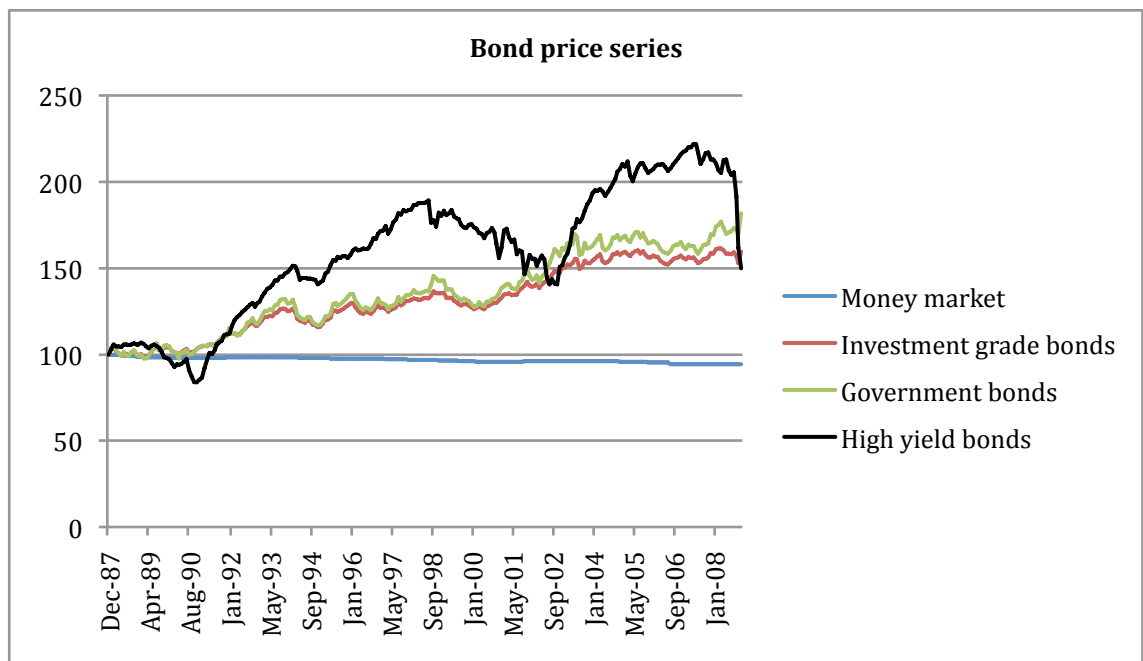


Figure 11. Bond price series calculated from monthly excess log-returns

4.3 Input estimation

As discussed in section 3.3, the traditional Black-Litterman model is based on the equilibrium returns, which are implied returns using the market weights. However, this is basically an impossible task in practice when we have several asset classes globally.

There just is no such data available. That is the reason why we have studied several other expected returns estimation methods.

The expected returns are estimated in six different ways (Table 2.). All methods are discussed in more detail in chapter 2. The estimation period is an in-sample period: 01/1988-12/1999. The first column represents the pure historical mean. It should be noted that the in-sample period ends just before the stock market crash in 2000, and it is obvious that historical means calculated from the part of the strongest bull market period in history, are strongly upward biased. Returns in the second column are calculated using James-Stein shrinkage method, where historical means are forced towards minimum variance portfolio return (recall that minimum variance portfolio does not need any information on expected returns). The shrinkage factor is calculated as in the equation (2.55). The third and fourth columns are based on the capital asset pricing model, where the market portfolio is 50% international bond index (Citygroup)¹⁷ and 50% international stocks index (MSCI World)¹⁸. As discussed earlier, it is hard to define what really is the “market portfolio”, but equally weighted international stock-bond portfolio could reasonably be argued to represent an average investor’s well-diversified portfolio.

Relatively high expected excess return on the Japanese stocks is a result from strong covariance with the market portfolio. The CAPM shrinkage returns are calculated using Wang’s (2006) method where historical returns and the CAPM returns are equally weighted. Implied returns from resampling are calculated by resampling long-only portfolios (target excess return 5%) 500 times and averaging the resulting portfolios (see chapter 2.2). Using these weights we calculated the market risk aversion coefficient ($\delta = 10.85$). Using reverse optimization we came up with the corresponding implied returns. Results in the last column are calculated using the implied returns from equally weighted portfolio ($\delta = 4.31$).

When taking a closer look on the estimated expected returns it can be said that implied returns from resampling seem to be most intuitive. Historical returns as well as James-Stein shrinkage implicates that in the long run, Japanese stocks yield quite strong negative excess returns. These results can easily be ruled out, since they were true, Japanese corporations are making loss in the long run and they would disappear (and basically the whole economy for the same reason). Even though Japanese stocks have been depressed since 1990, it is reasonable to expect higher returns than a risk free rate in the long haul. Also negative money market excess returns do not seem to have any economic reasoning, because risk free rate is calculated from 1-month t-bills while money market returns are calculated from 3-month t-bills. Investors require some kind

¹⁷ Source: Bloomberg

¹⁸ Source: www.msibarra.com

of a risk premium for longer investments even though the investment would be considered as “risk free” (this usually means that the investment is default free), because interest rate risk, inflation risk and (some kind of) default risk occurs. That is why we could conclude, that the money market return should be quite close to zero, but still above it. Since the maturity of the money market investment is only two months longer than the maturity of the risk free investment.

According to Arnott & Bernstein (2002), US stocks have yielded 2,4% return over the long-term government bonds since 1810. Dimson et al (2003) also studied historical returns and concluded that US stock market has performed slightly better than other countries. Using these studies as a reference, we could say that implied returns from the resampling seem to be again the most realistic compared to other results. Excess returns presented here are returns over the short-term risk free rate (1-month t-bill return) and hence can be considered as real returns since 1-month t-bills should be quite close to inflation. Arnott & Bernstein (2002) concluded that historically US stocks have yielded 6,1% real return and long term government bonds 3,7%. Thus, 6,72% excess return on US stocks seem to be quite close to their findings, but return on government bonds too low. The implied returns from the resampling suggest that Japanese and European stock markets would provide investors slightly less than 5% and the US stock market about 6,5% over the risk free rate. Higher emerging market equity returns are also quite sensible, since they are considered to be more risky investments than equity investments in developed countries.

