NORWEGIAN SCHOOL OF ECONOMICS
BERGEN, SPRING 2018

Norwegian Equities Mean Reversion

Statistical Arbitrage as a Source of Views in the Black-Litterman Model

André Wattø Sjuve

Supervisor: Gunnar Stensland

Master Thesis, MSc in Economics and Business Administration, Financial Economics

Norwegian School of Economics

Abstract

A trading strategy incorporating s-scores and conditional mean returns in the Black-Litterman model is backtested over a 13-year period, 01.01.2005-29.12.2017, on the OBX index at the Oslo Stock Exchange. Estimating the trading signals using different techniques, and conducting a constrained optimisation, daily NOK-neutral long-short active portfolio weights are computed. In combination with the benchmark weights, the strategy is found to yield substantial profits gross and net of costs, but statistical evidence in support of a superior strategy hypothesis is lacking, i.e. the results are with a high probability a product of randomness. Seemingly, the strategy seems to benefit from volatility, outperforming the benchmark during the financial crisis, to then underperform in the low volatility years, 2016 and 2017. Additionally, with the high concentration of Energy companies in the chosen benchmark, the possibilities to make profitable trades in other sectors are capped, as seen by the low percentage share of strong trading signals becoming active positions within these sectors, and the poor performance of non-Energy sector-based portfolios. This thesis finds some support for previous research, in that high volatility regimes are linked to better performance and which sectors are fitting for a mean-reversion strategy.

Preface

Before you lies the dissertation *"Norwegian Equities Mean Reversion - Statistical Arbitrage as a Source of Views in the Black-Litterman Model"*, representing the completion of my Master's degree with specialisation in Finance at the Norwegian School of Economics. It was written on the account of my interest in the field of applied finance, attained during my time as a student in Bergen. The process has been rewarding, but at times challenging and frustrating.

I would like to extend gratitude to my supervisor, Gunnar Stensland, for valuable input early in the process. Additionally, to whomever else that have contributed to my education, thank you.

André Wattø Sjuve

Bergen, 18th June, 2018

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Following the seminal work of Harry Markowitz on portfolio optimisation, Black and Litterman (1990) offered an approach that overcame some of the instabilities of the traditional mean-variance framework [1] by introducing the idea of an equilibrium model for asset returns and subjective views about assets performance. The equilibrium expected excess returns (prior distribution) are based on the market portfolio, which is a product of the CAPM and is as an intuitive starting point, since it is directly linked linked to the market. Views are then able to shift the portfolio weights away from the prior based on investor's confidence. Consequently, a practitioner must synthesise private information into explicit return estimates and quantify uncertainty.

A few years prior to the Black-Litterman model, it is said that pairs trading originated within Morgan Stanley Vidyamurthy (2004) as a way of automating investment decisions. Following the technological advances and exponential growth in data, a wide range of quantitatively oriented hedge funds have emerged, many focusing on the use of different statistical arbitrage strategies. These strategies have existed side by side with the Black-Litterman model, until Liew and Roberts (2013) attempted to combine them in their paper *"U.S. Equity Mean-Reversion Examined"*.

The aim here, is to explore the same combination they employ on a set of liquid Norwegian equities, the constituents of the OBX index, from 2005 through 2017 and identify if such a blend consistently outperforms the index (benchmark). Assuming a zero daily risk-free rate, the strategy utilises statistical arbitrage to form views on the one-day expected excess return of the constituents, which is then combined with zero-equilibrium expected returns (Da Silva et al., 2009), in the Black-Litterman model to produce posterior return estimates. From this, a set of active positions are generated,

[1] See Michaud (1989)

tilting the total portfolio weights away from the benchmark. Applying the findings of Liew and Roberts, view confidence is directly linked to the extent of mispricing between the constituents and their industries represented by SPDR MSCI Europe sector ETFs, measured by the s-score.

The literature [2], confirms the richness of research done within tactical asset allocation, statistical arbitrage and Black-Litterman as isolated fields of research, but coverage of their combination is limited. As such, my choice of topic contributes to uncharted territory, by blending pairs trading with the Black-Litterman model within active management. To the best of my knowledge, the only paper trying to bridge these is Liew and Roberts (2013).

Where previous research paint with a broad brush, [3] this thesis focuses on a single index, the OBX. Generating active positions to tilt the benchmark weights using statistical arbitrage and the Black-Litterman model appears to be the first of its kind. Additionally, existing literature is mainly focused on U.S equities, but here the Norwegian equity market is considered.

The work presented draws most of its inspiration from the papers by Avellaneda and Lee (2010) and (Liew and Roberts, 2013). Both present impressive results, measured by strategy returns, as well as risk adjusted measures, such as Sharpe ratios. However, none of them explicitly address the challenges of converting Sharpe ratios from one frequency to another. Following Lo, autocorrelation and non-IID returns are corrected for when aggregating Sharpe ratios from a daily to a yearly frequency.

Using daily data from Børsprosjektet, along with publicly available information obtained from Newsweb, I reconstruct the OBX index weights for the period 03.01.2005 to 31.12.2017. This serves as input to construct the passive component of the trading strategy. Then, I pair each constituent in the OBX index with an ETF based on the constituent's GICS [4] code. With daily returns on these ETFs, obtained from Datastream, I calibrate the Ornstein-Uhlenbeck process and compute daily s-scores for each constituent.

Based on expected one-day excess returns and covariance matrices, the latter being estimated using an equally weighted approach as well as an exponential weighted moving averages approach, the derived posterior return estimates results in a set of active weights, that together with the benchmark weights form the total portfolio. With these weights, portfolio returns, and performance measures are computed and reported.

[2]For a full review see appendix E

[3]Avellaneda and Lee consider a trading universe consisting of $1,417$ stocks across 15 sectors and Liew and Roberts utilise 500 securities, covering nine sectors.

[4]Global Industry Classification Standard

The remainder of this thesis is structured as follows: Section 2 provides the theoretical base for every strategy component, section 3 describes the data, and provides descriptive statistics, section 4 outlines how the results are obtained, before they are presented and discussed in section 5. Finally, in section 6, some limitations and possible extensions are reflected upon, before some concluding remarks are offered.

Chapter 2

Theory

This chapter introduces the central theory. Intuition for statistical arbitrage is provided, with emphasis on pairs trading. Then key elements of the Ornstein-Uhlenbeck process is considered, the original Black-Litterman model is derived, modifications used are highlighted and insights on the link between Black-Litterman and active management is given. Longer derivations are found in the appendices. For notation, bold uncapitalized letters, symbols and numbers represent vectors, bold capitalised letters and symbols represents matrices and regular symbols represent numbers.

2.1 The Fundamentals of Pairs Trading

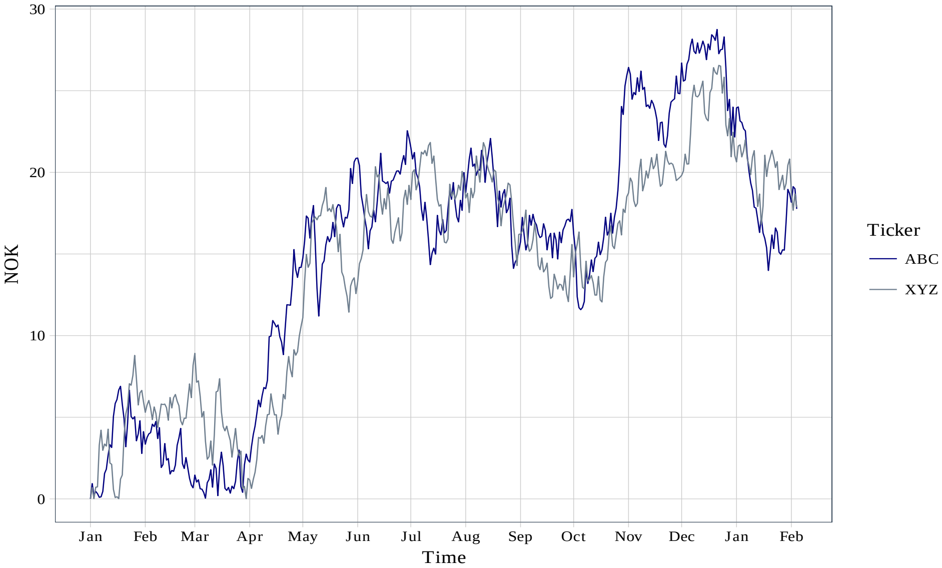
Pairs trading is based on the simple concept of exploiting price disparity between similar securities, with the expectation of profit when the imbalance is corrected. Asymmetries are detected by monitoring the spread [1] of pairs of securities with a history of co-movement. Profits are generated by going long the relatively undervalued security and shorting the other when the imbalance passes a predetermined threshold. This idea assumes the existence and continuation of an equilibrium relationship, and is exposed to divergence risk [2].

Consider the two simulated stock prices depicted in figure 2.1. Each follow a non-stationary stochastic process, but they never move too far apart, indicating that common factors are determining their growth.

[1]Distance from equilibrium relationship

[2]The risk of the securities not reverting to the equilibrium relationship.

Figure 2.1: Cointegrated Stock Prices Simulated



This plot illustrates two cointegrated securities. The series are simulated using a random walk with an imposed covariance structure.

2.2 Relative Value Model for Equity Pricing

Before modelling the cointegrated residual process, the systematic component must be factored out of the stock return process, with time, t, being measured from 03.01.2005. In accordance with Avellaneda and Lee (2010), returns are quantitatively compared with a set of factors, using an M factor linear model, exclusively based on price data. Thus, stock returns are assumed to follow

$$\frac{dS_i(t)}{S_i(t)} = \alpha_i dt + \sum_{k=1}^{M} \beta_{ik} \frac{dF_k(t)}{F_k(t)} + dX_i(t) \tag{2.1}$$

$$\forall i = 1, ..., N$$

$$\forall k = 1, ..., M$$

Where

$$\sum_{k=1}^{M} \beta_{ik} \frac{dF_k(t)}{F_k(t)}$$

Are the systematic components of the individual stock returns. Following Liew and Roberts (2013), the model employed here is simplified to

$$\frac{dS_i(t)}{S_i(t)} = \alpha_i dt + \beta_{ij}\frac{dF_j(t)}{F_j(t)} + dX_i(t) \tag{2.2}$$

$$\forall i = 1, ..., N$$

$$\forall j = 1, ..., J$$

Consequently, individual returns are assumed to depend on a single factor. F_j is the price of the j^th ETF used to describe the stock's sector. Note that β_{ij} is the factor loading of ETF j on stock i. To arrive at the idiosyncratic component, each security is regressed against its sector ETF, serving as a proxy for the stock's peer group. Hence the idiosyncratic term is given by

$$\alpha_i dt + dX_i(t) \tag{2.3}$$

Hence, the idiosyncratic term consists of a drift term and a residual component. The drift captures company specific qualities, which can make it consistently over- or underperform its industry. $dX_i(t)$ describes the irrational behaviour not captured by these factors or industry, and is what will be modelled using a mean-reverting and stationary stochastic process.

2.3 The Ornstein-Uhlenbeck Process

Results regarding the Ornstein-Uhlenbeck process is presented here, with full derivations in appendix A. As stated in section 2.2, $dX_i(t)$ is the increment of a mean-reverting stationary stochastic process. The chosen model need to incorporate the considerations discussed in section 2.2. Moreover, it must allow for observed anomalies in the relationship between a security and its sector ETF. By this, the increment, $dX_i(t)$ is assumed to satisfy

$$dX_i(t) = \kappa_i\left[\theta_i - X_i(t)\right]dt + \sigma_i dB_i(t), \quad \kappa_i > 0 \tag{2.4}$$

Which is the continuous time representation of a stationary autoregressive model of order one. The objective is to capture the dynamics related to overreactions and other factors driving temporal mispricing. Assuming efficient markets, such anomalies should be short-lived and the process should have a zero unconditional expectation to reflect this. This property is demonstrated in (A.9). Furthermore, the one-day conditional expectation is given by

$$\mathbb{E}\{dX_i(t)|\mathcal{F}_s, s \leq t\} = \kappa_i[\theta_i - X_i(t)]dt \tag{2.5}$$

With a sign depending on $\theta_i - X_i(t)$. This fits with the observation that the spread should mean revert from above if the stock is relatively overvalued compared to the ETF and vice versa. Assuming the parameters of the process to be constant over the chosen period, equation (2.4) admits the solution

$$X(t) = e^{-\kappa t}X(0) + \theta\left[1 - e^{-\kappa t}\right] + \sigma \int_0^t e^{-\kappa(t-s)}dB(s) \tag{2.6}$$

2.4 Signals

The cointegrated residual process aims at two things: Estimating the one-day expected excess return and generating signals- the relative distance between stock and ETF. Following Avellaneda and Lee (2010), the s-score is defined as

$$s_i = \frac{X_i(t) - \theta_i}{\sigma_{eq,i}} \tag{2.7}$$

And it measures the number of standard deviations the cointegrated residual is from its theoretical equilibrium level. By this representation, the effect of the drift term $\alpha_i dt$ is assumed insignificant to the overall process. See figure 2.2 for Statoil's s-scores during 2016.

Figure 2.2: S-score Statoil

This plot depicts equation (2.7) for Statoil over in 2016. The light blue bands are placed at the $\pm 1.5\sigma$ mark.

With a pure pairs trading strategy, the s-score would signal when to initiate long and short positions. Here, it will determine confidence in the one-day expected excess returns within the Black-Litterman model. If the drift component is not negligible, (2.7) must be modified. The drift is accounted for by considering the conditional expected residual return over dt.

$$\begin{aligned}\mathbb{E}\{dX_i(t)|\mathcal{F}_s, s \leq t\} &= \alpha_i dt + \kappa_i \left(\theta_i - X_i(t)\right) dt \\ &= \kappa_i \left(\frac{\alpha_i}{\kappa_i} + \theta_i - X_i(t)\right) dt\end{aligned}$$

This, in combination with (2.7), yields

$$\mathbb{E}\{dX_i(t)|\mathcal{F}_s, s \leq t\} = \kappa_i \left(\frac{\alpha_i}{\kappa_i} - \sigma_{eq,i} s_i\right) dt$$

Which suggests that the modified s-score is given by

$$s_{mod,i} = s_i - \frac{\alpha_i}{\kappa_i \sigma_{eq,i}} \tag{2.8}$$

Incorporating the drift in the signal ensures that if it is positive, the modified s-score is reduced, increasing the threshold before a short signal is created. Thus, it embeds a momentum-strategy.

2.5 Black-Litterman

The Black-Litterman model combines private views with the market view, resulting in blended expected excess returns. This should yield more stable portfolios than the Markowitz framework. The latter is known for being highly sensitive to inputs and is prone to produce extreme portfolios. In their original paper, Black and Litterman started with the mean-variance framework postulated by Markowitz and reverse engineered it. Within the model, optimal portfolio weights are given by

$$\boldsymbol{\omega}^* = (\delta \boldsymbol{\Sigma})^{-1} \boldsymbol{\mu} \tag{2.9}$$

Where

δ: Risk aversion parameter

$\boldsymbol{\Sigma}$: Covariance matrix of market portfolio

$\boldsymbol{\mu}$: Expected excess return vector

Instead of using this to compute the optimal weights, they assumed that the market portfolio, $\boldsymbol{\omega}_{mkt}$, was the equilibrium portfolio. With the market weights being optimal, they backed out the implied returns using

$$\boldsymbol{\Pi} = \delta \boldsymbol{\Sigma} \boldsymbol{\omega}_{mkt} \tag{2.10}$$

This was a substantial shift away from previous techniques. For the equilibrium portfolio to admit the form suggested by Black and Litterman, they used the mean-variance framework and the Capital Asset Pricing Model (CAPM), making the model internally inconsistent. By assuming the CAPM holds, it follows that investors have homogeneous beliefs, which is in clear contradiction with the aim of incorporating private information in the estimation of expected excess returns. Regardless of this tension, the model is popular in the industry.

Now, the model is briefly explained, and its main results presented. Then practical considerations when applying the model in active management is discussed. Derivations are found in appendices B and C.

The Master Formula

The base of the Black-Litterman model is that expected returns are

$$\mathbf{r} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{2.11}$$

Where the objective is to model these parameters. The expected value, $\boldsymbol{\mu}$ is a random vector, assumed to be distributed according to

$$\boldsymbol{\mu} \sim \mathcal{N}(\boldsymbol{\Pi}, \boldsymbol{\Sigma}_\pi) \tag{2.12}$$

Hence, expected market returns are assumed unknown, and estimated with uncertainty, $\boldsymbol{\Sigma}_\pi$, which is defined to be $\boldsymbol{\Sigma}_\pi = \tau\boldsymbol{\Sigma}$, stating that the uncertainty of the mean estimate is proportional to the uncertainty of the returns, with τ representing the scaling factor. Given that both the mean and the returns are stochastic variables, total uncertainty of the prior distribution is equal to $\boldsymbol{\Sigma}_r = \boldsymbol{\Sigma} + \boldsymbol{\Sigma}_\pi$ [3]. Thus, expected returns are modelled as

$$\mathbf{r} \sim \mathcal{N}(\boldsymbol{\Pi}, \boldsymbol{\Sigma}_r) \tag{2.13}$$

The posterior distribution is often referred to as the master formula. Assuming that the investment universe consists of N securities at time t, the following applies

$$\mathbb{E}(\mathbf{r}) = \left[(\tau\boldsymbol{\Sigma})^{-1} + \mathbf{P}^\intercal\boldsymbol{\Omega}^{-1}\mathbf{P}\right]^{-1}\left[(\tau\boldsymbol{\Sigma})^{-1}\boldsymbol{\Pi} + \mathbf{P}^\intercal\boldsymbol{\Omega}^{-1}\mathbf{Q}\right] \tag{2.14}$$

Where

$\boldsymbol{\Pi}_{Nx1}$:	Vector of equilibrium expected excess returns
$\boldsymbol{\Sigma}_{NxN}$:	Covariance matrix of equilibrium expected excess returns
$\mathbf{P}_{kxN}$:	Pick matrix
$\boldsymbol{\Omega}_{kxk}$:	Diagonal covariance matrix of views expressing uncertainty
$\mathbf{Q}_{kx1}$:	Vector of view's expected excess returns
τ_{1x1}:	Scaling parameter
$\mathbb{E}(\mathbf{r})_{Nx1}$:	The posterior returns estimate

[3] This relationship assumes that the uncertainty of the unknown mean is uncorrelated with the variance of the returns

From (2.14), the posterior estimate of expected excess returns is a weighted average between the equilibrium returns and the returns from investor's views (second term). The average is then scaled, so the weights sum to one (first term). This can also be written as

$$\mathbb{E}(\mathbf{r}) = \boldsymbol{\Pi} + \tau\boldsymbol{\Sigma}\mathbf{P}^{\intercal}\left[\mathbf{P}\tau\boldsymbol{\Sigma}\mathbf{P}^{\intercal} + \boldsymbol{\Omega}\right]^{-1}\left[\mathbf{Q} - \mathbf{P}\boldsymbol{\Pi}\right] \tag{2.15}$$

Which illustrates that, in the absence of views, the posterior estimate is equal to the equilibrium returns. The covariance matrix of $\mathbb{E}(\mathbf{r})$ is

$$\boldsymbol{\Sigma}_{pm} = \left[(\tau\boldsymbol{\Sigma})^{-1} + \mathbf{P}^{\intercal}\boldsymbol{\Omega}^{-1}\mathbf{P}\right]^{-1} \tag{2.16}$$

This represents the uncertainty in the posterior mean estimate, and not the variance of the returns (Walters et al., 2014). The variance of the posterior distribution is then the sum of the variance for the returns themselves and the variance of the estimate of the mean, $\boldsymbol{\Sigma}_p = \boldsymbol{\Sigma} + \boldsymbol{\Sigma}_{pm}$.

Parameters

$\boldsymbol{\Pi}$ is the implied expected excess returns from the constituents of the equilibrium portfolio and $\boldsymbol{\Sigma}$ is the covariance matrix. The pick matrix expresses which assets the investor has a view on, where each row represents a view for a total of k rows, and one column for each asset in the trading universe, N in total. Views can either be absolute (Statoil has a 5 % upside) or relative (Statoil will outperform Aker BP by 2 %). For the absolute view, the row in the pick matrix corresponding to that view would have 1 in the column for Statoil. In the second case the row corresponding to the view would have 1 in the column for Statoil and -1 in the column for Aker BP.

Assuming that views are uncorrelated, $\boldsymbol{\Omega}$ is the diagonal matrix capturing uncertainty. Hence, $\forall i \neq j$, $\omega_{ij} = 0$. $\mathbf{Q}$ is the vector with excess return estimates, specified by the investor, while τ is the proportionality factor, linking variance of assets' returns and the unknown mean return.

In this thesis, an alternative to the original model is used, where the unknown mean return, $\boldsymbol{\mu}$, is treated as a point estimate. Alas, this result in $\boldsymbol{\Sigma}_{\pi} = \tau\boldsymbol{\Sigma} = 0$, and that τ disappears. Therefore, the covariance matrix of the returns $\boldsymbol{\Sigma}$, and not $\boldsymbol{\Sigma}_p$ is taken as input in the optimisation routine. Accordingly, the posterior expected returns are defined as

$$\mathbb{E}(\mathbf{r}) = \boldsymbol{\Pi} + \boldsymbol{\Sigma}\mathbf{P}^\intercal \left[\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^\intercal + \boldsymbol{\Omega}\right]^{-1} \left[\mathbf{Q} - \mathbf{P}\boldsymbol{\Pi}\right] \tag{2.17}$$

Bridging Black-Litterman and Active Management

This section discusses the practical considerations of implementing the model. Since it was originally derived under the mean-variance framework, where the investor seeks to maximise expected utility of wealth as a function of wealth level, it does not directly map over to active management, where the objective is to maximise wealth in excess of the benchmark portfolio (Lee, 2000). Because of this discrepancy, naive implementation will result in suboptimal trades and unintentional risk taking. The points emphasised in this section is largely derived from Da Silva et al. (2009) and Lee (2000). Complete results and derivations are found in appendix C.

The following definitions apply

$\boldsymbol{\omega}_B$:	Vector of benchmark portfolio weights
$\boldsymbol{\omega}_a$:	Vector of active portfolio weights
$\boldsymbol{\omega}_{GMVP}$:	Vector of minimum variance portfolio weights
μ_{GMVP}:	Expected excess return minimum variance portfolio
$\boldsymbol{\mu}$:	Expected excess return vector
$\overline{\boldsymbol{\mu}}$:	Expected excess equilibrium return vector
δ:	Total risk aversion parameter
δ_T:	Active risk aversion parameter
$\boldsymbol{\Sigma}$:	Covariance matrix
λ:	Lagrangian multiplier (MV optimisation)
λ_T:	Lagrangian multiplier (Active optimisation)

Given the objective of the investor in the total risk and return framework, along with assumptions outlined in C, expected utility is determined by the portfolio's expected return and variance.

$$\mathbb{E}\{U(W)\} = -e^{-\delta\left(\mu_p - \frac{\delta}{2}\sigma_p^2\right)} \tag{2.18}$$

And is maximised by solving

$$\begin{aligned} \underset{\boldsymbol{\omega}}{\arg\max} \quad & \boldsymbol{\omega}^\intercal \boldsymbol{\mu} - \boldsymbol{\omega}^\intercal \boldsymbol{\Sigma} \boldsymbol{\omega} \\ \text{subject to} \quad & \boldsymbol{\omega}^\intercal \mathbf{1} = 1 \end{aligned} \tag{2.19}$$

Which admits the solution

$$\boldsymbol{\omega}^* = \boldsymbol{\omega}_{GMVP} + \boldsymbol{\omega}_S + \boldsymbol{\omega}_T \tag{2.20}$$

Where

$$\boldsymbol{\omega}_S = \frac{1}{\delta}\frac{\boldsymbol{\Sigma}^{-1}(\overline{\boldsymbol{\mu}}\mathbf{1}^\intercal - \mathbf{1}\overline{\boldsymbol{\mu}}^\intercal)\boldsymbol{\Sigma}^{-1}}{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{1}}\mathbf{1}$$

$$\boldsymbol{\omega}_T = \frac{1}{\delta}\frac{\boldsymbol{\Sigma}^{-1}\left[(\boldsymbol{\mu} - \overline{\boldsymbol{\mu}})\mathbf{1}^\intercal - \mathbf{1}(\boldsymbol{\mu} - \overline{\boldsymbol{\mu}})^\intercal\right]\boldsymbol{\Sigma}^{-1}}{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{1}}\mathbf{1}$$

Hence, in the mean-variance framework, the optimal portfolio consists of three components, the global minimum variance portfolio, a strategic ($\boldsymbol{\omega}_S$) and a tactical ($\boldsymbol{\omega}_T$) component. The strategic part increases the weights for assets with higher equilibrium excess returns, as seen by, $\boldsymbol{\omega}_S$, $\overline{\boldsymbol{\mu}}\mathbf{1}^\intercal - \mathbf{1}\overline{\boldsymbol{\mu}}^\intercal$, which compares the equilibrium returns for all assets, i and j, $i \neq j$. Each matrix element is defined as $\overline{\mu_{ij}} = \overline{\mu_i} - \overline{\mu_j}$. If the sign of the term is positive, the portfolio weight, ω_i is increased compared to its weight in the minimum variance portfolio $\omega_{GMVP,i}$, and vice versa for asset j.

$\boldsymbol{\omega}_{GMVP} + \boldsymbol{\omega}_S$, the strategic mix is known as the benchmark portfolio in active management (Lee, 2000). Given that the objective of an active investment manager is to outperform his benchmark, the expected utility can be expressed in terms of alpha and tracking error (TE) [4].

$$\mathbb{E}\{U(W)\} = -e^{-\delta_T(\alpha - TE^2)} \tag{2.21}$$

Which is maximised by solving

$$\begin{aligned} &\underset{\boldsymbol{\omega}_a}{\arg\max} \quad \boldsymbol{\omega}_a^\intercal\boldsymbol{\mu} - \boldsymbol{\omega}_a^\intercal\boldsymbol{\Sigma}\boldsymbol{\omega}_a \\ &\text{subject to} \quad \boldsymbol{\omega}_a^\intercal\mathbf{1} = 0 \end{aligned} \tag{2.22}$$

The solution can be written as

$$\boldsymbol{\omega}_a^* = \frac{1}{\delta_T}\frac{\boldsymbol{\Sigma}^{-1}\left[\boldsymbol{\mu}\mathbf{1}^\intercal - \mathbf{1}\boldsymbol{\mu}^\intercal\right]\boldsymbol{\Sigma}^{-1}}{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{1}}\mathbf{1} \tag{2.23}$$

[4]TE is the volatility of the alpha.

Comparing this to $\boldsymbol{\omega}_T$ in the mean variance framework, it is apparent why using the Black-Litterman model in an active setting will result in unwanted positions.

$$\boldsymbol{\omega}_T = \frac{1}{\delta}\frac{\boldsymbol{\Sigma}^{-1}\left[(\boldsymbol{\mu}-\overline{\boldsymbol{\mu}})\mathbf{1}^\intercal - \mathbf{1}(\boldsymbol{\mu}-\overline{\boldsymbol{\mu}})^\intercal\right]\boldsymbol{\Sigma}^{-1}}{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{1}}\mathbf{1}$$

$$\boldsymbol{\omega}_a^* = \frac{1}{\delta_T}\frac{\boldsymbol{\Sigma}^{-1}\left[\boldsymbol{\mu}\mathbf{1}^\intercal - \mathbf{1}\boldsymbol{\mu}^\intercal\right]\boldsymbol{\Sigma}^{-1}}{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{1}}\mathbf{1}$$

When maximising utility in the active risk-return domain, a position in asset i is opened if $\mu_i > \mu_j$. In the mean-variance framework the same action only occurs if the difference in expected excess returns between asset i and j is larger than what it is expected to be in equilibrium. This is clearly seen by limiting the scope to two assets, where the j^{th} element of the i^{th} row in

$$(\boldsymbol{\mu}-\overline{\boldsymbol{\mu}})\mathbf{1}^\intercal - \mathbf{1}(\boldsymbol{\mu}-\overline{\boldsymbol{\mu}})^\intercal$$

Is equal to

$$(\mu_i - \mu_j) - (\overline{\mu_i} - \overline{\mu_j})$$

Which goes to show that a tactical position is only opened if the difference in expected excess returns for two assets is larger than what the difference should be in equilibrium, whereas in the active management paradigm, the equilibrium is ignored. Consequently, any difference is perceived as an alpha opportunity.

To illustrate how the differences in objectives affect the optimal portfolio when using the Black-Litterman model as is, assume that the investor does not have any views, and derive the optimal portfolio starting with (C.14)

$$\begin{aligned}\boldsymbol{\omega}_a^* &= \frac{\boldsymbol{\Sigma}^{-1}}{\delta_T}\left(\boldsymbol{\mu} - \mathbf{1}\frac{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{1}}\right)\\ &= \frac{\boldsymbol{\Sigma}^{-1}}{\delta_T}(\boldsymbol{\mu} - \boldsymbol{\omega}_{GMVP}^\intercal\boldsymbol{\mu}\mathbf{1})\end{aligned}$$

Without any views, expected excess returns ($\boldsymbol{\mu}$) should equal equilibrium expected excess returns ($\overline{\boldsymbol{\mu}}$), where equilibrium returns are given by (2.10), and can be substituted into (C.14), giving

$$\boldsymbol{\omega}_a^* = \frac{\delta}{\delta_T}(\boldsymbol{\omega}_B - \boldsymbol{\omega}_{GMVP}) \tag{2.24}$$

No views should not lead to active positions. As shown in (2.24), this is not the case, unless the minimum variance portfolio is the benchmark, which is rarely the case. Thus, implementing the Black-Litterman model without any adjustments results in unwarranted active positions. Essentially the discrepancy arises because of how mean-variance and active management define equilibrium. The former allows for differences between asset i and j in equilibrium, while the latter does not. Consequently, in active management, active trades are initiated to restore equilibrium from what is perceived as a disequilibrium.

Next, the consequences with regards to risk and return are outlined. In appendix C the perceived upsides and hidden downsides of naively using the Black-Litterman model is derived. What it shows is that with no information the active portfolio will yield an information ratio (IR) equal to

$$IR = \delta\sqrt{\sigma_B^2 - \sigma_{GMVP}^2}$$

Clearly, having no information yields a positive information ratio, and the observed upside is proportional to the benchmark's volatility. A riskier benchmark increases the number of active positions and gives the illusion of a positive alpha, but with a net beta exposure to the benchmark, as seen by

$$\beta_a = \frac{\delta}{\delta_T}\left(1 - \frac{\sigma_{GMVP}^2}{\sigma_B^2}\right) > 0$$

And total beta is

$$\beta_p = 1 + \beta_a > 1$$

With the portfolio volatility being

$$\sigma_p = \sigma_B \sqrt{\beta_a \left(\frac{\delta}{\delta_T} + 2 \right) + 1} > \sigma_B$$

In conclusion, the illusionary alpha opportunity embeds unwanted systematic and total portfolio risk. Since the objective of active management cannot be achieved in the mean-variance framework, potential corrections are summarised below.

A potential countermeasure is to switch out the implied excess expected returns that solve the reverse Sharpe Ratio optimisation with a vector of implied expected excess returns that implicitly solve the information ratio optimisation problem (Da Silva et al., 2009). Formally, use a $\boldsymbol{\Pi}$ that solves

$$\begin{aligned} 0 = &\arg\max_{\boldsymbol{\omega}_a} \quad (\boldsymbol{\omega}_a + \boldsymbol{\omega}_B)^\intercal \boldsymbol{\Pi} - \lambda \boldsymbol{\omega}_a^\intercal \boldsymbol{\Sigma} \boldsymbol{\omega}_a \\ &\text{subject to all other constraints} \end{aligned} \tag{2.25}$$

Instead of solving the problem proposed in Black and Litterman (1992)

$$\begin{aligned} 0 = &\arg\max_{\boldsymbol{\omega}_a} \quad (\boldsymbol{\omega}_a + \boldsymbol{\omega}_B)^\intercal \boldsymbol{\Pi} - \lambda (\boldsymbol{\omega}_a + \boldsymbol{\omega}_B)^\intercal \boldsymbol{\Sigma} (\boldsymbol{\omega}_a + \boldsymbol{\omega}_B) \\ &\text{subject to } \textit{no} \text{ constraints} \end{aligned}$$

Choosing $\boldsymbol{\Pi}$ such that all elements are equal, $\Pi_i = \Pi_j,\ \forall i \neq j$ result in the equilibrium return vector becoming constant $\boldsymbol{\Pi} = c$. This in conjunction with the constraint $\boldsymbol{\omega}_a \mathbf{1} = 0$ makes the first term of (2.25) disappear, leaving a tracking error minimisation problem, which can always be solved by taking no active positions ($\boldsymbol{\omega}_a = 0$). Therefore, any constant $\boldsymbol{\Pi}$ will implicitly solve (2.25) regardless of other constraints. The rationale for fixing $\boldsymbol{\Pi}$ is that in an uninformed case, the belief a priori for the expected excess returns should be that they are equal. Since any constant vector satisfies this, setting $\boldsymbol{\Pi} = 0$ is an appropriate choice [5]. Additionally, $\boldsymbol{\omega}_B$ is removed from the objective function. The portfolio is then constructed by solving the optimisation problem to arrive at $\boldsymbol{\omega}_a$ and defining

$$\boldsymbol{\omega} = \boldsymbol{\omega}_a + \boldsymbol{\omega}_B \tag{2.26}$$

[5] Corresponding to saying that without information the assets should yield the risk-free rate.