Table 2. Expected annual excess returns using different return estimation methods.
Time period in sample data 1/1988-12/1999

	Historical mean	James-Stein shrinkage	CAPM	CAPM shrinkage	Implied returns from resampling	Implied returns from 1/n
Emerging markets	7.23%	4.29%	4.68%	5.95%	7.57%	6.34%
Europe	5.82%	3.42%	4.76%	5.29%	4.94%	4.10%
Japan	-4.51%	-2.91%	6.95%	1.22%	4.40%	6.36%
USA	9.24%	5.52%	4.18%	6.71%	6.72%	3.56%
Money market	-0.34%	-0.35%	0.01%	-0.16%	0.03%	0.01%
Investment grade bonds	2.10%	1.14%	0.83%	1.46%	1.35%	0.49%
Government bonds	2.31%	1.27%	1.10%	1.71%	1.69%	0.61%
High yield	4.54%	2.64%	1.02%	2.78%	2.95%	1.06%
Average	3.30%	1.88%	2.94%	3.12%	3.71%	2.82%

4.4 Strategic weights

Using the above calculated expected excess returns we calculated the optimal short sell constrained portfolios using each return estimation method studied in chapter 2 (Figure 12). Surprisingly, James-Stein shrinkage portfolio seems to be less diversified than the portfolio, where the pure historical returns are used. In addition, James-Stein portfolio does not contain any money market investments. The CAPM shrinkage portfolio does not include any money market investments either and seems to be less diversified than the pure CAPM portfolio. It should also be noted, that Japan is missing all the portfolios except CAPM portfolios, which is clearly due to large beta with the market portfolio.

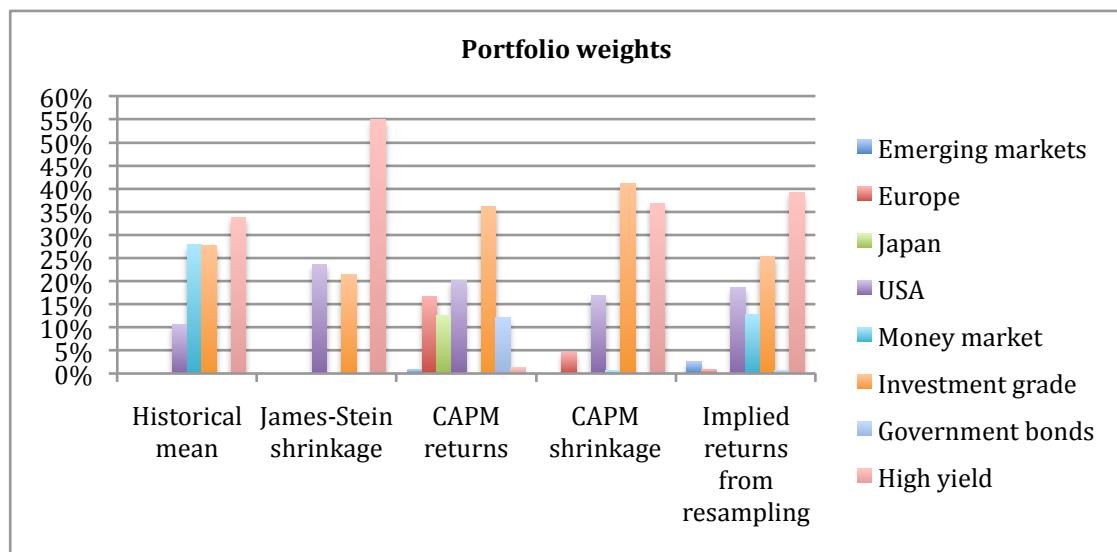


Figure 12. Portfolio weights when short sales are not allowed. Annual target excess return is 3%.

In Figure 13. the portfolios are optimized using the same inputs as above, but the covariance matrix is replaced with the shrinkage covariance matrix. It can be seen that extreme weights are now less extreme and some new asset classes have stepped into some portfolios. For example government bonds are taken into the last portfolio and stock weight is increased. At the same time, the investment grade bond weight is decreased. The portfolio, where the historical means are used, is more diversified as a slight percentage is allocated to emerging market equity. In the CAPM portfolio, the weight on the investment grade bonds is reduced while weights on the other assets are increased.

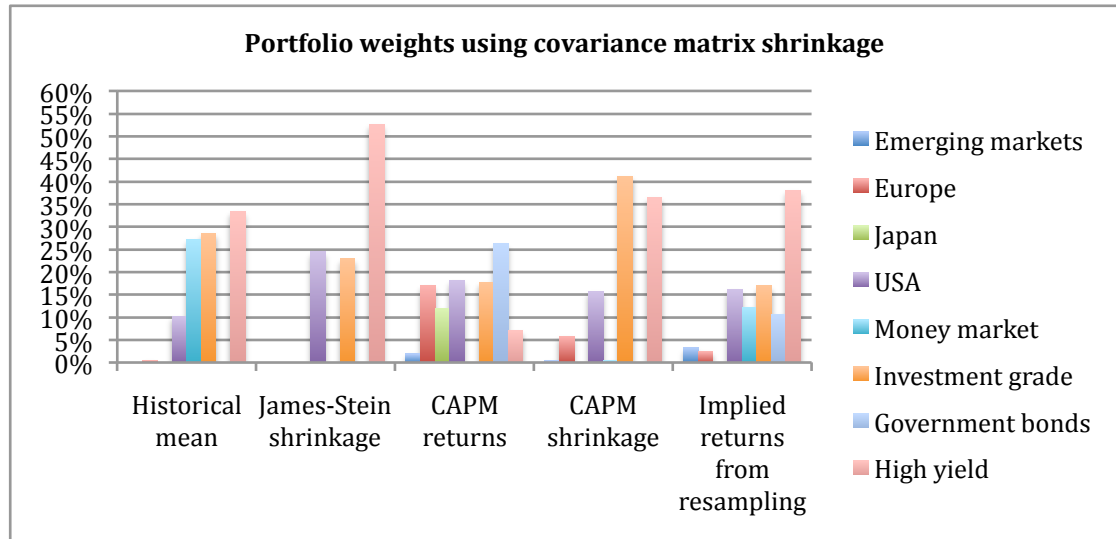


Figure 13. Portfolio weights when short sales are not allowed and covariance matrix shrinkage is used. Annual target excess return is 3%.