Chapter 3

Data

The two main data sources are Børsprosjektet and Datastream. From the former, daily data on the constituents of the OBX index was pulled, and daily data on sector ETFs was downloaded from the latter for the time period 01.01.2005 to 31.12.2017.

3.1 Securities Information

From Børsprosjektet, daily data for Oslo Børs (OSE) was downloaded and relevant information from this was pulled by first constructing a semi-annual list of all the constituents in the OBX from 1H 2005 to 2H 2017. This was done using Newsweb, OSE's messaging system for listed companies and financial instruments, searching for messages with the title *"Oslo Børs- Index constituents OBX xH 20xx"* under the category *"Melding fra Oslo Børs"*. These press releases [1] informed of the upcoming index constituents and the representative number of shares for each company. The extracted variables were

Ticker:	Identification symbol
Company name:	Public name
ISIN:	Unique company identifier
Sub industry:	According to Global Industry Classification Standard (GICS)
Shares:	Free float number of shares
Dates:	Start and end date for each constituent's period

For the studied period, there are 67 unique companies that constitutes the OBX index. Knowing each period's composition, the following variables were gathered from the OSE

[1] For 1H 2005 to 2H 2006 the relevant information was downloaded from Datastream.

dataset.

LogReturnAdjGeneric	Logarithmic returns based on close prices and adjusted for dividends and corporate actions[2]
Bid	Unadjusted Bid prices
Offer	Unadjusted Offer prices
Last	Unadjusted Close prices
AdjOpen	Opening prices adjusted for dividends and corporate actions
GICS	Eight digit code describing which sub sector the company belongs to

Also, a dataset on all stock splits and reversals occurring for listed companies between 1980 and 2017 was downloaded from Børsprosjektet to correct for issues outlined in section 3.3

Table 3.1: Descriptive Statistics- OBX Constituents

Statistic	N	Mean	St.Dev	Min	Max	Skewness	Kurtosis	No.Stocks
Energy	2,280	-0.104	0.584	-1.094	0.822	-0.378	43.537	29
Materials	3,258	-0.180	0.635	-0.894	0.644	-1.418	70.630	2
Industrials	2,932	-0.026	0.575	-0.573	0.775	0.396	30.485	8
Consumer Discretionary	2,496	-0.014	0.480	-0.744	0.416	-1.126	54.282	4
Consumer Staples	2,740	0.151	0.410	-0.572	0.268	-0.394	20.490	7
Health Care	1,167	0.319	0.601	-0.559	0.494	0.624	41.740	2
Financials	2,307	0.019	0.405	-0.401	0.283	-0.350	15.580	6
Technology	1,944	0.087	0.515	-1.234	0.342	-4.440	159.867	8
Telecom	3,262	0.085	0.306	-0.300	0.136	-1.156	22.094	1
Utilities	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	0

This table provides descriptive statistics for the constituents of the OBX index between 2005 and 2017, divided into sectors. N denotes the average number of observations within each sector, Mean is the daily mean return for all stocks in the sector annualised, as is the St.Dev. No.Stocks is the number of companies within each sector.

3.2 Exchange Traded Funds Information

From Datastream, daily adjusted [3] prices (P), denominated in EUR for State Street Global Advisors' MSCI Europe GICS sector ETFs were downloaded. These funds track the performance of European large and mid-sized companies in different sectors, according to GICS (SPDR, 2018).

[3]Dividends and other events to ensure comparability over time

Table 3.2: Descriptive Statistics- Exchange Traded Funds

Statistic	N	Mean	St.Dev	Min	Max	Skewness	Kurtosis	Constituents	Return	Mkt.Cap
Energy	3,265	0.050	0.238	-0.097	0.100	-0.190	4.790	17	0.029	76,868.300
Materials	3,265	0.076	0.286	-0.148	0.130	-0.010	8.190	43	0.077	32,304.690
Industrials	3,265	0.076	0.238	-0.142	0.125	-0.110	10.590	90	0.083	24,320.490
Consumer Discretionary	3,262	0.076	0.222	-0.119	0.134	0.040	9.360	70	0.074	28,068.180
Consumer Staples	3,264	0.101	0.159	-0.064	0.061	-0.250	4	42	0.076	78,883.010
Health Care	3,265	0.076	0.175	-0.077	0.078	-0.270	5.180	27	0.045	77,393.410
Financials	3,265	0	0.286	-0.139	0.144	-0.010	9.920	85	-0.002	47,008.870
Technology	3,265	0.050	0.238	-0.131	0.108	-0.320	6.840	18	0.019	42,173.880
Telecom	3,263	0.050	0.190	-0.094	0.085	-0.180	5.370	20	0.031	35,145.740
Utilities	3,264	0.050	0.206	-0.138	0.112	-0.480	10.040	21	0.048	22,468.520

This table displays descriptive statistics for the SPDR MSCI Europe GICS ETFs. N is the number of observations, Mean is the daily mean return annualised, St.Dev is the daily volatility annualised, Constituents is the number of securities in each ETF according to SPDR, Return is the fund's annualised return since inception and MktCap is the average constituent's market capitalisation per end of April 2018.

3.3 Adjustments and Other Considerations

During the data processing, several issues arose. When constructing the OBX constituents list it became clear that companies were added and subtracted between scheduled revisions, making it necessary to adjust the list on a daily basis, resulting in active use of Newsweb to search for information on inter-period changes.

Additionally, some companies performed stock splits or reversals while being a constituent, requiring the adjustment of the number of representative shares. Lastly, the price data had some missing observations, which were padded using the last price forward.

Chapter 4

Methodology

The objective of this chapter is to clearly state how the results presented in section 5 are generated. Each phase of the process is reviewed, emphasising the link between theory and implementation. First, the reconstruction of the OBX index is outlined, before signal estimation is explained. Following this, parameter settings in the Black-Litterman model are discussed, and the optimisation problem is formalised along with practical considerations. Lastly, some issues with backtesting are considered.

4.1 Reconstructing the OBX Index

Given the high frequency, long-short nature of the strategy, selecting a benchmark is important, since the trading universe must consist of liquid securities. Therefore, the OBX index, comprised of the 25 most liquid companies on OSE is chosen. Hence, each constituent's weight must be computed, to obtain $\boldsymbol{\omega}_B$, and is given by

$$\omega_{i,t}^{OBX} = \frac{MC_{i,t}}{\Sigma_{i=1}^{n} MC_{i,t}}, \ \forall i = 1, \ldots, n \tag{4.1}$$

Where $MC_{i,t}$ is the market capitalisation of company i at day t, and is computed according to

$$MC_{i,t} = p_{i,t} q_{i,t} \tag{4.2}$$

$q_{i,t}$ is the number of free float shares each constituent is represented by in the index at any given day, and $p_{i,t}$ is the corresponding price, which is determined by the pricing algorithm given by OSE (2017) [1].

$$p_{i,t} = \begin{cases} bid_{i,t}, & \text{if } bid_{i,t} > trade_{i,t} \\ ask_{i,t}, & \text{if } ask_{i,t} > 0 \text{ and } ask_{i,t} < trade_{i,t} \\ trade_{i,t} & \text{otherwise} \end{cases}$$

Using data on bid, offer and last (trade) prices, a merged dataset with correct prices according to the algorithm was produced. In their methodology, OSE adjust prices for dividend and other corporate actions. Using adjusted price series from Børsprosjektet resulted in wrong index weights, because Børsprosjektet seems to adjust the entire price series for such actions, creating large price discrepancies between the downloaded prices and the prices used by OSE. Thus, unadjusted prices were used, as they were closer to the actual prices. Taking a random sample of days uncovered that the estimated weights do not deviate much from the actual weights [2]. To determine $q_{i,t}$, adjustments outlined in section 3.3 were necessary, and a few occurrences of irregular changes to the index composition were discovered. These are summarised in table E.4.

It is further assumed that companies are tradeable up until the day they leave the index, effectively disregarding the notion of illiquidity in stocks about to exit. In addition to compositional changes, some companies perform stock splits or reversals while being a constituent. This requires the adjustment of the representative number of shares, for the affected companies. In the case of a stock split/reversal at $t_{\Delta q}$, which is $t_{rev1} \leq t_{\Delta q} \leq t_{rev2}$, and the company is a constituent for $t \geq t_{rev2}$, then the number of shares for the next period is assumed to apply from $t_{\Delta q}$ instead of t_{rev2} [3]. If a company is not a constituent in the period after the split/reversal, the split/reversal ratio is used, i.e. if there is a 2:1 split, the number of index shares is multiplied by two. These adjustments are summarised in table E.5.

With daily data on each constituent's representative number of shares and the correct prices according to the pricing algorithm, the weights were computed using equation (4.1).

[1] Note that *trade* is often referred to as the close price and *bid* the offer price

[2] The estimation error seemed to occur at the fourth decimal for the sample.

[3] t_{rev1} and t_{rev2} represents times of index revisions.

4.2 The Mean-Reversion Model

This section provides details on estimating the parameters of the Ornstein-Uhlenbeck process and computation of s-scores. First, stock returns are decomposed in accordance with (2.2), in which each company is matched with a sector ETF [4], and the residuals are estimated from

$$r_{i,t} = \beta_{0,i} + \beta_i r_{ETF_j,t} + \nu_{i,t},\ t = 1, \ldots, 70 \tag{4.3}$$

Returns are chronologically ordered, such that $r_{ETF_j,70}$ is the last observed return. Given the model in (2.2)

$$\alpha_i dt = \beta_{0,i} \Leftrightarrow \alpha_i = \frac{1}{dt}\beta_{0,i} = \frac{1}{\Delta t}\beta_{0,i}$$

Δt corresponds to $\frac{1}{days_{year}}$, and the number of days in a year is set to 252, as is done by Avellaneda and Lee and Liew and Roberts. Accordingly, $\alpha_i = \beta_0 * 252$ is the estimate of permanent differences between security i and its sector. Security returns are logarithmic and based on close prices adjusted for dividends and other corporate actions, following

$$r_{i,t} = \ln\left(\frac{p_{adjclose,i,t}}{p_{adjclose,i,t-1}}\right)$$

Both Avellaneda and Lee and Liew and Roberts use a 60- day regression window on the basis that it incorporates one earnings cycle, which likely reflects natural fluctuations in the share price over the cycle. In this thesis, exponential weighted moving averages (EWMA) is one of two techniques used to estimate the covariance matrices. To include a sufficient amount of history in these matrices, a 70 days estimation window is chosen. For consistency, the same estimation window is used for the stock decomposition and the computation of the s-scores.

The sum of residuals process for security i at time t is defined as

$$X_{i,t} = \Sigma_{k=1}^{70} \nu_{t-k+1}$$

[4]Some companies are reclassified during the period making daily matching a necessity.

This can be seen as a discrete adaptation of the continuous Ornstein-Uhlenbeck process, with parameters: κ_i, θ_i, σ_i and $\sigma_{eq,i}$, which can be estimated by

$$X_{i,t} = a_i + b_i X_{i,t-1} + \zeta_t$$

Compared with equation (2.6), it follows that

$$\begin{aligned} a_i &= \theta_i \left(1 - e^{-\kappa_i \Delta t}\right) \\ b_i &= e^{-\kappa_i \Delta t} \\ \mathbb{V}ar(\zeta_t) &= \sigma_i^2 \frac{1 - e^{-2\kappa_i \Delta t}}{2\kappa_i} \end{aligned}$$

Leading to

$$\begin{aligned} \kappa_i &= -\ln(b) * 252 \\ \theta_i &= \frac{a}{1-b} \\ \sigma_i &= \sqrt{\frac{\mathbb{V}ar(\zeta_t) 2\kappa_i}{1-b^2}} \\ \sigma_{eq,i} &= \sqrt{\frac{\mathbb{V}ar(\zeta_t)}{1-b^2}} \end{aligned}$$

It follows from the use of OLS regression that, $X_{i,t} = X_{i,70} = 0$. Consequently, the s-score is equal to

$$s_i = \frac{-\theta_i}{\sigma_{eq,i}}$$

As opposed to many techniques reviewed in section D.2, the existence of a cointegrated relationship between a security and its sector-ETF is assumed, rather than tested for. To limit uncritical use, securities need to mean revert fast, resulting in signals only being used if the characteristic time-scale for mean reversion (Avellaneda and Lee, 2010)

$$\tau_i = \frac{1}{\kappa_i} \tag{4.4}$$

is lower than $T_1 = \frac{70}{252} = 0.28$. Another approach is to limit the number of stocks to the ones with mean reversion times less than or equal to half a period ($\kappa_i \geq \frac{252}{35} = 7.2$), which implies that b from the AR(1) regression needs to be $0 \leq b \leq 0.97$ [5].

The s-score above assumes that the drift term is negligible. Therefore, a set of scores that considers the drift is calculated. Moreover, given the potential high variance of the θ_i estimator, a centred version is also computed, expressed as

$$\bar{\theta}_i = \frac{a_i}{1 - bi} - \left\langle \frac{a}{1 - b} \right\rangle$$

The brackets denote averaging over all the stocks in the OBX index at time t.

4.3 Parameter Settings in Black-Litterman

Equilibrium/Prior Expected returns

As outlined in section 2.5, prior returns, $\boldsymbol{\Pi}$, are set to zero. Holding a market-neutral portfolio leads to returns stemming from $dX_{i,t}$, where the unconditional expectation is zero, i.e. with no information, the expected net-return is zero [6]. The Ornstein-Uhlenbeck process aligns with this reasoning [7].

Covariance Matrix

Using a 70-days estimation window, the covariance matrices are estimated using an equally weighted as well as the EWMA approach, where $\lambda = 0.94$ in accordance with Morgan et al. (1996). The securities that comprise the matrix at time t are the ones registered with a signal at time t, meaning that a stock might be a constituent without being in the set where active positions might be taken.

Pick Matrix

Here, views are expressed on expected returns in absolute terms for all securities with an estimated s-score. Therefore, the pick matrix equals I_{NxN} at time t.

Views

The views vector, $\mathbf{Q}$ is estimated according to (2.5), where the elements are the one-day expected excess returns for assets with an estimated s-score at time t.

[5] $-252\ln(b) \geq 7.2 \Leftrightarrow b \leq 0.97$

[6] Assuming $\alpha_i = 0$

[7] Given that its unconditional mean is zero.

Confidence in Views

He and Litterman (1999) proposes to quantify confidence, $\boldsymbol{\Omega}$, according to

$$\boldsymbol{\Omega} = \mathrm{diag}(\mathbf{P}^\intercal(\tau\boldsymbol{\Sigma})\mathbf{P})$$

Relying on Liew and Roberts, confidence is quantified by exploiting the information embedded in the s-scores regarding the mean-reversion state of nature. A larger s-score indicates a larger deviation from equilibrium. Consequently, the likelihood of mean-reversion should increase confidence in the expected excess return estimate. Therefore, absolute s-scores are daily grouped into quintiles, and confidence values, $[0.1, 0.01, 0.001, 0.0001, 0.00001]$, are assigned based on these, where a smaller number corresponds to greater confidence.

Posterior Distribution of Expected Returns

Conditioned on the parameter settings, the posterior returns vector is given by

$$\mathbb{E}(\mathbf{r}) = \mathbf{\Sigma}\mathbf{P}^\intercal\left[\mathbf{P}\mathbf{\Sigma}\mathbf{P}^\intercal + \mathbf{\Omega}\right]^{-1}\mathbf{Q} \tag{4.5}$$

4.4 Portfolio Optimisation

From section 2.5, the optimisation problem becomes

$$\underset{\boldsymbol{\omega}_a}{\arg\max} \quad \boldsymbol{\omega}_a^\intercal \mathbb{E}(\mathbf{r}) - \lambda \boldsymbol{\omega}_a^\intercal \boldsymbol{\Sigma} \boldsymbol{\omega}_a$$

$$\text{subject to all other constraints}$$

The constraints follow Liew and Roberts, and are

Beta neutrality within each industry

$$\boldsymbol{\beta}_{\mathbb{S}_j}^\intercal \boldsymbol{\omega}_a = 0,\ \forall j \text{ where } \mathbb{S}_j \triangleq \{i : i \in \text{sector } j\}$$

All active positions must sum to zero

$$\boldsymbol{\omega}_a^\intercal \mathbf{1} = 0$$

When shorting, there must be constraints on how much leverage is allowed for

$$1 \leq (\boldsymbol{\omega}_a + \mathbf{u})^\intercal \mathbf{1} \leq \Lambda \tag{4.6}$$

$$1 \leq \mathbf{u}^\intercal \mathbf{1} \leq \Lambda \tag{4.7}$$

$$u_i \geq 0,\ \forall i \tag{4.8}$$

Furthermore, it is not allowed to short (go long) securities expected to mean revert from below (above). Hence

$$\omega_{a,i} \geq 0,\ \forall i \in \text{BTO},\ \text{where BTO} \triangleq \{i : s_{i,t} < 0\}$$

$$\omega_{a,i} \leq 0,\ \forall i \in \text{STO},\ \text{where STO} \triangleq \{i : s_{i,t} > 0\}$$

Where BTO denotes the set of securities in the universe at time t that are relatively undervalued and expected to mean-revert from below. STO denotes the set of securities that are relatively overvalued and expected to mean-revert from above.

Demanding true equality of the beta neutrality constraint rarely leads to a solution, so the sector exposure has a slack of ± 0.001. Leverage is given by

$$\Lambda = \frac{\text{Long Market Value} + |\text{Short Market Value}|}{\text{Equity}}$$

And set to two, resulting in a NOK-neutral long-short active portfolio.

During the financial crisis, there were instances where all s-scores were negative, giving no valid solution to the optimisation problem, i.e. exploiting mean-reversion is not possible. In these instances all active weights were set to zero, making the portfolio equal to the benchmark.

4.5 Backtesting

It is easy to (unintentionally) cheat when conducting a backtest, and report superior results, which are not reproducible when managing real capital. A common mistake is to trade on the same returns used to generate the signal. $r_t = \frac{p_t}{p_{t-1}} - 1$ is the return form

$t-1$ to t. If r_t generates the signal and profits are made using r_{t+1}, then the implicit assumption is that the strategy can form a signal and trade on p_t, since $r_{t+1} = \frac{p_{t+1}}{p_t}$.

Thus, signals are generated in the afternoon using day t's close prices and trading is done the following morning, using the previous afternoon's generated signals and the opening prices at time $t + \frac{1}{2}$, before profits are made at time $t + \frac{3}{2}$. Given the 70-days estimation window, the first signal is created on 20.04.2005 from returns between 10.01.2005 and 20.04.2005, with the portfolio being initiated using adjusted opening prices on 21.04.2005. 28.12.2017 is the last day where signals are generated. The last rebalancing occurs using the following morning's opening prices, while that day's returns are given by

$$r_{29.12.2017} = \frac{p^{adjclose}_{29.12.2017}}{p^{adjopen}_{29.12.2017}} - 1$$

Even though this is a high frequency strategy, where trading costs are important to consider, a simple approach is taken. To not underestimate these unknowns a total cost of 50 basis points (bp) is used and incorporates trading costs and slippage. Portfolio equity at time t is assumed to be given by the following P&L [8]

$$E_{t+\Delta t} = E_t + E_t r_f \Delta t + \Sigma_{i=1}^{N} C_{it} r_{it} - \left(\Sigma_{i=1}^{N} C_{it}\right) r_f \Delta t \tag{4.9}$$

$$- \Sigma_{i=1}^{N} |C_{i(t+\Delta t)} - C_{it}| \varepsilon \tag{4.10}$$

Where, E_t is the equity at time t, C_{it} is the NOK amount invested in stock i at time t, r_{it} is the return on stock i over the period $(t, t + \Delta t)$ [9]. r_f is the interest rate. For simplicity, long rates are assumed equal to short rates, and set equal to the risk free rate [10]. ε is the combined one way term incorporating transaction costs and slippage.

4.6 Measuring Performance

Most of the performance measures reported in section 5 are computed using standard techniques, but for transparency an overview is provided. Among the return statistics, holding period return is the annualised return and is calculated following

[8] It is the same as Avellaneda and Lee uses.

[9] Effectively from open at day t to open at time $t + 1$, i.e. $\Delta t = \frac{1}{252}$

[10] Assumed to be zero.

$$HPR \text{ p.a.} = \left[1 + \frac{Equity_T - Equity_{t_0}}{Equity_{t_0}}\right]^{\frac{1}{m}} - 1$$

Where m is the number of years. The yearly mean and median is simply the mean and median daily returns annualised by

$$Mean/Medianreturn \text{ p.a.} = (1 + Mean/Mediandailyreturn)^n - 1$$

Where n is the number of compounding days. Annual volatility is the daily volatility of the portfolio's returns scaled up by the square root of the number of days in a year [11]. Skewness is measured using the sample estimator, given by

$$Ske\hat{w}ness = \frac{\frac{1}{n}\Sigma_{i=1}^{n}(r_i - \bar{r})^3}{\left[\frac{1}{n-1}\Sigma_{i=1}^{n}(r_i - \bar{r})^2\right]^{\frac{3}{2}}}$$

The reported kurtosis, measuring the shape of the returns' probability distribution, is in excess of the normal distribution. Estimation is done in accordance with

$$Kur\hat{t}osis = \frac{\frac{1}{n}\Sigma_{i=1}^{n}(r_i - \bar{r})^4}{\left[\frac{1}{n}\Sigma_{i=1}^{n}(r_i - \bar{r})2\right]^{2}} - 3$$

Maximum drawdown is measured as any deviation below the maximum cumulative return at time t, with the size being measured in percentage of that maximum cumulative return (Bacon, 2011, p 88). Alpha is computed as the difference between strategy and benchmark annualised returns, with tracking error being the volatility of the excess return, computed by

$$TrackingError = \sqrt{\Sigma \frac{(r_{strategy} - r_{benchmark})^2}{\sigma_{daily}\sqrt{252}}}$$

Using alpha and tracking error, information ratio is the ratio between them. Converting Sharpe ratios from daily to yearly frequency is often done by scaling up by the square root of the number of days in a year, which is valid if returns are IID (Lo, 2002). Correct

[11] Equal to 252 trading days.

aggregation when returns are not IID involves adjusting for autocorrelation in returns, assuming they follow an AR(1) process, the time conversion is consequently given by

$$SR_{yearly} = SR_{daily}\sqrt{n}\left[1 + \frac{2\rho}{1-\rho}\left(1 - \frac{1-\rho^n}{n(1-\rho)}\right)\right]^{-\frac{1}{2}}$$

Where ρ is the autocorrelation coefficient, and n is the number of days in a year. For more details, see Lo (2002).

4.7 Testing Strategy Returns

To test if the strategy significantly outperform the benchmark, a martingale test, the compound excess return test (CERT), put forward in Foster and Young (2010), is applied. This test is strict, saying that to believe that the strategy is superior, it must outperform the benchmark by a factor of 20 [12], because the strategy's historical returns may have come from a process with positive probability of loosing everything [13]. The strategy's return generating process is assumed to be a black-box, as it was proposed as a method of separating skilled from non-skilled fund managers. The test makes no assumptions regarding the parametric distribution of returns or serial dependence, it is easy to compute and it corrects for unobserved tail risk, hence it is appealing.

Based on Foster and Young (2010) the benchmark portfolio starts at one, and at time t it has grown by a factor, $(1 + r_{t,a})$. Given a zero daily risk free-rate, the ratio

$$A_t = \frac{1 + r_{t,a}}{1 + r_{t,f}} = 1 + r_{t,a} \geq 0$$

is the multiplicative excess return from the benchmark in relation to the risk-free rate. For the strategy, the ratio is

$$B_t = \frac{1 + r_{t,s}}{1 + r_{t,f}} = 1 + r_{t,s} \geq 0$$

Next, by setting $B_t \triangleq M_t A_t$, the factor M_t shows by how much the strategy over- or underperform the benchmark in period t. $M_t \geq 0$, is a non negative martingale. An alpha

[12] Given a significance level of 5%

[13] Black Swan exposure.

of zero translates into the set of martingales, M_t, where the conditional expectation is one, as opposed to a set of martingales with conditional expectation larger than one being characteristic of strategies with a positive alpha Foster and Young (2010). As such, the null hypothesis of zero alpha becomes

$$H_0 : \forall t, \ \mathbb{E}[M_t | A_{t-1}, \ldots, A_1, M_{t-1}, \ldots, M_1] = 1 \tag{4.11}$$

For each t, $1 \leq t \leq T$, define

$$C_t \triangleq \prod_{1 \leq s \leq t} M_s \tag{4.12}$$

To be the compound excess return of the strategy to the benchmark up to time t. The probability that this series was generated by a strategy unable beat the market is at most $\min_{1 \leq t \leq T} \frac{1}{C_t}$. Here, the significance level is set to 5%.

Chapter 5

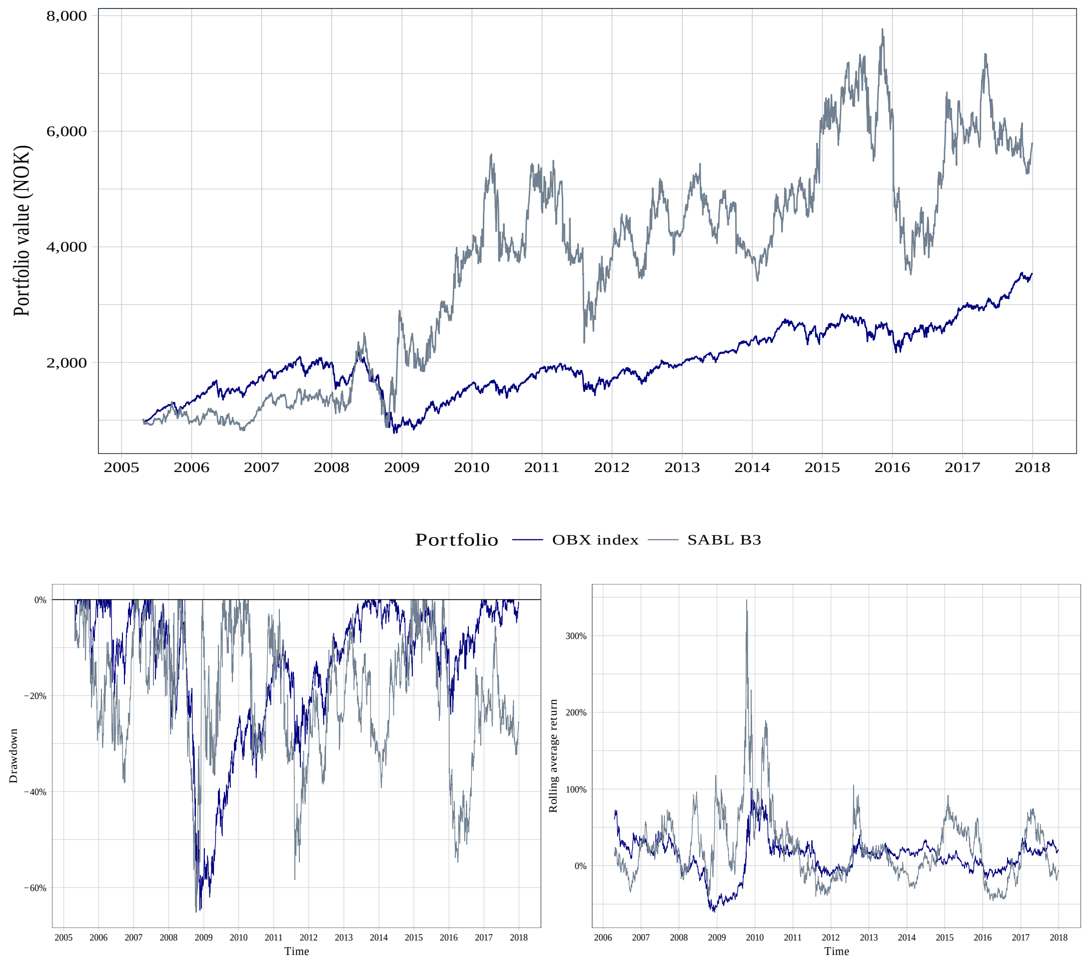
Results

Many trading strategies look superior in a backtest, but fails out of sample. Therefore, relying only on an in-sample test is not adequate to unequivocally establish profitability. Despite this, only the "would-have-been" historical performance is considered here. In this section, performance results for the best performing strategy variation is presented with a discussion on sources of profitability and the statistical significance of the return generating process is treated. For a discussion of results regarding mean-reversion for the constituents and gross performance of all portfolios (benchmark, reference and strategy variations), see appendix E

Out of the eighth strategy variations tested [1], the combination of EWMA covariance, drift and non-centred means (SABL B3) performed best based on total return and risk adjusted performance measures. For a full discussion, see section E.2. Thus, in this section, its results after transaction costs are presented and compared to the benchmark portfolio. Performance is broken down by year and sector. The latter entails forming portfolios that are limited to active bets within each sector and mapping their theoretical returns, to see which sector(s) are responsible for value creation. Furthermore, drawdowns, rolling returns and statistical significance are discussed. Transaction costs are assumed to be constant. Although unrealistic, the estimate is conservative, to limit the overestimation of results after costs. Potential improvements and remedies are discussed in chapter 6.

[1] For a description of all variations, see table E.2

Figure 5.1: Comparing SABL With the OBX Index



The top panel plots the cumulative returns of the benchmark portfolio and the SABL B3 portfolio from 2005 to 2017. The lower left panel shows the drawdowns for both portfolios, while the lower right panel plots the one-year rolling average return for the same portfolios. All plots are after transaction costs.

It is apparent that during the strong markets of 2005 and 2006, the strategy took some unprofitable positions, growing slower than the benchmark. By mid-2008 the SABL B3 and the benchmark were close in value, before experiencing the same drawdown. The largest discrepancy occurs in the last quarter of 2008, when the market saw some extreme intraday movements. The strategy portfolio had daily movements of ±20% and saw 12 days of movements of more than 10%, where the market only saw one. This resulted in the SABL B3 neutralising the entire downturn, ending up with positive value creation for the year.

The market was better in 2009, returning 60% while the strategy returned 52%. Where the large increase in value in 2008 was due to some extreme movements, the strong performance in 2009 seems to be caused by steady compounding, as the strategy only saw eight ±10% intraday moves. Furthermore, post 2009, the strategy and the benchmark share a common trend, with the strategy exhibiting higher volatility, as is also seen by the lower right panel, in which the one-year rolling average return in 2009-2010 far exceeds the benchmark, only to fluctuate around it post 2010.

In the lower left panel, the relatively steep increase in SABL B3's value in 2009 compared to the benchmark is visible, as its drawdowns are consistently lower. The market's drawdowns in 2012-2014 and 2017 are much lower than for the strategy, explaining the strategy's relatively poor performance these years, also seen from table 5.1.