Table 3. below illustrates the real performance of the above derived strategic allocations in-sample and out-of-sample. Highest Sharpe ratios are underlined>. Surprisingly, the portfolio where the historical means were used performed best of all out-of-sample portfolios. However, when covariance matrix shrinkage was used, the historical mean portfolio was second best portfolio in-sample while CAPM shrinkage was the best. One important insight is that covariance matrix shrinkage enhanced all of the in-sample portfolio performances. Moreover, three out-of-sample portfolios out of six were better performing when covariance matrix shrinkage was used, compared to the sample covariance matrix performance. 1/n portfolio was the worst performing portfolio in-sample, no matter what covariance matrix was used. This contradicts the DeMiguel et al survey (2007). The figures 14. and 15. show the time series of the portfolio performances.

Table 3. Strategic monthly performance of different strategic weights.

Covariance matrix shrinkage						
Out-of-sample	Historical mean	James-Stein	CAPM	CAPM shrinkage	Resampled portfolio	1/n
Mean	-0.0006	-0.0018	-0.0020	-0.0012	-0.0012	-0.0022
Std	0.0153	0.0246	0.0233	0.0207	0.0201	0.0277
Sharpe	<u>-0.0387</u>	-0.0722	-0.0869	-0.0571	-0.0593	-0.0788

In-sample	Historical mean	James-Stein	CAPM	CAPM shrinkage	Resampled portfolio	1/n
Mean	0.0029	0.0036	0.0030	0.0039	0.0045	0.0029
Std	0.0115	0.0179	0.0218	0.0156	0.0183	0.0234
Sharpe	0.2493	0.2008	0.1379	<u>0.2497</u>	0.2442	0.1261

Historical covariance matrix						
Out-of-sample	Historical mean	James-Stein	CAPM	CAPM shrinkage	Resampled portfolio	1/n
Mean	-0.0005	-0.0018	-0.0022	-0.0011	-0.0016	-0.0024
Std	0.0154	0.0249	0.0225	0.0206	0.0251	0.0273
Sharpe	<u>-0.0348</u>	-0.0712	-0.0983	-0.0542	-0.0624	-0.0860

In-sample	Historical mean	James-Stein	CAPM	CAPM shrinkage	Resampled portfolio	1/n
Mean	0.0024	0.0036	0.0020	0.0032	0.0036	0.0017
Std	0.0117	0.0179	0.0220	0.0159	0.0188	0.0237
Sharpe	<u>0.2042</u>	0.2006	0.0923	0.2037	0.1928	0.0703

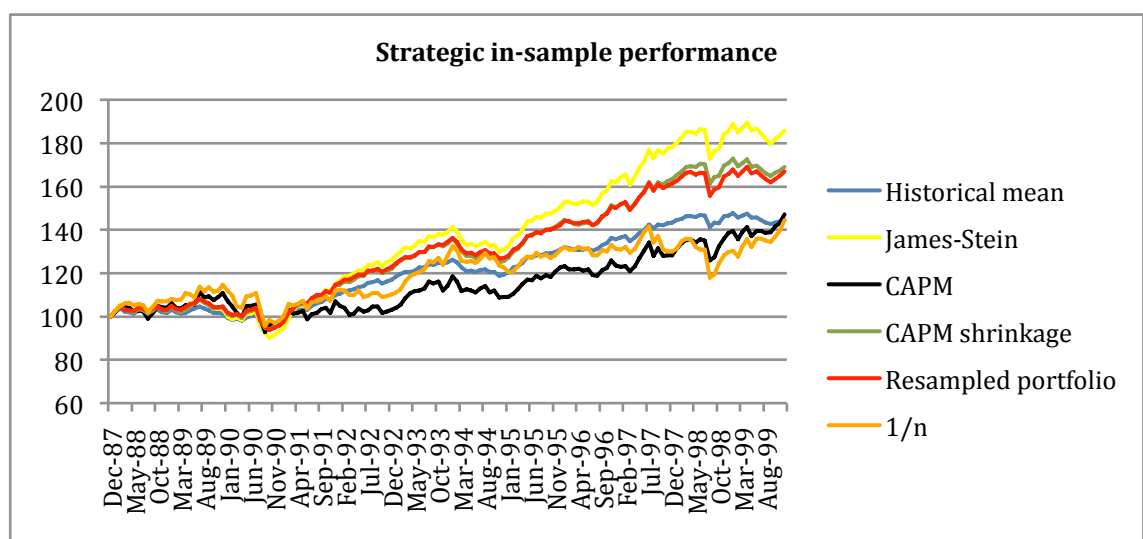


Figure 14. Strategic in-sample performance. Time series are calculated on excess return basis.