Table 5.1: Performance SABL B3 and OBX

Year	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
SABL B3													
HPR	-0.008	0.252	0.151	0.761	0.523	0.183	-0.176	0.133	-0.122	0.656	0.056	-0.008	−0.059
Yearly mean	0.067	0.386	0.248	1.857	0.877	0.344	0.012	0.209	-0.082	0.764	0.149	0.118	−0.025
Ann.Volatility	0.389	0.447	0.397	0.989	0.644	0.503	0.636	0.358	0.301	0.340	0.411	0.486	0.271
Skewness	-0.250	-0.255	-0.514	0.338	0.133	-0.290	-0.639	-0.089	-1.456	0.779	0.388	-0.840	0.421
Kurtosis	0.558	3.830	3.243	2.195	2.079	3.110	3.713	0.343	6.697	3.321	1.429	4.589	2.420
Ann.SR	-0.105	0.567	0.427	0.749	0.852	0.362	-0.462	0.384	-0.576	1.658	0.110	-0.073	-0.348
Alpha	-0.439	-0.033	0.126	2.349	-0.002	0.140	0.077	0.021	-0.315	0.685	0.092	-0.079	-0.245
Tracking Error	0.360	0.405	0.394	0.934	0.623	0.462	0.531	0.328	0.283	0.320	0.374	0.467	0.263
Information Ratio	-1.220	-0.082	0.321	2.514	-0.003	0.303	0.144	0.063	-1.114	2.139	0.247	-0.169	-0.934
MaxDD	0.294	0.322	0.201	0.652	0.313	0.336	0.575	0.256	0.313	0.134	0.252	0.450	0.284
OBX													
HPR	0.317	0.336	0.121	-0.529	0.603	0.152	-0.115	0.134	0.199	0.043	0.024	0.164	0.188
Yearly mean	0.506	0.382	0.149	-0.459	0.717	0.186	-0.082	0.155	0.210	0.056	0.043	0.192	0.198
Ann.Volatility	0.186	0.250	0.214	0.531	0.361	0.241	0.271	0.187	0.115	0.156	0.190	0.216	0.117
Skewness	-0.379	-0.260	-0.211	-0.167	-0.211	0.072	-0.194	-0.146	-0.278	0.369	-0.025	-0.088	-0.207
Kurtosis	2.186	3.139	0.016	1.902	0.273	1.588	1.041	0.612	0.711	2.828	1.481	1.294	0.091
Ann.SR	2.464	1.474	0.530	-1.703	1.747	0.634	-0.579	0.727	1.800	0.153	0.021	0.759	1.778
MaxDD	0.157	0.202	0.166	0.648	0.211	0.188	0.281	0.152	0.081	0.161	0.171	0.152	0.057

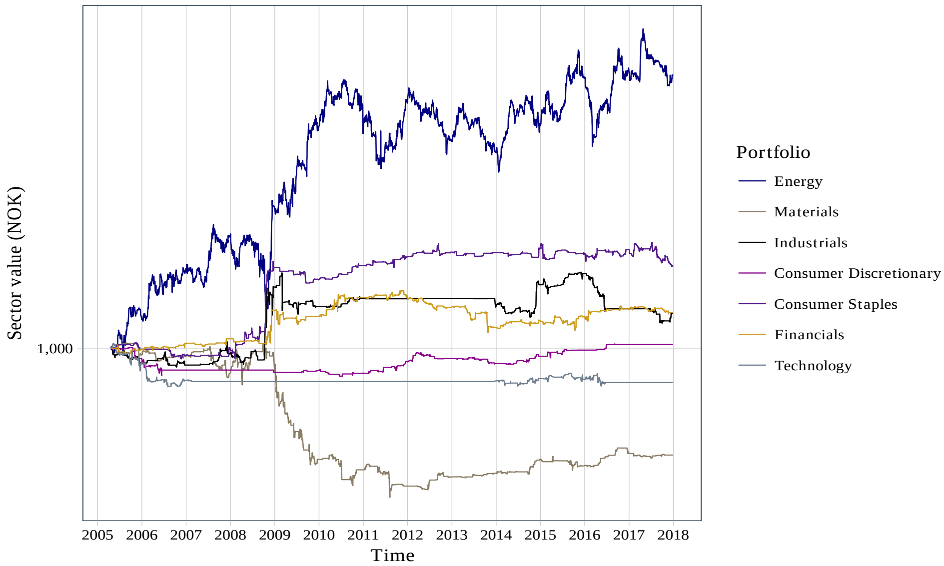
This table breaks down the performance of the overall best strategy and the benchmark by year. Strategy results are reported after costs.

In table 5.1 it is shown how the strategy yields positive returns during the financial crisis while the OBX falls by 53%, and how it experiences five years of negative returns against two years for the benchmark. Resulting from the crash of the oil price in 2014, the benchmark, with a heavy concentration of Energy companies, yielded 4.3%, meanwhile the strategy returned 66%, indicating that the expected return estimates had high accuracy.

Given that two out the top three years for the strategy, measured by Sharpe Ratio, is two of the most volatile years for the market, and two of the worst years are two of the market's calmest years, it might be that the strategy performs best in high volatility regimes. Despite this, 2011 being a volatile year, was not lucrative for the strategy, returning -17% and 2014 being a low-volatility year earned high returns. Risk adjusted, the strategy outperforms on Sharpe ratio four years, delivers top quartile information ratios in two (Grinold and Kahn, 2000, p.114) and returns positive alpha in six out of 13 years. In other words, it does not consistently outperform the benchmark on neither a total return or a risk adjusted basis.

A decomposition of portfolio performance by sector can uncover which factors drive it, and figure 5.2 plots the performance of the sector portfolios.

Figure 5.2: Return Decomposition SABL B3



This panel plots the cumulative returns for different sectors, i.e. the performance of sector portfolios. Each one is constructed by only taking the active bets in that sector. Performance is reported before costs. Note that the y axis is log scale with a base of 10.

The dominance of Energy companies in the trading universe is further emphasised here, showing how this sector is the largest driver for overall portfolio value, although the large returns of late 2008 is also due to Financials, Consumer Staples and Industrials. Supplementing the insights from table 5.1, the high returns 2014 is mainly attributed to the Energy sector, its portfolio value increasing sharply at the beginning of the year. Additionally, Industrials and Consumer Staples increased towards the end of the year, while all other sectors were flat. The poor performance in 2011 is also mostly driven by the Energy sector, as well as a decline in the Materials portfolio.

One of the most profitable sectors in Liew and Roberts (2013) was Technology, which is one of the least profitable here, mainly because the index consists of few Technology companies. From 2005 to 2007, only three index constituents are in the sector, two of them from the Tandberg system. Following this, the number of relevant companies in the OBX is too low to take positions, until 2014. This lack of sector representatives seems to be a general problem for all industries except Energy. Additionally, the cross-section of companies within each sector being in the index at the same time is narrow. Effectively, at any given time, there is not enough diversity in the benchmark for positions to be taken.

Furthermore, as shown in table 3.2 the Utilities sector has no representatives, Telenor is the only Telecom company and in Health Care there are two companies, Algeta and Nordic Nanovector, which are not in the index at the same time. Consequently, no active trades are made in these, effectively reducing the sector sample from ten to seven. The remaining ones have a total of 36 companies, 29 of them being Energy companies. To highlight the extent of this, active positions are decomposed, comparing trading signals with active positions, to see how effectively strong signals are utilised in each industry.

Table 5.2: Utilization of Strong Signals

Sector	Energy	Materials	Industrials	Consumer Discretionary	Consumer Staples	Financials	Technology
Number of strong signals	6,555	824	1,302	1,018	1,604	1,459	508
Number of active positions	19,838	1,614	1,991	1,079	2,398	2,735	1,134
Number of strong trades	4,034	339	431	151	471	551	245
Trade share	0.203	0.210	0.216	0.140	0.196	0.201	0.216
Signal share	0.615	0.411	0.331	0.148	0.294	0.378	0.482

This table shows if and how strong signals end up in active trades. Number of strong signals is the sum of signals within each sector larger than 1.5 in absolute value, number of strong trades is the sum of all the strong signals that actually results in an active position being taken. Trade share is the ratio between the number of strong trades and the total number of active trades. Signal share is the ratio between the number of strong trades and the total number of strong signals.

Table 5.2 shows how the number of positions and strong signals in the Energy sector is x1.8 and x0.97 the number of active positions and strong signals in all other industries combined. In the Energy sector, 62% of all strong signals culminate in an active position, while utilisation drops below 50% for the others, demonstrating how high concentration of Energy companies leads to few stock pairs in the other sectors. This, in combination with the beta-neutrality and long-short constraints, likely reduces the utilisation of strong signals, since a long (short) position within one sector requires the opposite position in another company in the same sector, and a small set of potential matches means fewer outcomes where the opposite position is feasible. Consequently, it could be interesting to see if a broader universe or less restrictive constraints increase signal utilisation.

The strategy increases the investor's investment by a factor, M_t larger than the benchmark, so the statistical significance of this factor is tested [2], to see if alpha is the product of a superior trading strategy or if it is a product of randomness. The null hypothesis is given by equation (4.11).

Figure 5.3: Multiplicative Excess Strategy Return to Benchmark

This panel plots the ratio of cumulative returns of the strategy to the benchmark, C_t as given by equation (4.12), which represents the multiplicative excess return of the strategy relative to the OBX.

Based on this, the probability that the strategy's return series cannot beat the market is $\min_{1 \leq t \leq T} \frac{1}{C_t} = 0.29$. Hence, there is not enough evidence to confidently state that the strategy is superior.

In conclusion, there have been years where the combination of statistical arbitrage and the Black-Litterman model has extracted positive mean-reversion alpha before and after transaction costs, with the SABL B3 portfolio delivering a total return in excess of the benchmark, although not always large enough to compensate for the increased risk. In addition to lacking statistical evidence in support of a superior strategy hypothesis, the dominant Energy sector, along with optimisation constraints and low variability in the composition of constituents, results in low utilisation of strong signals in other sectors. This effectively caps the possibility for profitable trades outside the Energy industry, hinging its success on that sector. Hence, while it seems to have potential, it might benefit from a larger trading universe with less skew towards a single industry, meaning it is probably not a good fit for the Norwegian stock market.

[2]See section 4.7.

Chapter 6

Summary and Conclusion

6.1 Limitations and Areas of Further Research

While the results provide interesting insights, much can be improved or expanded on. Thus, the aim here is to briefly outline some limitations and aspects that could be interesting to investigate further.

Costs

Using constant transaction costs is too simplistic for a real-world representation, given dynamic markets, where liquidity is an important driver, which disappears at the first sign of trouble. Therefore, assuming constant costs during the financial crisis is unrealistic, and true costs would likely reduce the returns of 2008 substantially. Moreover, the role of slippage is ignored, assuming that 50 bp will cover it. Effectively, what is a real challenge for any sizeable fund is ignored. As such, not addressing slippage directly likely leads to an underestimation of the strategy's real market impact.

In addition to the points mentioned above, transaction costs could also be addressed directly in the optimisation routine, by introducing a penalty term, restricting the number of trades taken. Another approach would be to obtain exposure to the benchmark by holding an ETF and taking the active positions by trading options and/or swaptions. This was briefly mentioned in Dahlquist and Harvey (2001).

Following the profit & loss equation given in section 4, stated by Avellaneda and Lee, the long and short rates are assumed equal and equal to the risk-free rate (which is assumed to be zero), effectively saying that the investor can short for free. This is clearly not realistic.

Essentially, there are two areas of improvement within the cost aspect. One is to use dynamic trading costs and more realistic long and short rates, the other is to actively manage them. The latter either directly in the optimisation routine or indirectly, using derivatives.

Estimation

Strategy implementation required the estimation of a large number of parameters, permitting a wide range of techniques, but given the time limit, some interesting alternatives were left out. The residual process estimation was done using a single factor model consisting of European sector ETFs as the explanatory variable within the OLS framework, with a 70-days estimation window. Alternatively, a multifactor model or PCA could be applied. Moving away from OLS, Kalman filter shows potential according to Chen et al., and it would be interesting to see how it performs in conjunction with the Black-Litterman model. Both Avellaneda and Lee and Liew and Roberts state that the 60-days window they use may not be the optimal, and that such a choice should be carefully considered. Using 70 days was the result of the trade-off between including enough information into the EWMA covariance matrices and not taking too long for a stock to be included in the trading universe.

The covariance matrix is a central input variable in the Black-Litterman model and the optimisation routine, and was here estimated using two standard techniques. To improve performance, one could explore the effect of using volatility models from the GARCH class. Within the literature on the Black-Litterman model there is little coverage on how to quantify view certainty, with Liew and Roberts being one of the few to provide a concrete solution. Expanding on their approach it could be interesting to see how other functional forms would alter the results, given how it affects the weight of the views in the blending of the prior and conditional distributions.

Furthermore, both papers [1] introduce volume weighted average returns when estimating factor exposure, finding this to improve performance, meaning such an implementation could be beneficial. The argument for this is, according to Avellaneda and Lee, to discourage open-to-short signals for stocks that rally on high trading volume and open-to-buy signals that form on stocks that fall on high volumes.

Survivorship bias is not present as the trading universe is coherent with the historical

[1] Avellaneda and Lee (2010) and Liew and Roberts (2013)

composition of the index, but data mining is an issue, since the strategy is implemented using different estimation techniques. Thus, randomness being the profit driver is a possibility. Hence, the strict CERT test was used as a remedy, which showed that it could not be excluded as the source of performance. Despite this, the variations in the strategy were decided upon before downloading and looking at the data, and not fit to it, somewhat reducing the bias.

6.2 Conclusion

Applying a strategy that uses statistical arbitrage as a source of views in the Black-Litterman model, this thesis has investigated its performance on the constituents of the OBX index, and is the first to do so. The key finding is that six of eight strategy variations yields annualised holding period returns above benchmark before costs with positive information ratios. After costs, the best version still outperforms, but the statistical evidence for a superior strategy hypothesis is lacking. As such one cannot conclude that the performance is not a product of randomness.

Yearly performance shows how profits are mainly generated during 2008, 2009 and 2014. This, in combination with the years of weakest performance being years of low volatility, indicates that the strategy performs better in high volatility regimes. Upon closer investigation, the positive net return of 2008 was largely due to a few extreme days in the last quarter, while the high return in 2009 was the product of steady compounding. A structural problem with Norwegian equities is the large concentration of companies in the Energy sector, resulting in low variability within the other sectors represented in the index. Along with optimisation constraints, this led to a minor percentage share of strong trading signals being transformed into active portfolio positions. Consequently, the possibility for profitable trades outside the Energy industry is capped, hinging the strategy's success on the Energy sector. On that notion, it became clear that this strategy may be more suited for a larger market with a greater number of liquid stocks more evenly distributed across different industries.

From the estimation perspective, the main takeaway was that, in conflict with others, using a shrinkage estimator for the means of the Ornstein-Uhlenbeck processes underperformed the standard estimator. Seemingly, centred means reduce the size of the expected returns and confidences, producing active positions that does not benefit fully from the high volatility and erratic market regimes of 2008 and 2009. Consequently, the strategy may better serve as a specialised strategy to investment managers searching to profit from high volatility regimes than on a stand-alone basis.

Bibliography

Andrade, S., Di Pietro, V. and Seasholes, M. (2005), 'Understanding the profitability of pairs trading', *Unpublished working paper, UC Berkeley, Northwestern University* .

Arnott, R. D. and Fabozzi, F. J. (1988), *Asset allocation: A handbook of portfolio policies, strategies & tactics*, Probus Professional Pub.

ASA, O. B. (2017), 'Index methodology- equities', `https://www.oslobors.no/Oslo-Boers/Regelverk/Hoeringer/Consultation-Changes-to-the-Index-Methodology-Equities`. Online accessed: 07-May-2018.

Avellaneda, M. and Lee, J.-H. (2010), 'Statistical arbitrage in the us equities market', *Quantitative Finance* **10**(7), 761–782.

Bacon, C. R. (2011), *Practical portfolio performance measurement and attribution*, Vol. 568, John Wiley & Sons.

Bevan, A. and Winkelmann, K. (1998), 'Using the black-litterman global asset allocation model: three years of practical experience', *Fixed Income Research* pp. 1–19.

Black, F. and Litterman, R. (1990), Asset allocation: combining investor views with market equilibrium, Technical report, Discussion paper, Goldman, Sachs & Co.

Black, F. and Litterman, R. (1992), 'Global portfolio optimization', *Financial analysts journal* **48**(5), 28–43.

Bondt, W. F. and Thaler, R. (1985), 'Does the stock market overreact?', *The Journal of finance* **40**(3), 793–805.

Brinson, G. P., Hood, L. R. and Beebower, G. L. (1995), 'Determinants of portfolio performance', *Financial Analysts Journal* **51**(1), 133–138.

Chan, K. (1988), 'On the contrarian investment strategy', *journal of Business* pp. 147–163.

Chen, Y., Ren, W. and Lu, X. (n.d.), 'Machine learning in pairs trading strategies'.

Chordia, T. and Shivakumar, L. (2002), 'Momentum, business cycle, and time-varying expected returns', *The Journal of Finance* **57**(2), 985–1019.

Da Silva, A. S., Lee, W. and Pornrojnangkool, B. (2009), 'The black-litterman model for active portfolio management', *Journal of portfolio management* **35**(2), 61.

Dahlquist, M. and Harvey, C. (2001), 'Global tactical asset allocation'. Available at SSRN: https://ssrn.com/abstract=795376 or http://dx.doi.org/10.2139/ssrn.795376.

Do, B. and Faff, R. (2010), 'Does simple pairs trading still work?', *Financial Analysts Journal* **66**(4), 83–95.

Do, B. and Faff, R. (2012), 'Are pairs trading profits robust to trading costs?', *Journal of Financial Research* **35**(2), 261–287.

Elliott, R. J., Van Der Hoek*, J. and Malcolm, W. P. (2005), 'Pairs trading', *Quantitative Finance* **5**(3), 271–276.

Engelberg, J., Gao, P. and Jagannathan, R. (2008), 'An anatomy of pairs trading: the role of idiosyncratic news, common information and liquidity'.

Engle, R. F. and Granger, C. W. (1987), 'Co-integration and error correction: representation, estimation, and testing', *Econometrica: journal of the Econometric Society* pp. 251–276.

Foster, D. P. and Young, H. P. (2010), 'Gaming performance fees by portfolio managers', *The Quarterly Journal of Economics* **125**(4), 1435–1458.

Fusai, G. and Meucci, A. (2003), 'Assessing views', *Risk Magazine* pp. S18–S21.

Galenko, A., Popova, E. and Popova, I. (2012), 'Trading in the presence of cointegration', *The Journal of Alternative Investments* **15**(1), 85.