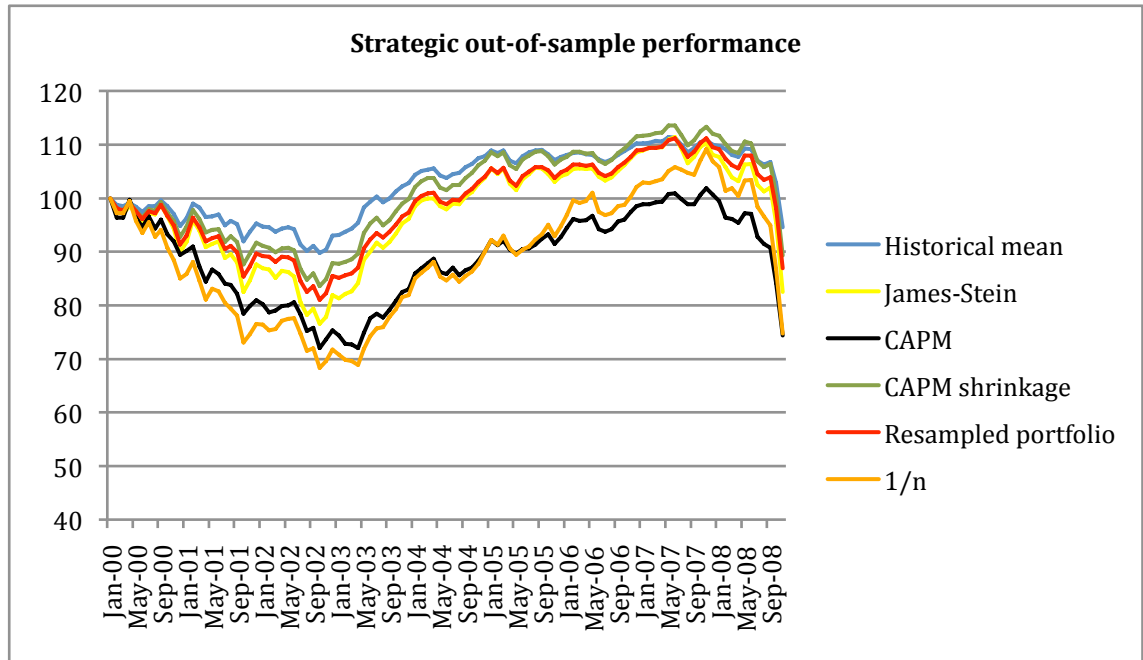


Figure 15. Strategic out-of-sample portfolio performance. Time series are calculated on excess return basis.

4.5 Vector autoregressive model for return forecasting

“Active management is forecasting. Without forecasts, manager would invest passively and choose the benchmark.” (Grinold & Kahn, 1999)

Vector autoregressive models (VARs) are so called multivariate models that include more than one independent variable. The VAR is a systems regression model that can be considered a kind of hybrid between the univariate time series models and the simultaneous equation models. The simplest case is the bivariate VAR-model, where there are only two variables y_{1t} and y_{2t} . The current values of both of these dependent variables depend on different combinations of the previous k values of both variables and error terms

$$\begin{aligned} y_{1t} &= \beta_{10} + \beta_{11}y_{1t-1} + \dots + \beta_{1k}y_{1t-k} + \alpha_{11}y_{2t-1} + \dots + \alpha_{1k}y_{2t-k} + u_{1t} \\ y_{2t} &= \beta_{20} + \beta_{21}y_{2t-1} + \dots + \beta_{2k}y_{2t-k} + \alpha_{21}y_{1t-1} + \dots + \alpha_{2k}y_{1t-k} + u_{2t} \end{aligned} \quad (4.1)$$

where u_{it} is a white noise disturbance term with

$$E(u_{it}) = 0, \quad (i = 1, 2), \quad E(u_{1t}, u_{2t}) = 0 \quad (4.12)$$

VAR models have several advantages compared with univariate (e.g. ARIMA) time series models of simultaneous equations structural models. First of all, the research doesn't need to specify which variables are endogenous or exogenous because all are endogenous. In other words, all the equations in the system are identified and thus, requirements for solving simultaneous equations structural models are met. Second, VARs allow the value of the variable to depend on more than just its own lags or combinations of white noise terms. VARs are more flexible than univariate AR models and therefore VARs offer a very rich structure, implying that they may be able to capture more features of the data. If there are no contemporaneous terms on the right hand side of the equation i.e. no y_{2t} term on the right hand side of the equation y_{1t} and vice versa, it is possible to simply use OLS separately on each equation. This arises from the fact that all variables on the right hand side are pre-determined, i.e. at the time t they are all known. There is also evidence that the forecasts generated by VARs are often better than traditional structural models. (Brooks, 2002)

Some researchers have argued that the a-theoretical nature of VARs without contemporaneous terms (reduced form VARs) leaves them unstructured and their results difficult to interpret theoretically. So, lets assume that we have a simple bivariate VAR(2) model and we place contemporaneous feedback terms in both of the equations, we get

$$\begin{aligned} y_{1t} &= \beta_{10} + \beta_{11}y_{1t-1} + \alpha_{11}y_{2t-1} + \alpha_{12}y_{2t} + u_{1t} \\ y_{2t} &= \beta_{20} + \beta_{21}y_{2t-1} + \alpha_{21}y_{1t-1} + \alpha_{22}y_{1t} + u_{2t} \end{aligned} \quad (4.13)$$

This is easy to stack up using matrices and vectors

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \alpha_{1t-1} \\ \alpha_{21} & \beta_{21} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} \alpha_{12} & 0 \\ 0 & \alpha_{22} \end{pmatrix} \begin{pmatrix} y_{2t} \\ y_{1t} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \quad (4.14)$$

This would be known as a VAR in primitive form. The contemporaneous terms can be taken over to the left hand side and written as

$$\begin{pmatrix} 1 & -\alpha_{12} \\ -\alpha_{22} & 1 \end{pmatrix} \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \alpha_{1t-1} \\ \alpha_{21} & \beta_{21} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \quad (4.15)$$

or

$$Ay_t = \beta_0 + \beta_1 y_{t-1} + u_t \quad (4.16)$$

We discussed several risk factors that explain stock and bond returns in chapter 3.4. The following risk factors are chosen for the asset return forecasting in VAR model:

1. Term spread: Market yield between 10-year and 3-month US treasury bond (constant maturity)¹⁹
2. Default spread: Spread between US Baa corporate bond yield and 10-year US Treasury note yield
3. TED spread: Spread between Eurodollar deposit rate (London) and US 3-month t-bill
4. Logarithmic change in real dividend (US)²⁰

The first factor tries to catch the information included in the shape of a yield curve. It is a widely proven fact, that the yield curve shape forecasts economic growth as well as recessions relatively well²¹. It also includes inflation expectations that are a very important factor concerning the bond and stock returns. Second and third factors are well known measures to illustrate the general risk aversion of the investors. These factors measure risk premiums that investors require for their investments. The spreads get high values when investors see the general riskiness of the economy arisen. When companies are expected to face hard times, spreads are high because risky investments became more risky as probability to default increases. Hence investors require higher premiums on low-grade corporate bonds. For example default spreads have sky rocketed to all time high during the global financial crisis, that we are having at the moment. The TED-spread illustrates the risk that lies in the interbank market. When risks among the financial institutions are expected to increase, so does the TED-spread. The cost of borrowing becomes higher as risk premiums required increase. This is obviously the case in the credit crises. The fourth factor is a real dividend growth rate, which Arnott and Bernstein (2002) used in their study examining the historical stock risk premium. They argue that in the long run, the dividend growth rate and the total economy growth rate need to be close to each other. Dividends and their growth rate measure the ability of the companies to create value for the stockholders. Hence it can be argued to be a strong determinant for the stock market return.

¹⁹ Data for the first three factors: <http://www.federalreserve.gov/datadownload>

²⁰ Source: www.irrationalexuberance.com

²¹ see Cambell, Harvey 1991 and Andrew Ang et al 2006

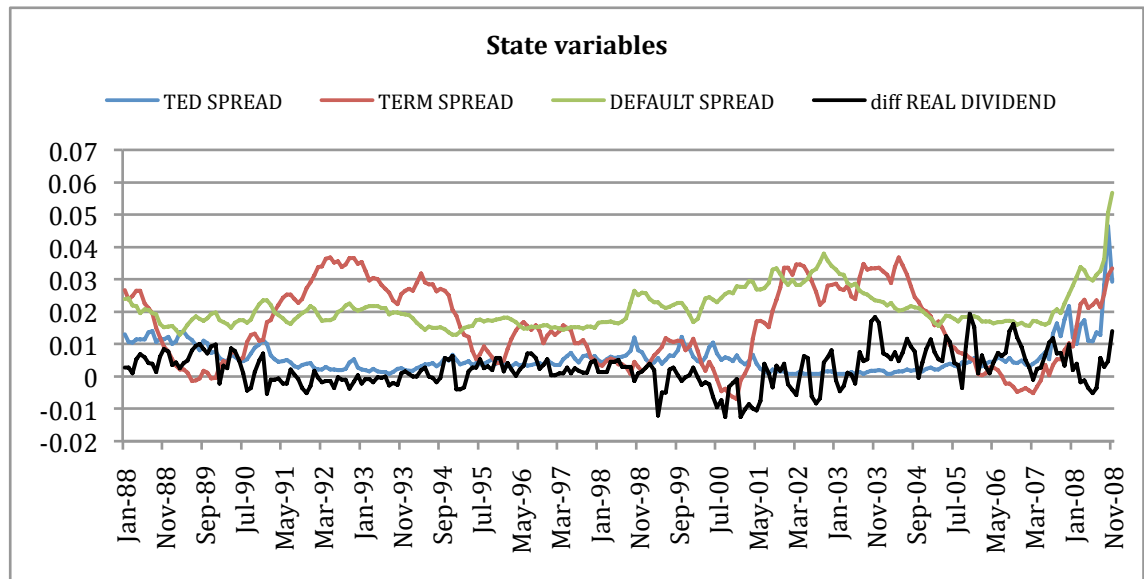


Figure 16. State variable behavior over sample

Table 3. illustrates the estimated VAR-model. It contains eight asset classes and four fundamental factors discussed above (and constant). *Em* stands for emerging market equity, *Europe* for European equity, *Japan* for Japanese equity and *Us* for the United States equity. *MM* stands for money market investment, *IG* (US) for investment grade bonds (US), *Gvt* for government bonds (US) and *HY* for high yield (US). T-values are in parenthesis and values over +/-2 are underlined.

Table 4. Estimated VAR-model for return forecasting. Estimation period is in-sample 1/1988-12/1999.

	Em_t	$Europe_t$	$Japan_t$	Us_t	MM_t	IG_t	Gvt_t	HY_t	TED_t	$Term_t$	$Default_t$	$dRealD_t$	Const.	R^2
Em_{t+1}	0.063 (0.57)	0.057 (0.28)	-0.067 (-0.65)	-0.003 (-0.01)	10.58 (0.82)	-4.290 (-1.30)	2.479 (1.06)	0.396 (0.90)	1.598 (0.66)	0.047 (0.07)	6.284 <u>(2.88)</u>	0.578 (0.29)	-0.117 (-2.76)	0.11
$Europe_{t+1}$	-0.038 (-0.54)	-0.160 (-1.24)	0.009 (0.13)	0.028 (0.18)	1.441 (0.18)	1.415 (0.67)	-0.820 (-0.55)	0.268 (0.95)	0.368 (0.24)	-0.272 (-0.64)	0.690 (0.50)	0.795 (0.64)	-0.008 (-0.29)	0.06
$Japan_{t+1}$	-0.096 (-0.85)	-0.446 <u>(-2.17)</u>	0.046 (0.45)	0.406 (1.60)	13.061 (1.00)	-1.940 (-0.58)	1.426 (0.60)	-0.035 (-0.08)	0.697 (0.29)	0.460 (0.68)	5.501 <u>(2.49)</u>	1.567 (0.79)	-0.115 (-2.66)	0.13
Us_{t+1}	-0.062 (-1.03)	-0.094 (-0.84)	0.028 (0.50)	-0.087 (-0.63)	6.269 (0.88)	0.346 (0.19)	-0.205 (-0.16)	0.097 (0.40)	1.899 (1.43)	-0.223 (-0.60)	0.422 (0.35)	-0.418 (-0.38)	-0.004 (-0.19)	0.10
MM_{t+1}	1.50e-03 (1.87)	-2.2e-03 (-1.50)	-8.9e-04 (-1.19)	-4.3e-03 <u>(-2.35)</u>	-1.6e-02 (-0.17)	1.42e-02 (0.59)	-4.7e-05 (-0.01)	5.86e-03 (1.81)	-7.5e-02 <u>(-4.26)</u>	-4.5e-03 (-0.92)	9.27e-04 (0.06)	1.237e-02 (0.86)	1.6e-04 (0.53)	0.44
IG_{t+1}	0.001 (0.06)	-0.026 (-0.82)	0.003 (0.19)	-0.070 (-1.79)	2.138 (1.06)	1.104 <u>(2.14)</u>	-0.610 (-1.68)	-0.015 (-0.22)	0.609 (1.62)	0.153 (1.46)	-0.767 <u>(-2.25)</u>	0.070 (0.23)	0.01 (1.52)	0.20
Gvt_{t+1}	-0.001 (-0.025)	-0.026 (-0.59)	0.001 (0.07)	-0.127 <u>(-2.36)</u>	1.622 (0.59)	1.613 <u>(2.27)</u>	-0.856 (-1.71)	-0.035 (-0.363)	0.736 (1.42)	0.196 (1.36)	-1.127 <u>(-2.40)</u>	0.049 (0.115)	0.015 (1.68)	0.22
HY_{t+1}	-0.019 (-0.72)	-0.043 (-0.89)	0.010 (0.40)	0.125 <u>(2.07)</u>	-1.215 (-0.39)	0.747 (0.94)	-0.425 (-0.76)	0.101 (0.95)	0.260 (0.45)	0.437 <u>(2.70)</u>	0.224 (0.43)	-0.261 (-0.55)	-0.01 (-0.98)	0.24

R-squared measures are relatively high, especially for bond returns. Statistically the most important factor seems to be the default spread. Even though term spread has only one t-value over two (for high yield bonds), the t-values are quite high for bonds returns. In general it seems to be easier to forecast bond returns than stock returns since almost all t-values are higher for bond returns than stock returns.