Gatev, E., Goetzmann, W. N. and Rouwenhorst, K. G. (2006), 'Pairs trading: Performance of a relative-value arbitrage rule', *The Review of Financial Studies* **19**(3), 797–827.

Grinold, R. C. and Kahn, R. N. (2000), 'Active portfolio management'.

He, G. and Litterman, R. (1999), 'The intuition behind black-litterman model portfolios'.

Herold, U. (2003), 'Portfolio construction with qualitative forecasts', *The Journal of Portfolio Management* **30**(1), 61–72.

Idzorek, T. (2007), A step-by-step guide to the black-litterman model: Incorporating user-specified confidence levels, *in* 'Forecasting expected returns in the financial markets', Elsevier, pp. 17–38.

Jacobs, H. and Weber, M. (2013), 'Losing sight of the trees for the forest', *Attention Allocation and Anomaliesm* .

Jegadeesh, N. (1990), 'Evidence of predictable behavior of security returns', *The Journal of finance* **45**(3), 881–898.

Jegadeesh, N. and Titman, S. (1993), 'Returns to buying winners and selling losers: Implications for stock market efficiency', *The Journal of finance* **48**(1), 65–91.

Lee, W. (2000), *Theory and methodology of tactical asset allocation*, Vol. 65, John Wiley & Sons.

Liew, J. and Roberts, R. (2013), 'Us equity mean-reversion examined', *Risks* **1**(3), 162–175.

Lo, A. W. (2002), 'The statistics of sharpe ratios', *Financial analysts journal* **58**(4), 36–52.

Markowitz, H. (1952), 'Portfolio selection', *The journal of finance* **7**(1), 77–91.

Meucci, A. (2006), 'Beyond black-litterman in practice: A five-step recipe to input views on non-normal markets'. Available at SSRN: https://ssrn.com/abstract=872577 or http://dx.doi.org/10.2139/ssrn.872577.

Meucci, A. (2008), 'Fully flexible views: Theory and practice'. Available at SSRN: https://ssrn.com/abstract=1213325.

Michaud, R. O. (1989), 'The markowitz optimization enigma: Is 'optimized'optimal?', *Financial Analysts Journal* **45**(1), 31–42.

Morgan, J. et al. (1996), 'Riskmetrics technical document'.

Satchell, S. and Scowcroft, A. (2000), 'A demystification of the black–litterman model: Managing quantitative and traditional portfolio construction', *Journal of Asset Management* **1**(2), 138–150.

SPDR (2018), 'Spdr® msci europe financials ucits etf', `https://no.spdrs.com/en/professional/etf/spdr-msci-europe-financials-ucits-etf-STZ-FP`. Online accessed: 21-April-2018.

Theil, H. (1971), Principles of econometrics, Technical report.

Vidyamurthy, G. (2004), *Pairs Trading: quantitative methods and analysis*, Vol. 217, John Wiley & Sons.

Walters, C. et al. (2013), 'The factor tau in the black-litterman model'.

Walters, C. et al. (2014), 'The black-litterman model in detail'.

Appendix A

The Ornstein-Uhlenbeck Process

In this appendix, the solution of the stochastic differential equation governing the Ornstein-Uhlenbeck process is derived. Different statistical properties of the process are also considered. Starting with the SDE

$$dX_i(t) = \kappa_i \left[\theta_i - X_i(t)\right] dt + \sigma_i dB_i(t), \quad \kappa_i > 0 \tag{A.1}$$

It can be solved using a general framework for solving linear stochastic differential equations. The setup requires the arithmetic Brownian motion and the geometric Brownian motion.

Solving General Linear Stochastic Differential Equations

To show its solution using a general framework, one can use the methodology on how to solve general linear SDEs. Starting with a general linear SDE

$$\begin{cases} dX(t) & = [b_1(t)X(t) + b_2(t)]dt + [\sigma_1(t)X(t) + \sigma_2(t)]dB(t), \quad 0 \leq t \leq T \\ X(0) & = x_0 \end{cases} \tag{A.2}$$

Where

$x_0 \in L_2(\Omega)$ independent of $B = B(t),\ 0 \leq t \leq$

b_1, b_2, σ_1 and σ_2 are deterministic functions on $[0, T]$

Note that the coefficients also satisfy the linear growth condition and the global lipschitz condition. Alas, for a given $(\Omega, \mathcal{F}, \mathcal{P})$ and B Brownian motion, there exists a (unique) strong solution.

By looking at (A.2) it is evident that the SDE is a linear combination of the SDEs for an arithmetic and geometric Brownian motion. Hence, solving for $X(t)$ entails expressing the process as the product of an ABM ($Z(t)$) and a GBM ($Y(t)$). Based on the solutions to these, the drift and dispersion terms of $X(t)$ can be determined such that $X(t) = Y(t)Z(t)$. This means that

$$dX(t) = Y(t)dZ(t) + Z(t)dY(t) + dY(t)dZ(t) \tag{A.3}$$

The solution to the arithmetic Brownian motion SDE is given by

$$X(t) = X(0) + \mu t + \sigma B(t) \tag{A.4}$$

And the geometric Brownian motion SDE admits the following solution

$$X(t) = X(0)e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma B(t)} \tag{A.5}$$

Substituting in (A.4) and (A.5) in (A.3), yields

$$\begin{aligned} dX(t) &= Y(t)[\mu_Z(t)dt + \sigma_Z(t)dB(t)] + Z(t)[\mu_Y(t)Y(t)dt + \sigma_Y(t)Y(t)dB(t)] \\ &\quad + [\mu_Z(t)dt + \sigma_Z(t)dB(t)][\mu_Y(t)Y(t)dt + \sigma_Y(t)Y(t)dB(t)] \\ &= \{[\mu_Z(t) + \sigma_Z(t)\sigma_Y(t)]\, Y(t) + \mu_Y(t)X(t)\}\, dt + [\sigma_Z(t)Y(t)\ \sigma_Y(t)X(t)]\, dB(t) \end{aligned} \tag{A.6}$$

Equating this with (A.2), the drift and diffusion coefficients are given by

$$\mu_Y(t) = b_1(t), \quad \sigma_Y(t) = \sigma_1(t), \quad \mu_Z(t) = \frac{b_2(t) - \sigma_Z(t)\sigma_Y(t)Y(t)}{Y(t)}, \quad \sigma_Z(t) = \frac{\sigma_2(t)}{Y(t)} \tag{A.7}$$

Solution

Based on this, obtaining the solution for (A.1) is straightforward. From (A.7), it is clear that

$$b_1(t) = \kappa, \quad b_2(t) = \kappa\theta, \quad \sigma_1(t) = \sigma_Y(t) = 0, \quad \sigma_2(t) = \sigma$$

Since $\sigma_Y(t) = 0$, the geometric Brownian motion SDE becomes and ordinary differential equation, with solution

$$\frac{dY(t)}{Y(t)} = -\kappa dt \Rightarrow Y(t) = e^{-\kappa t}, \quad Y(0) = 1$$

And

$$\mu_Z(t) = \frac{\kappa\theta}{Y(t)} = \kappa\theta e^{\kappa t}$$

$$\sigma_Z(t) = \frac{\sigma_2(t)}{Y(t)} = \sigma e^{\kappa t}$$

Leading to

$$\begin{cases} Z(t) = Z(0) + \int_0^t \kappa\theta e^{\kappa s} ds + \int_0^t \sigma e^{\kappa s} dB(s) \\ Z(0) = X(0) \end{cases}$$

$X(t)$ then becomes

$$\begin{aligned} X(t) &= Z(t)Y(t) \\ &= e^{-\kappa t}\left[X(0) + \int_0^t \kappa\theta e^{\kappa s} ds + \int_0^t \sigma e^{\kappa s} dB(s)\right] \\ X(t) &= e^{-\kappa t}X(0) + \theta\left[1 - e^{-\kappa t}\right] + \sigma \int_0^t e^{-\kappa(t-s)} dB(s) \end{aligned} \tag{A.8}$$

Properties

The unconditional expectation of the increments of the Ornstein-Uhlenbeck process is zero

$$\begin{aligned} \mathbb{E}\{dX_i(t)\} &= \mathbb{E}\left\{\kappa_i\left[\theta_i - X_i(t)\right]dt + \sigma_i dB_i(t)\right\} \\ &= \kappa_i\theta_i dt - \kappa_i\mathbb{E}[X_i(t)]dt + \sigma_i\mathbb{E}[dB_i(t)] \\ &= \kappa_i\{\theta_i - \mathbb{E}[X_i(t)]\}dt \\ &= \kappa_i[\theta_i - \theta_i] = 0 \end{aligned} \tag{A.9}$$

Conditional mean equal to

$$\mathbb{E}\{dX_i(t)|\mathcal{F}_s, s \le t\} = \kappa_i[\theta_i - X_i(t)]dt \tag{A.10}$$

Which is the daily predicted returns. The sign of the forecast depends on the sign of $\theta_i - X_i(t)$, since $\kappa_i > 0$. $\mathcal{F}_s$ is the σ-algebra generated by the Brownian motion of the process.

The process, (A.8), has an expected value

$$\begin{aligned}\mathbb{E}[X_i(t)] &= \mathbb{E}\left\{e^{-\kappa_i t}X_i(0) + \theta_i\left(1 - e^{-\kappa_i t}\right) + \sigma_i\int_0^t e^{-\kappa_i(t-s)}dB(s)\right\}\\ &= \theta_i\left(1 - e^{-\kappa_i t}\right)\end{aligned} \tag{A.11}$$

And a variance of

$$\begin{aligned}\mathbb{V}ar\left[X_i(t)\right] &= \mathbb{V}ar\left\{e^{-\kappa_i t}X_i(0) + \theta_i\left(1 - e^{-\kappa_i t}\right) + \sigma_i\int_0^t e^{-\kappa_i(t-s)}dB(s)\right\}\\ &= \mathbb{V}ar\left\{\sigma_i\int_0^t e^{-\kappa_i(t-s)}dB(s)\right\}\\ &= \sigma_i^2\left\{\mathbb{E}\left[\left(\int_0^t e^{-\kappa_i(t-s)}dB(s)\right)^2\right] - \left[\mathbb{E}\left(\int_0^t e^{-\kappa_i(t-s)}dB(s)\right)\right]^2\right\}\\ &= \sigma_i^2\left[\frac{1}{2\kappa_i}e^{-\kappa_i(t-s)}\right]_0^t = \frac{\sigma_i^2}{2\kappa_i}\left(1 - e^{-\kappa_i t}\right)\end{aligned} \tag{A.12}$$

Letting time increase towards infinity, the equilibrium distribution for this process becomes normal with

$$\mathbb{E}\{X_it(t)\} = \theta_i \tag{A.13}$$

$$\mathbb{V}ar\{X_i(t)\} = \frac{\sigma_i^2}{2\kappa_i} \tag{A.14}$$

Appendix B

Black-Litterman

In this section, the master formula in the Black-Litterman model, equation (2.14) is derived. The approach in this thesis relies on using point estimates for the expected excess returns. As such only the formula for $\mathbb{E}(\mathbf{r})$ is derived and not the formula for posterior variance, since this is irrelevant. There are many different approaches to derive the Black-Litterman model, but here the way of Theil's mixed estimation method will be used, which relies on the use of generalised least squares estimation, where the beta estimate represents the $\mathbb{E}(\mathbf{r})$. As a starting point, recall the objective of ordinary least squares, which is to minimise the sum of squared residuals. The following will primarily use matrix notation, and all letters and symbols in **bold** represents vectors and matrices.

The error term is defined as

$$\boldsymbol{\epsilon} = \mathbf{y} - \mathbf{x}\boldsymbol{\beta}$$

and residuals as

$$\mathbf{e} = \mathbf{y} - \mathbf{x}\hat{\boldsymbol{\beta}}$$

This leads to sum of squared residuals being $SSR = \mathbf{e}^\intercal\mathbf{e}$. The best estimate for beta, $\hat{\boldsymbol{\beta}}$ is given by the first order condition of minimising SSR.

$$\frac{\partial SSR}{\partial \hat{\boldsymbol{\beta}}} = 0$$

$$\hat{\boldsymbol{\beta}}_{OLS} = (\mathbf{x}^\intercal\mathbf{x})^{-1}\mathbf{x}^\intercal\mathbf{y} \tag{B.1}$$

If $\mathbb{E}(\boldsymbol{\epsilon}) = 0$ and $\mathbb{V}ar(\boldsymbol{\epsilon}) = \sigma^2\mathbf{I}$, then (B.1) is BLUE. In the case of financial time series, it is long been established that asset returns exhibit volatility clustering, resulting in $\mathbb{V}ar(\boldsymbol{\epsilon}) \neq \sigma^2\mathbf{I}$. As such, OLS is not suited for financial time series econometrics. With heteroskedasticity the expectation is that $\mathbb{V}ar(\boldsymbol{\epsilon}) = \boldsymbol{\Xi} \neq \sigma^2\mathbf{I}$. The improved approach should therefore be a transformation ensuring that the error term is homoskedastic. Assuming $\boldsymbol{\Xi}$ is symmetric and positive definite, there exists a matrix $\mathbf{A}$ such that

$$\boldsymbol{\Xi} = \mathbf{A}\mathbf{A}^\intercal \Rightarrow \mathbf{A}^{-1}\boldsymbol{\Xi}(\mathbf{A}^{-1})^\intercal = \mathbf{I}$$

Premultiplying the OLS model by $\mathbf{A}^{-1}$ yields

$$\mathbf{A}^{-1}\mathbf{y} = \mathbf{A}^{-1}\mathbf{x}\boldsymbol{\beta} + \mathbf{A}^{-1}\boldsymbol{\epsilon}$$

With

$$\tilde{\mathbf{y}} = \mathbf{A}^{-1}\mathbf{y}, \quad \tilde{\mathbf{x}} = \mathbf{A}^{-1}\mathbf{x}, \quad \boldsymbol{\xi} = \mathbf{A}^{-1}\boldsymbol{\epsilon}$$

The distributional parameters of the new error term, $\boldsymbol{\xi}$ is

$$\begin{aligned}\mathbb{E}(\boldsymbol{\xi}) &= \mathbb{E}(\mathbf{A}^{-1}\boldsymbol{\epsilon}) = \mathbf{A}^{-1}\mathbb{E}(\boldsymbol{\epsilon}) = 0 \\ \mathbb{V}ar(\boldsymbol{\xi}) &= \mathbb{V}ar(\mathbf{A}^{-1}\boldsymbol{\epsilon}) = \mathbf{A}^{-1}\boldsymbol{\Xi}(\mathbf{A}^{-1})^\intercal = \mathbf{I}\end{aligned}$$

Showing that the transformed model has a homoskedastic error term. Relying on (B.1), the solution to the new model is

$$\begin{aligned}\hat{\boldsymbol{\beta}}_{GLS} &= (\tilde{\mathbf{x}}^\intercal\tilde{\mathbf{x}})^{-1}\tilde{\mathbf{x}}\tilde{\mathbf{y}} \\ &= \left[(\mathbf{A}^{-1}\mathbf{x})^\intercal\mathbf{A}^{-1}\mathbf{x}\right]^\intercal(\mathbf{A}^{-1}\mathbf{x})^\intercal\mathbf{A}^{-1}\mathbf{y} \\ &= \left[\mathbf{x}^\intercal(\mathbf{A}^{-1})^\intercal\mathbf{A}^{-1}\mathbf{x}\right]^{-1}\mathbf{x}^\intercal(\mathbf{A}^{-1})^\intercal\mathbf{A}^{-1}\mathbf{y}\end{aligned}$$

$$\hat{\boldsymbol{\beta}}_{GLS} = \left[\mathbf{x}^\intercal\boldsymbol{\Xi}^{-1}\mathbf{x}\right]^{-1}\mathbf{x}^\intercal\boldsymbol{\Xi}^{-1}\mathbf{y} \tag{B.2}$$

Applying this framework to obtain (2.14), it is assumed that the prior distribution of returns follows a simple linear model on the form

$$\boldsymbol{\Pi} = \mathbf{x}\boldsymbol{\beta} + \boldsymbol{\zeta}$$

Where

$\boldsymbol{\Pi}_{Nx1}$: Equilibrium returns
$\mathbf{x}_{NxN}$: Identity matrix ($\mathbf{I}_N$) with factor loadings
$\boldsymbol{\beta}_{Nx1}$: Unknown means of the return generating process
$\boldsymbol{\zeta}_{NxN}$: Residuals matrix with $\mathbb{E}(\boldsymbol{\zeta}) = 0,\ \mathbb{V}ar(\boldsymbol{\zeta}) = \mathbb{E}(\boldsymbol{\zeta}^\intercal\boldsymbol{\zeta}) = \boldsymbol{\Theta}$ (invertible)

Along with the prior, the investor has private beliefs to be blended with the prior information. This is modelled in the same way as the prior, and the relationship admits the form

$$\mathbf{Q} = \mathbf{P}\boldsymbol{\beta} + \boldsymbol{\eta}$$

Where

$\mathbf{Q}_{kx1}$: Private views returns
$\mathbf{P}_{kxN}$: Matrix mapping views to assets
$\boldsymbol{\beta}_{Nx1}$: Unknown means of return generating process
$\boldsymbol{\eta}_{kx1}$: Residuals vector with $\mathbb{E}(\boldsymbol{\eta}) = 0,\ \mathbb{V}ar(\boldsymbol{\eta}) = \mathbb{E}(\boldsymbol{\eta}^\intercal\boldsymbol{\eta}) = \boldsymbol{\Omega}$ (invertible)

Writing the two models in combination gives

$$\begin{bmatrix} \boldsymbol{\Pi} \\ \mathbf{Q} \end{bmatrix} = \begin{bmatrix} \mathbf{x} \\ \mathbf{P} \end{bmatrix} \boldsymbol{\beta} + \begin{bmatrix} \boldsymbol{\zeta} \\ \boldsymbol{\eta} \end{bmatrix} \tag{B.3}$$

In order to apply (B.2) to (B.3), one needs to know the expression for the variance of the error term.

$$\mathbb{V}ar\left\{ \begin{bmatrix} \boldsymbol{\zeta} \\ \boldsymbol{\eta} \end{bmatrix} \right\} = \mathbb{E}\left\{ \begin{bmatrix} \boldsymbol{\zeta} \\ \boldsymbol{\eta} \end{bmatrix} \begin{bmatrix} \boldsymbol{\zeta}^\intercal & \boldsymbol{\eta}^\intercal \end{bmatrix} \right\} = \mathbb{E}\left\{ \begin{bmatrix} \boldsymbol{\zeta}\boldsymbol{\zeta}^\intercal & \boldsymbol{\zeta}\boldsymbol{\eta}_\intercal \\ \boldsymbol{\eta}\boldsymbol{\zeta}^\intercal & \boldsymbol{\eta}\boldsymbol{\eta}^\intercal \end{bmatrix} \right\}$$

Since prior information and private views are assumed to be independent of each other, the error terms are also independent, resulting in the above expression being reduced to

$$\mathbb{E}\left\{\begin{bmatrix}\boldsymbol{\zeta}\boldsymbol{\zeta}^{\intercal} & 0 \\ 0 & \boldsymbol{\eta}\boldsymbol{\eta}^{\intercal}\end{bmatrix}\right\} = \begin{bmatrix}\mathbb{E}(\boldsymbol{\zeta}\boldsymbol{\zeta}^{\intercal}) & 0 \\ 0 & \mathbb{E}(\boldsymbol{\eta}\boldsymbol{\eta}^{\intercal})\end{bmatrix} = \begin{bmatrix}\boldsymbol{\Theta} & 0 \\ 0 & \boldsymbol{\Omega}\end{bmatrix} \tag{B.4}$$

With this, the GLS estimate of beta becomes

$$\hat{\boldsymbol{\beta}}_{GLS} = \left\{\begin{bmatrix}\mathbf{x}^{\intercal} & \mathbf{P}^{\intercal}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Theta} & 0 \\ 0 & \boldsymbol{\Omega}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{x}^{\intercal} \\ \mathbf{P}^{\intercal}\end{bmatrix}\right\}^{-1}\begin{bmatrix}\mathbf{x}^{\intercal} & \mathbf{P}^{\intercal}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Theta} & 0 \\ 0 & \boldsymbol{\Omega}\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{\Pi} \\ \mathbf{Q}\end{bmatrix}$$

$$= \left\{\begin{bmatrix}\mathbf{x}^{\intercal}\boldsymbol{\Theta}^{-1} & \mathbf{P}^{\intercal}\boldsymbol{\Omega}^{-1}\end{bmatrix}\begin{bmatrix}\mathbf{x} \\ \mathbf{P}\end{bmatrix}\right\}^{-1}\left\{\begin{bmatrix}\mathbf{x}^{\intercal}\boldsymbol{\Theta}^{-1} & \mathbf{P}^{\intercal}\boldsymbol{\Omega}^{-1}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Pi} \\ \mathbf{Q}\end{bmatrix}\right\}$$

$$\hat{\boldsymbol{\beta}}_{GLS} = \left[\mathbf{x}^{\intercal}\boldsymbol{\Theta}^{-1}\mathbf{x} + \mathbf{P}^{\intercal}\boldsymbol{\Omega}^{-1}\mathbf{P}\right]^{-1}\left[\mathbf{x}^{\intercal}\boldsymbol{\Theta}^{-1}\boldsymbol{\Pi} + \mathbf{P}^{\intercal}\boldsymbol{\Omega}^{-1}\mathbf{Q}\right] \tag{B.5}$$

In the Black-Litterman model there is only one factor per asset. As such $\mathbf{x}$ is an identity matrix and it can be dropped from (B.5). This leaves

$$\hat{\boldsymbol{\beta}}_{GLS} = \left[\boldsymbol{\Theta}^{-1} + \mathbf{P}^{\intercal}\boldsymbol{\Omega}^{-1}\mathbf{P}\right]^{-1}\left[\boldsymbol{\Theta}^{-1}\boldsymbol{\Pi} + \mathbf{P}^{\intercal}\boldsymbol{\Omega}^{-1}\mathbf{Q}\right]$$

To arrive at (2.14), set $\boldsymbol{\Theta}^{-1} = (\tau\boldsymbol{\Sigma})^{-1}$ and $\mathbb{E}(\mathbf{r}) = \hat{\boldsymbol{\beta}}_{GLS}$

Appendix C

Active Management

The aim of this section is to provide transparent derivations of results used in section 2.5. To ensure this, the derivations are done in small steps, but key results will be marked by numbered equations. This chapter is largely based on Lee (2000), chapter 2 and Da Silva et al. (2009), but connects the two to provide insight on the relationship between the Black-Litterman model and active management.

Mean-Variance Paradigm

As described by von Neumann and Morgenstern, the objective of the investor is to maximise expected utility of wealth. Assuming constant relative risk aversion and that asset returns follow a multivariate normal distribution, the expected utility of wealth can be represented as

$$\mathbb{E}\{U(W)\} = -e^{-\delta\left(\mu_p - \frac{\delta}{2}\sigma_p^2\right)} \tag{C.1}$$

Here, $\mu_p = \boldsymbol{\omega}^\intercal \boldsymbol{\mu}$ is the expected return on the portfolio, and $\sigma_p^2 = \boldsymbol{\omega}^\intercal \boldsymbol{\Sigma} \boldsymbol{\omega}$ is the portfolio variance. Maximising expected utility is achieved by

$$\begin{aligned} \underset{\boldsymbol{\omega}}{\arg\max} \quad & \boldsymbol{\omega}^\intercal \boldsymbol{\mu} - \boldsymbol{\omega}^\intercal \boldsymbol{\Sigma} \boldsymbol{\omega} \\ \text{subject to} \quad & \boldsymbol{\omega}^\intercal \mathbf{1} = 1 \end{aligned} \tag{C.2}$$

Solving this optimisation problem is done through the first order conditions of the Lagrangian.

$$\mathcal{L} = \boldsymbol{\omega}^\intercal \boldsymbol{\mu} - \frac{\delta}{2} \boldsymbol{\omega}^\intercal \boldsymbol{\Sigma} \boldsymbol{\omega} - \lambda(\boldsymbol{\omega}^\intercal \mathbf{1} - 1)$$

First order conditions are

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\omega}} = \boldsymbol{\mu} - \delta \boldsymbol{\Sigma} \boldsymbol{\omega} - \mathbf{1}\lambda = 0$$

$$\Rightarrow \boldsymbol{\omega}^* = \frac{1}{\delta} \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \mathbf{1}\lambda) \tag{C.3}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \boldsymbol{\omega}^\intercal \mathbf{1} = 1$$

$$\Leftrightarrow \mathbf{1}^\intercal \boldsymbol{\omega} = 1 \tag{C.4}$$

Inserting (C.3) in (C.4) yields

$$\mathbf{1}^\intercal \boldsymbol{\omega} = \mathbf{1}^\intercal \left[\frac{1}{\delta} \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{\mu} - \mathbf{1}\lambda \right) \right] = 1$$

$$\Rightarrow \delta = \mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - \mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \mathbf{1} \lambda$$

$$\Rightarrow \lambda = \frac{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}}{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \mathbf{1}} - \frac{\delta}{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \mathbf{1}} \tag{C.5}$$

Setting (C.5) back into (C.3) results in

$$\begin{aligned}
\boldsymbol{\omega}^* &= \frac{\boldsymbol{\Sigma}^{-1}}{\delta} \left[\boldsymbol{\mu} - \left(\frac{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}}{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \mathbf{1}} - \frac{\delta}{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \mathbf{1}} \right) \mathbf{1} \right] \\
&= \frac{\boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}}{\delta} - \frac{\boldsymbol{\Sigma}^{-1} \mathbf{1}}{\delta} \left(\frac{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}}{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \mathbf{1}} - \frac{\delta}{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \mathbf{1}} \right) \\
&= \frac{\boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}}{\delta} + \frac{\boldsymbol{\Sigma}^{-1} \mathbf{1}}{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \mathbf{1}} \left(1 - \frac{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}}{\delta} \right)
\end{aligned}$$

Note that $\boldsymbol{\omega}_{GMVP} = \frac{\boldsymbol{\Sigma}^{-1} \mathbf{1}}{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \mathbf{1}}$. As such, the above can be rearranged, such that

$$
\begin{aligned}
\boldsymbol{\omega}^* &= \boldsymbol{\omega}_{GMVP} + \frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\delta} - \frac{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\delta}\frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{1}} \\
&= \boldsymbol{\omega}_{GMVP} + \frac{\boldsymbol{\Sigma}^{-1}}{\delta}\left(\boldsymbol{\mu} - \mathbf{1}\frac{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}}{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{1}}\boldsymbol{\mu}\right)
\end{aligned} \tag{C.6}
$$

To progress, the concept of equilibrium returns $\overline{\boldsymbol{\mu}}$, which is often used as input in portfolio optimisation will often not equal expected returns, as the economy rarely is in equilibrium. With this, (C.6) can be rewritten as

$$
\begin{aligned}
\boldsymbol{\omega}^* &= \boldsymbol{\omega}_{GMVP} + \frac{\boldsymbol{\Sigma}^{-1}}{\delta}\left[\boldsymbol{\mu} - \overline{\boldsymbol{\mu}} + \overline{\boldsymbol{\mu}} - \mathbf{1}\frac{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \overline{\boldsymbol{\mu}} + \overline{\boldsymbol{\mu}})}{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{1}}\right] \\
&= \boldsymbol{\omega}_{GMVP} + \frac{\boldsymbol{\Sigma}^{-1}}{\delta}\left(\overline{\boldsymbol{\mu}} - \mathbf{1}\frac{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\overline{\boldsymbol{\mu}}}{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{1}}\right) + \frac{\boldsymbol{\Sigma}^{-1}}{\delta}\left(\boldsymbol{\mu} - \overline{\boldsymbol{\mu}} - \mathbf{1}\frac{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \overline{\boldsymbol{\mu}})}{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{1}}\right) \\
&= \boldsymbol{\omega}_{GMVP} + \frac{1}{\delta}\frac{\boldsymbol{\Sigma}^{-1}}{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{1}}\left(\overline{\boldsymbol{\mu}}\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{1} - \mathbf{1}\overline{\boldsymbol{\mu}}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{1}\right) + \\
&\frac{1}{\delta}\frac{\boldsymbol{\Sigma}^{-1}}{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{1}}\left[(\boldsymbol{\mu} - \overline{\boldsymbol{\mu}})\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{1} - \mathbf{1}(\boldsymbol{\mu} - \overline{\boldsymbol{\mu}})^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{1}\right] \\
\boldsymbol{\omega}^* &= \boldsymbol{\omega}_{GMVP} + \frac{1}{\delta}\frac{\boldsymbol{\Sigma}^{-1}(\overline{\boldsymbol{\mu}}\mathbf{1}^\intercal - \mathbf{1}\overline{\boldsymbol{\mu}}^\intercal)\boldsymbol{\Sigma}^{-1}}{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{1}}\mathbf{1} + \frac{1}{\delta}\frac{\boldsymbol{\Sigma}^{-1}\left[(\boldsymbol{\mu} - \overline{\boldsymbol{\mu}})\mathbf{1}^\intercal - \mathbf{1}(\boldsymbol{\mu} - \overline{\boldsymbol{\mu}})^\intercal\right]\boldsymbol{\Sigma}^{-1}}{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{1}}\mathbf{1}
\end{aligned} \tag{C.7}
$$

The optimal portfolio weights when maximising the expected utility when regarding total return and total risk, consists of three terms. This can be rewritten as

$$
\boldsymbol{\omega}^* = \boldsymbol{\omega}_{GMVP} + \boldsymbol{\omega}_S + \boldsymbol{\omega}_T \tag{C.8}
$$

Where

$$
\begin{aligned}
\boldsymbol{\omega}_S &= \frac{1}{\delta}\frac{\boldsymbol{\Sigma}^{-1}(\overline{\boldsymbol{\mu}}\mathbf{1}^\intercal - \mathbf{1}\overline{\boldsymbol{\mu}}^\intercal)\boldsymbol{\Sigma}^{-1}}{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{1}}\mathbf{1} \\
\boldsymbol{\omega}_T &= \frac{1}{\delta}\frac{\boldsymbol{\Sigma}^{-1}\left[(\boldsymbol{\mu} - \overline{\boldsymbol{\mu}})\mathbf{1}^\intercal - \mathbf{1}(\boldsymbol{\mu} - \overline{\boldsymbol{\mu}})^\intercal\right]\boldsymbol{\Sigma}^{-1}}{\mathbf{1}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{1}}\mathbf{1}
\end{aligned}
$$

This shows that the optimal portfolio weights consists of three different components when maximising expected utility of wealth in the mean-variance paradigm. The optimal weights are decided by the weights in the global minimum variance portfolio, a strategic bet and a tactical bet.

Active management Paradigm

In this part, the focus is shifted from total risk and return, and instead the optimal portfolio is derived when the investor only concerns himself with active risk and return. Such a situation is common in the investment industry, where managers are ranked based on by how much they can beat their benchmark portfolio. In this setting, $\boldsymbol{\omega}_{GMVP} + \boldsymbol{\omega}_S$ makes up the benchmark portfolio. In this part, the assumption is that the investor does not measure utility based on level of wealth, as determined by the total risk and return of the portfolio, but rather by how much risk and return is generated in excess of a benchmark index.