4.6 Overview on the result

The expected returns from resampling were used as the prior vector in the Black-Litterman model, because they were intuitively the most realistic returns as discussed in chapter 4.4. The optimal portfolio weights derived from these returns were used as the strategic allocation in the empirical analysis (see figure 17). If there were no forecasts for the asset returns, and thus any deviations from the prior vector were made, the portfolio performance would have been the same as the passive benchmark performance. Even though the resampled portfolio was not the best performing portfolio out-of-sample (see Table 3), we are more interested if the tactical allocation using the Black-Litterman model is able to create positive alpha, i.e. deviate from the strategic allocation and enhance the risk adjusted return.

The Black-Litterman model was updated end of the each month and portfolio rebalanced accordingly. The VAR-model was estimated in-sample in order to forecast excess returns for the first out-of-sample month (January 2000). During the out-of-sample, the VAR-model was estimated for $t+1$ at the end of each month adding the latest state variable observations to the model. Covariance matrix was estimated only in-sample and it did not vary during the out-of-sample. The value for τ was chosen to be conservatively 0.5, because there is no consensus in the literature whether it should be close to zero or close to one.

The Black-Litterman model and its derived inputs in this study beat the strategic allocation (benchmark) out-of-sample, in terms of risk-adjusted returns. Tactical allocation was able to create annual alpha out-of-sample 1.96% (t-value 1.31). This is calculated using the equation (3.48) as follows:

$$\alpha = r_p - \beta_P \cdot r_B,$$

where $\alpha = 0.001634$ and $\beta = 0.7845$. This is a linear regression of the monthly TAA returns against the monthly benchmark returns. Monthly Sharpe ratio of the TAA was 0,034 and benchmark -0,059. The following table 5. and figure 16. conclude the performance:

Table 5. Monthly TAA out-of-sample results

	Return	Std	Sharpe	Active return	Active risk	Residual return α	Residual risk	IR
TAA	0.0007	0.0203	0.0343	0.0019	0.0135	0.0016	0.0147	0.1398
Benchmark	-0.0012	0.0201	-0.0593	-	-	-	-	-

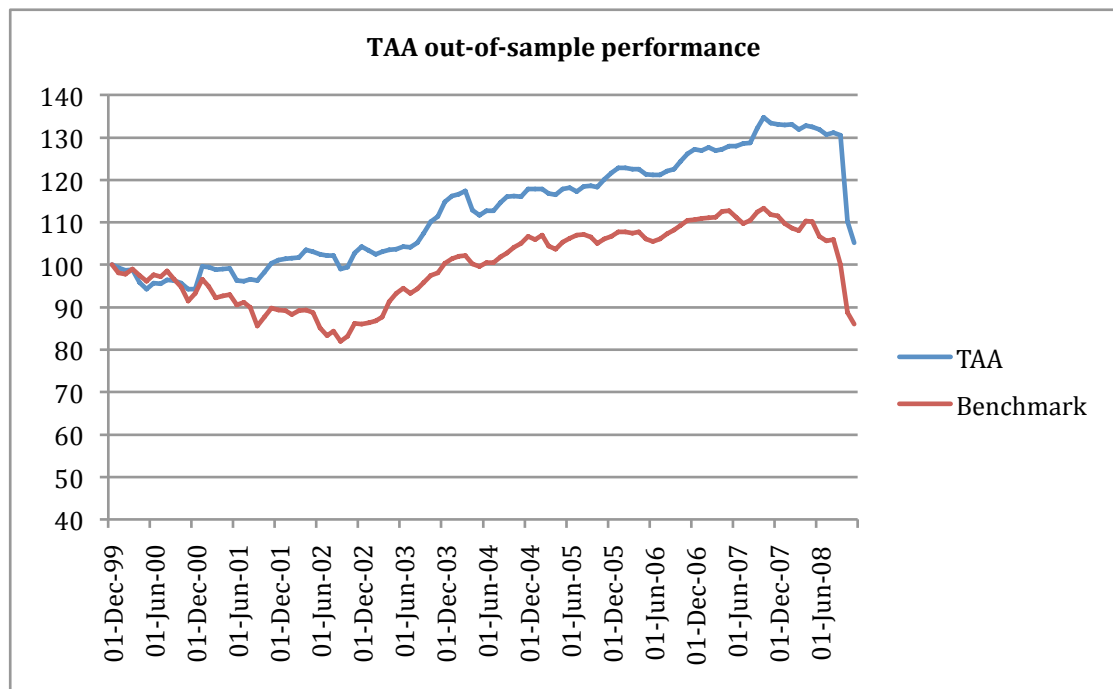


Figure 17. TAA and SAA out-of-sample performance

The Figure 17. depicts how the weights have varied during the out-of-sample period. The money market weight is quite high until June 2003. After that, the money market weight drops quite dramatically. This is quite accurately the time when the first, strong bull market of the 21st century started. High yield bonds and investment grade bonds replace largely the money market weight and also the stock market weight is increased after June 2003. In fall 2004, the money market weight is increased again, and is replaced by investment grade bonds a few times until May 2008. Note that the stock market weight is significantly increased between June 2003 and January 2008, being almost 50% at its best. In December 2007, the stock market weight is dramatically reduced and the money market weight increased. However, the European stock market weight is quite high in May and November 2008.

At least one practical insight can be drawn from these results. The money market investment acts a quite important role in the portfolio, even though it has expected excess return over the long run near zero (see table 1). Given this fact, the high money market weight needs to be due to attractive correlation characteristics with other asset

classes and its low risk. Also high yield bonds need to have very attractive risk, return or correlation characteristics, as their high weight in the portfolio indicates. High yield actually has the highest Sharpe ration in-sample, so it is no wonder that high yield bonds act a very significant role in the portfolio during the whole out-of-sample. Most of the time, stock market weight seems to be roughly 10-30%, which sounds quite low. However, investment grade bonds and high yield bonds largely compensate this as they have very attractive portfolio characteristics compared to stocks.

As a conventional investor tends to figure out which stock is the winning stock, this study suggests that using the sensible quantitative portfolio management model with broad variety of different asset indices (for example etf:s), one could have survived from this financial turmoil (by so far) with relatively dry feet. The Black-Litterman model with vector autoregressive model forecast can lead to favorable results quite easily, and still, further improvements may be possible. For example separate models for stocks and bonds could be used, since the VAR-model easily suffers from over parameterization when number of state variables is increased. In addition, we used constant, shrinkage covariance matrix estimated in-sample, which could be replaced with a time varying one, or it could be even forward looking (see Scherer 2002, 129). In addition, we did not take into consideration currencies for simplicity, but active currency risk management could enhance results quite dramatically (see Mertens & Zimmermann 2002).

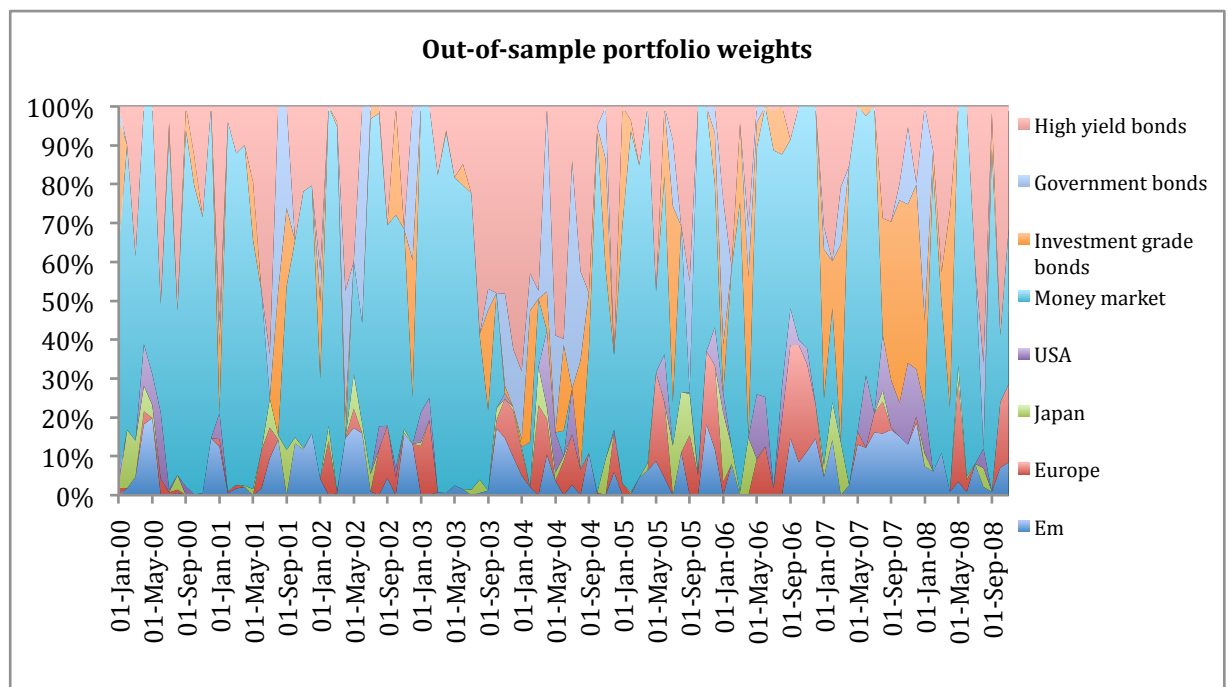


Figure 18. Out-of-sample portfolio weights

5 CONCLUSION OF THE STUDY

We have studied several issues that quantitative portfolio managers face in practice. Even though the theory behind the optimal portfolio choice is very straightforward, it certainly is not that in practice. The main problem in finding an optimal portfolio is estimating the expected returns, which is primarily a statistical problem. Using pure historical mean returns results in undesired portfolio weights because of an estimation error. Moreover, estimating the covariance structure of assets is also problematic in a statistical sense, but fortunately that plays a relatively minor role compared to the expected returns. We have investigated several approaches to estimate the expected returns and how to enhance historical estimates, including the covariance matrix. However, the out-of-sample test did show, that the optimal portfolio using the historical returns performed better than using the alternative estimation methods. It is still confusing, that the corresponding portfolio was not the best performing portfolio in-sample.

The empirical study strengthens the argument of Ledoit & Wolf (2003), that no one should be using the historical covariance matrix in portfolio optimization, when we are able to use shrinkage estimation on a sample covariance matrix. We were able to enhance Sharpe ratios of all strategic portfolios in-sample using the covariance matrix shrinkage, and out-of-sample three out of six. Even though portfolio resampling has no economic reasoning, it produced intuitively and empirically the most interesting results. The use of the reverse optimization on resampled portfolio weights resulted in the most intuitive expected returns of all six estimation methods. In addition, the long only resampled optimal portfolios along the efficient frontier were clearly more diversified than the traditional mean-variance optimized portfolios.

The most important finding of this study is that a model based portfolio management is able to produce positive alpha and beat the passively managed benchmark. Also the performance of the $1/n$ portfolio was very poor compared to the results of the Black-Litterman model. A naive $1/n$ portfolio performed very poorly compared to other strategic portfolios, which contradicts the survey of DeMiguel et al (2007). Even though the Black-Litterman model is theoretically quite complicated, we can conclude that it certainly is relatively easy to implement and results are encouraging. The VAR-based forecasts were surprisingly good and these forecasts were successfully implemented using the Black-Litterman model.

Further research could focus on the covariance matrix forecasting and taking currencies into consideration. Mertens & Zimmermann (2002) states that currencies are a very important part of generating alpha in the global context – probably more important than equity TAA strategies.

6 REFERENCES

- Ang, A., Piazzesi, M., & Wei, M. (2006). What does the Yield Curve Tell us about GDP Growth? *Journal of Econometrics* .
- Arnott, R., & Bernstein, P. (2002, January). What risk premium is "normal"? *Journal of Portfolio Management* .
- Best, M. J., & Grauer, R. R. (1991, January). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. *The review of financial studies* , 315-342.
- Black, F., & Litterman, R. (1992). Global portfolio optimization. *Financial Analysts Journal* , 28-43.
- Brandt, M. W. (2004). Portfolio choice problems. *Y. Ait-Sahalia and L.P Hansen, eds., Handbook of Financial Econometrics, Elsevier Science: Amsterdam* .
- Brinson, G., Hood, R., & Beebower, G. (1986). Determinants of portfolio performance. *Financial Analysts Journal* .
- Britten-Jones, M. (1999). The Sampling Error in Estimates of Mean-Variance Efficient Portfolio Weights. *The Journal of Finance* , 54 (2), 665-671 .
- Brooks, C. (2002). *Introductory econometrics for finance*. New York: Cambridge university press.
- Cambell, H. (1991). The Term Structure and World Economic Growth. *Journal of Fixed Income* .
- Campbell, J. Y., & Ammer, J. (1993). What moves the stock and bond markets? Variance decomposition for long-term asset returns. *The journal of finance* , 48 (1), 3-37.
- Chen, N.-F., Roll, R., & Ross, A. S. (1986). Economic forces and the stock market. *Journal of business* , 59, 383-404.
- Cochrane, J. H. (2005). *Asset pricing*. New Jersey: Princeton University Press.
- DeMiguel, V., Garlappi, L., & Uppal, R. (2007). *Optimal versus Naive Diversification: How Inefficient Is the 1/N Portfolio Strategy?*
- Dimson, E., Marsh, P., & Staunton, M. (2003). *Irrational optimism*. LBS Institute of Finance and Accounting Working Paper No. IFA397.
- Drobetz, W. (2001). How to the avoid pitfalls in portfolio optimization? Putting the Black-Litterman approach at work. *Financial markets and portfolio management* , 15, 59-75.
- Elton, E. J., Gruber, M. J., & Blake, C. R. (1995). Fundamental Economic Variables, Expected Returns, and Bond Fund Performance. *The Journal of Finance* , 1229-1256.

- Fama, E. F., & Gibbons, M. R. (1984). A comparison of inflation forecasts. *Journal of monetary economics* , 13, 327-348.
- Fama, E., & French, K. R. (1989). Business conditions and expected returns on stock and bonds. *Journal of financial economics* , 25, 23-49.
- Ferson, W. E., & Campbell, H. R. (1993). The risk and predictability of international equity returns. *The review of financial studies* , 6 (3), 527-566.
- Ferson, W. E., & Campbell, H. R. (1991). The variation of risk premiums. *The journal of political economy* , 99 (2), 385-415.
- Frankfurter, G. M., Phillips, M. H., & Seagle, J. P. (1971). Portfolio selection: the effects of uncertain means, variances and covariances. *Journal of financial and quantitative analysis* , 6, 1251-1262.
- Frost, P. A., & Savarino, J. (1988). For better performance: Constrain portfolio weights. *Journal of portfolio management* , 15, 29-34.
- Grinold, R., & Kahn, R. (1999). *Active portfolio management* (Vol. 2 edition). New York: McGraw-Hill.
- He, G., & Litterman, R. (1999). *The Intuition Behind Black-Litterman Model Portfolios*. Goldman Sachs & Company. Investment Management Research.
- Idzorek, T. M. (2004). *A Step-By-Step Guide to the Black-Litterman Model*. Zephyr Associates Inc.
- Jagannathan, R., & Ma, T. (2002, January). Risk reduction in large portfolios: a role for portfolio weight constraints.
- James, W., & Stein, C. (1961). *Estimation with quadratic loss*. Proceeding of the Fourth Berkeley Symposium on Mathematics and Statistics.
- Jorion, P. (1986). Bayes-Stein estimation for portfolio analysis. *Journal of financial and quantitative analysis* , 21 (3), 279-291.
- Ledoit, O., & Wolf, M. (2003). Honey, I shrunk the sample covariance matrix.
- Lee, W. (2000). *Theory and Methodology of Tactical Asset Allocation*. New Hope, Pennsylvania, USA: Frank J. Fabozzi Associates.
- Litterman, R. (2003). *Modern Investment Management: An equilibrium approach*. New Jersey: John Wiley & Sons.
- Luenberger, D. G. (1998). *Investment Science*. New York: Oxford University Press.
- Mankert, C. (2006). *The Black-Litterman Model*. Licentiate thesis in KTH: Royal Institute of Technology, Stockholm.
- Markowitz, H. (1987). *Mean-variance analysis in portfolio choice and capital markets*. Cambridge: Blackwell.
- Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance* , 7 (1), 77-91.

- Mertens, E., & Zimmermann, H. (2002). *International TAA strategies and currencies*. WWZ/Department of Finance. Working paper.
- Merton, R. C. (1980). On estimating the expected return on the market. *Journal of financial economics* , 8, 323-361.
- Michaud, R. O. (1998). *Efficient Asset Management*. Boston: Harvard Business School Press.
- Philips, T. K., Rogers, G. T., & Capaldi, R. E. (1996). Tactical Asset Allocation: 1977-1994. *Journal of Portfolio Management* , 57-64.
- Rasmussen, M. (2003). *Quantitative portfolio optimisation, asset allocation and risk management*. New York: Palgrave Macmillan.
- Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of economic theory* , 13, 341-360.
- Rubinstein, M. (2002). Markowitz's "Portfolio Selection": A Fifty-Year Retrospective. *The Journal of Finance* , 57 (3), 1041-1045.
- Satchell, S., & Scowcroft, A. (2000, September). A demystification of the Black-Litterman model: Managing Quantitative and Traditional Construction. *Journal of Asset Management* , 138-150.
- Scherer, B. (2002). *Portfolio Constructing and Risk Budgeting*. London, UK: Risk Books.
- Scherer, B. (2004). Resampled efficiency and portfolio choice. *Swiss society for financial market research* , 382-398.
- Sharpe, W. F. (1963). A Simplified Model for Portfolio Analysis. *Management in Science* , 9 (2), 277-293.
- Stein, C. (1955). Inadmissibility of the usual estimator for the mean of multivariate normal distribution. *Berkeley, University of California Press* , pp. 197-206.
- Wang, Z. (2005). A shrinkage approach to model uncertainty and asset allocation. *The review of financial studies* , 18 (2), 673-705.