In this setting, it is sufficient to express the utility of the investor in terms of alpha and tracking error. The portfolio's alpha is the excess return over then benchmark portfolio, and tracking error is the standard deviation of the alpha. Assuming either quadratic utility, or that alphas from active positions are follows a multivariate normal distribution with a constant relative risk aversion concerning tracking error, expected utility is

$$\mathbb{E}[U(W)] = -e^{-\delta_T(\alpha - TE^2)} \tag{C.9}$$

Where

$$\alpha = \boldsymbol{\omega}_a^\intercal \boldsymbol{\mu}$$
$$TE^2 = \mathbb{V}ar(\alpha) = \boldsymbol{\omega}_a^\intercal \boldsymbol{\Sigma} \boldsymbol{\omega}_a$$

δ_T measures the risk tolerance for active risk [1], and is to be distinguished from δ which is the tolerance for *total* risk. To maximise (C.9), the following must be solved

$$\begin{aligned} \underset{\boldsymbol{\omega}_a}{\arg\max} \quad & \boldsymbol{\omega}_a^\intercal \boldsymbol{\mu} - \boldsymbol{\omega}_a^\intercal \boldsymbol{\Sigma} \boldsymbol{\omega}_a \\ \text{subject to} \quad & \boldsymbol{\omega}_a^\intercal \mathbf{1} = 0 \end{aligned} \tag{C.10}$$

The Lagrangian is thus

$$\mathcal{L} = \boldsymbol{\omega}_a^\intercal \boldsymbol{\mu} - \frac{\delta_T}{2} \boldsymbol{\omega}_a^\intercal \boldsymbol{\Sigma} \boldsymbol{\omega}_a - \lambda_T \boldsymbol{\omega}_a^\intercal \mathbf{1}$$

[1] Tracking error

And first order conditions

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\omega}_a} = \boldsymbol{\mu} - \delta_T \boldsymbol{\omega}_a \boldsymbol{\Sigma} - \lambda_T \mathbf{1} = 0$$

$$\Rightarrow \boldsymbol{\omega}_a^* = \frac{\boldsymbol{\Sigma}^{-1}}{\delta_T}(\boldsymbol{\mu} - \mathbf{1}\delta_T) \tag{C.11}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_T} = \boldsymbol{\omega}_a^\intercal \mathbf{1} = 0 \Leftrightarrow \mathbf{1}^\intercal \boldsymbol{\omega}_a = 0 \tag{C.12}$$

Inserting (C.11) in (C.12) yields

$$0 = \mathbf{1}^\intercal \frac{\boldsymbol{\Sigma}^{-1}}{\delta_T}(\boldsymbol{\mu} - \mathbf{1}\lambda_T)$$

$$0 = \mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - \mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \mathbf{1} \lambda_T$$

$$\lambda_T = \frac{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1}}{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \mathbf{1}} \boldsymbol{\mu} = \boldsymbol{\omega}_{GMVP} \boldsymbol{\mu} \tag{C.13}$$

Substituting (C.13) back in (C.11), results in

$$\begin{aligned} \boldsymbol{\omega}_a^* &= \frac{\boldsymbol{\Sigma}^{-1}}{\delta_T}\left(\boldsymbol{\mu} - \mathbf{1}\frac{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}}{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \mathbf{1}}\right) \qquad \text{(C.14)} \\ &= \frac{1}{\delta_T} \frac{\boldsymbol{\Sigma}^{-1}}{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \mathbf{1}} \boldsymbol{\mu} \mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \mathbf{1} - \mathbf{1} \frac{1}{\delta_T} \frac{\boldsymbol{\Sigma}^{-1}}{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \mathbf{1}} \boldsymbol{\mu}^\intercal \boldsymbol{\Sigma}^{-1} \mathbf{1} \end{aligned}$$

$$\boldsymbol{\omega}_a^* = \frac{1}{\delta_T} \frac{\boldsymbol{\Sigma}^{-1} \left[\boldsymbol{\mu}\mathbf{1}^\intercal - \mathbf{1}\boldsymbol{\mu}^\intercal\right] \boldsymbol{\Sigma}^{-1}}{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1} \mathbf{1}} \mathbf{1} \tag{C.15}$$

Also, note that (C.15) can be rewritten as

$$\boldsymbol{\omega}_a^* = \frac{\boldsymbol{\Sigma}^{-1}}{\delta_T}(\boldsymbol{\mu} - \mu_{GMVP}\mathbf{1}) \tag{C.16}$$

Bringing it Home

Optimal solution

To show how the two parts combine, start with (C.14) and substitute $\boldsymbol{\mu}$ with $\overline{\boldsymbol{\mu}}$, which in the Black-Litterman model is determined by (2.10). This results in

$$\begin{aligned}
\boldsymbol{\omega}_a^* &= \frac{\boldsymbol{\Sigma}^{-1}}{\delta_T}(\overline{\boldsymbol{\mu}} - \boldsymbol{\omega}_{GMVP}^\intercal \overline{\boldsymbol{\mu}}\mathbf{1}) \\
&= \frac{\boldsymbol{\Sigma}^{-1}}{\delta_T}(\delta \boldsymbol{\Sigma}\boldsymbol{\omega}_B - \boldsymbol{\omega}_{GMVP}^\intercal \delta \boldsymbol{\Sigma}\boldsymbol{\omega}_B)\mathbf{1}) \\
&= \frac{\delta}{\delta_T}(\boldsymbol{\omega}_B - \boldsymbol{\Sigma}^{-1}\boldsymbol{\omega}_{GMVP}^\intercal \boldsymbol{\Sigma}\boldsymbol{\omega}_B\mathbf{1})
\end{aligned} \tag{C.17}$$

Using the budget constraint from the mean-variance optimisation problem, the following applies

$$\begin{aligned}
\boldsymbol{\omega}_{GMVP}^\intercal \boldsymbol{\Sigma}\boldsymbol{\omega}_B &= \frac{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1}}{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1}\mathbf{1}} \boldsymbol{\Sigma}\boldsymbol{\omega}_B \\
&= \frac{\mathbf{1}^\intercal \boldsymbol{\omega}_B}{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1}\mathbf{1}} \\
&= \frac{1}{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1}\mathbf{1}} = \sigma_{GMVP}^2
\end{aligned} \tag{C.18}$$

Setting this back into (C.17), gives

$$\boldsymbol{\omega}_a^* = \frac{\delta}{\delta_T}\left(\boldsymbol{\omega}_B - \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\intercal \boldsymbol{\Sigma}^{-1}\mathbf{1}}\right)$$

$$\boldsymbol{\omega}_a^* = \frac{\delta}{\delta_T}(\boldsymbol{\omega}_B - \boldsymbol{\omega}_{GMVP}) \tag{C.19}$$

Perceived upside

With this dynamic, in a no-information situation, the portfolio will seemingly return an alpha governed by

$$\begin{aligned}\alpha &= \boldsymbol{\omega}_a^\intercal \overline{\boldsymbol{\mu}} \\ &= \frac{\delta}{\delta_T}(\boldsymbol{\omega}_B - \boldsymbol{\omega}_{GMVP})^\intercal \delta \boldsymbol{\Sigma} \boldsymbol{\omega}_B \\ &= \frac{\delta^2}{\delta_T}(\boldsymbol{\omega}_B^\intercal \boldsymbol{\Sigma} \boldsymbol{\omega}_B - \boldsymbol{\omega}_{GMVP}^\intercal \boldsymbol{\Sigma} \boldsymbol{\omega}_B)\end{aligned}$$

$$\alpha = \frac{\delta^2}{\delta_T}(\sigma_B^2 - \sigma_{GMVP}^2) \tag{C.20}$$

The uninformed portfolios tracking error is given by

$$\begin{aligned}TE &= \sqrt{\boldsymbol{\omega}_a^\intercal \boldsymbol{\Sigma} \boldsymbol{\omega}_a} \\ &= \sqrt{\frac{\delta}{\delta_T}(\boldsymbol{\omega}_B - \boldsymbol{\omega}_{GMVP})^\intercal \boldsymbol{\Sigma} \frac{\delta}{\delta_T}(\boldsymbol{\omega}_B - \boldsymbol{\omega}_{GMVP})} \\ &= \frac{\delta}{\delta_T}\sqrt{\sigma_B^2 + \sigma_{GMVP}^2 - 2\sigma_{GMVP,B}}\end{aligned}$$

$$TE = \frac{\delta}{\delta_T}\sqrt{\sigma_B^2 - \sigma_{GMVP}^2} \tag{C.21}$$

Observed information ratio is thus

$$IR = \frac{\alpha}{TE} = \frac{\frac{\delta^2}{\delta_T}(\sigma_B^2 - \sigma_{GMVP}^2)}{\frac{\delta}{\delta_T}\sqrt{\sigma_B^2 - \sigma_{GMVP}^2}}$$

$$IR = \delta\sqrt{\sigma_B^2 - \sigma_{GMVP}^2} \tag{C.22}$$

Hidden downside

The no-investment view scenario will lead to systematic risk exposure to benchmark portfolio given by

$$\sigma_{a,B} = \boldsymbol{\omega}_a^\intercal \boldsymbol{\Sigma} \boldsymbol{\omega}_B = \frac{\delta}{\delta_T}(\sigma_B^2 - \sigma_{GMVP}^2)$$

$$\beta_a = \frac{\sigma_{a,B}}{\sigma_B^2} = \frac{\frac{\delta}{\delta_T}(\sigma_B^2 - \sigma_{GMVP}^2)}{\sigma_B^2}$$

$$\beta_a = \frac{\delta}{\delta_T}\left(1 - \frac{\sigma_{GMVP}^2}{\sigma_B^2}\right) \tag{C.23}$$

The total beta of the portfolio is determined by

$$\beta_p = \frac{(\boldsymbol{\omega}_a + \boldsymbol{\omega}_B)^\intercal \boldsymbol{\Sigma}\boldsymbol{\omega}_B}{\sigma_B^2} = 1 + \beta_a \tag{C.24}$$

Portfolio volatility is given by

$$\sigma_p^2 = (\boldsymbol{\omega}_a + \boldsymbol{\omega}_B)^\intercal \boldsymbol{\Sigma}(\boldsymbol{\omega}_a + \boldsymbol{\omega}_B)$$

$$\sigma_p = \sigma_B\sqrt{\beta_a\left(\frac{\delta}{\delta_T} + 2\right) + 1} \tag{C.25}$$

Appendix D

Literature Review

The central role of investment programs in asset management has led to a diverse body of literature. This literature review will not cover the field in its entirety, but focus on the main elements of the applied strategy: Statistical arbitrage and the Black-Litterman model. A brief review of the different levels of asset allocations is also provided to outline where this strategy fits in.

D.1 Tactical Asset Allocation

Achieving the optimal blend of risk and return is the central task of any investment decision. Policy choice can account for more than 90 % of the variation in quarterly returns, according to Brinson et al. (1995). This shows that a fund manager's investment plan should be a determining factor in the allocation process. Dahlquist and Harvey (2001) states that there are three investment categories: Benchmark, strategic or tactical asset allocation. Benchmark investing encapsulates the index fund style of investment, strategic allocation is the formation of views on the future performance of different assets or asset classes, before taking long-term bets and deviating from the chosen benchmark. Tactical asset allocation (TAA) is best defined by (Arnott and Fabozzi, 1988, p 4):

> Tactical asset allocation broadly refers to active strategies which seek to enhance performance by opportunistically shifting the asset mix of a portfolio in response to the changing patterns of reward available in the capital markets. Notably, tactical asset allocation tends to refer to disciplined processes for evaluating prospective rates of return on various asset classes and establish an asset allocation response intended to capture higher rewards. In the various implementations of tactical asset allocation, there are different investment

horizons and different mechanisms for evaluating the asset allocation decision.

The prospects of TAA being able to increase portfolio performance is explored in Dahlquist and Harvey (2001). In it the authors outline a framework using conditional information in the asset allocation process. They find that information only needs a small degree of predictive power before inflicting large alterations to capital allocation. Consequently, they conclude that active strategies can exceed benchmark investing. Exploring if TAA in practice can outperform passive investing, is too wide to fully cover. Therefore, two popular TAA signals are reviewed: Momentum and contrarian.

The momentum strategy has received considerable attention in the literature. Jegadeesh and Titman document its results and find that going long previous winners and shorting previous losers yield an excess return of 12% p.a. on average between 1965 and 1989. They also point out that these returns are not caused by the systematic risk of the strategy, resulting in a puzzle that was answered by Chordia and Shivakumar (2002), where they showed how the anomalous returns can be explained by a set of lagged macroeconomic variables associated with the business cycle.

Opposite of the momentum strategy is the contrarian strategy. Its philosophy is to go long poor performers and short the strong performers, on the assumption that the market overreacts to news, implying that winners and losers are relatively over- and undervalued. Bondt and Thaler (1985) find that portfolios of historical losers outperformed portfolios of prior winners. Chan (1988) explains this by proposing that the portfolio risk is time-variant. Hence, abnormal returns are affected by how risk is estimated. Using CAPM in conjunction with other approaches, he reveals how following heavy losses (gains), a stock's beta increases (decreases), rendering the estimation of betas using historical data inappropriate. As a result, the seemingly abnormal returns are normal compensation for the strategy's inherent risk.

D.2 Statitical Arbitrage

The term statistical arbitrage covers a set of quantitative trading strategies, aiming at exploiting temporal statistical mispricing. Since its inception it has been embraced by the industry, in contrast to its relatively low popularity by academics [1].

"Pairs Trading: Performance of a Relative-Value Arbitrage Rule" by Gatev et al. is

[1] As of 22-April-2018, there are 9,895 citations for the central momentum paper by Jegadeesh and Titman (1993) and 2,347 citations for the central contrarian paper by Jegadeesh (1990) whilst Gatev et al. (2006) only has 640 citations.

central in the literature on statistical arbitrage. Using the sum of Euclidean squared distances (SSD), the authors rank all possible pairs of liquid U.S. stocks between 1962 and 2002 [2]. From $\frac{n(n-1)}{2}$ possible pairs, the 20 pairs with the lowest SSDs are traded on over the next six months. This rule generated excess returns of up to 11% p.a. before costs. The paper ascribes the returns to an unknown risk factor, after showing high portfolio correlations between distinct pairs. Moreover, after adjusting the returns using a version of the Fama and French three factor model, the correlations were still present.

Expanding on the work of Gatev et al., Galenko et al. (2012) develop a similar strategy, but it is based on a cointegration framework, since the strategy hinges on the assumption that the pairs revert back to a long-run equilibrium relationship [3]. Their methodology reduces divergence risk present in Gatev et al. (2006).

This source of risk is confirmed in Do and Faff (2010), where the approach of Gatev et al. is applied on an expanded dataset. They find that 32% of the pairs, formed by Euclidean distance, do not converge. Moreover, in a later paper, they include transaction costs and find the strategy by Gatev et al. to be unprofitable (Do and Faff, 2012), while also trying to establish economically intuitive pairs by only matching securities in the same industries [4].

Andrade et al. identifies the unknown risk factor proposed by Gatev et al. in their paper *"Understanding the profitability of pairs trading"*. Investigating the Taiwanese stock market between 1994 and 2004, they find the excess return from pairs trading to be compensation for providing liquidity to uninformed buyers. The connection is established by strong correlation between strategy profit and uninformed demand shocks in the underlying securities.

Modifying the approach by Gatev et al., Engelberg et al. (2008) demonstrate how the strategy's profitability decays with time and its dependence on the nature of the events occurring when securities diverge. Essentially, it is suggested that macro level information affects both securities in a pair at different rates, which is stated to be the main profit driver. The explanation being that, because of market frictions [5], the information is not reflected in prices at an equal pace, causing temporal mispricing.

Jacobs and Weber (2013) strengthens the notion that different absorption rates are

[2]Each price series is constructed as a cumulative total return index, normalized to the first day of a 12 months period, with reinvested dividends.

[3]For details, see Engle and Granger (1987)

[4]48 industries as proposed by Fama and French, see `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_48_ind_port.html`

[5]Low liquidity etc.

driving pairs trading returns by observing that trading opportunities are more likely to open on days where sizeable amounts of new information is released. This alters the focus from company to market level. Consequently, prices reflect common information at different velocities, creating temporary deviations from equilibrium.

In addition to cointegration being extensively used to determine suitable pairs, other techniques are also studied, such as modelling the pair spread as a mean-reverting process. Elliott et al. (2005) applies such a methodology, relying on the Ornstein-Uhlenbeck process to describe the spread yields, which generates interesting insights. With it, forecasts regarding mean-reversion times can be estimated and probabilities for divergence is also forecastable. Their paper is theoretical and offers no empirical results.

Finally, one of the most influential paper for this thesis is *"Statistical Arbitrage in the U.S. Equities Market"* by Avellaneda and Lee. The main idea presented here is to decompose stock returns into systematic and idiosyncratic factors, before modelling the assumed cointegrated idiosyncratic component as an Ornstein-Uhlenbeck process.

Based on this trading signals are generated. Furthermore, they use different sets of risk factors, to investigate the effect on the residual component and the strategy's performance. To determine different risk factors, Avellaneda and Lee use Principle Component Analysis (PCA) on the correlation matrix of the broad equities market, and ETFs. With their mean-reversion model, they create a dimensionless variable, s-score, measuring the distance of the residual from its equilibrium value in standard deviations. Trading positions are initiated when this value crosses certain thresholds [6].

Covering the entire period from 1997 to 2007, Sharpe ratios of 1.44 and 1.10 are reported for the PCA and ETF approach, respectively. They also find that low volatility regimes are associated with more factors being needed to explain the variation in the equity premium. This is linked to strategy performance, with a low number of required factors being advantageous.

D.3 Black-Litterman

Fisher Black and Robert Litterman contributed to asset allocation with their paper *"Asset Allocation: Combining Investor Views with Market Equilibrium"* (Black and Litterman, 1990), in which they renewed the interest in Modern Portfolio Theory, a field with little

[6] The values are determined empirically and depend on the nature of the trade, i.e. long/short and/or open/close

industry interest following its introduction by Harry Markowitz [7]. The lacking interest was attributed to a series of practical challenges [8]. Using the framework introduced by Theil (1971), they provided investors with a model for how they could blend private information with the implied excess equilibrium returns. Since its introduction, the model has been widely studied, and alterations have been put forward. Thus, the literature on the original model is first reviewed, before a selection of papers with alternative definitions are highlighted.

The paper originated in an internal Research Note within Goldman Sachs Fixed income. It was later extended and published in the Financial Analysts Journal. In it, intuition for the model is provided along with partial derivations. Following this, He and Litterman (1999) provides a reproducible example. Bevan and Winkelmann (1998) describe how Goldman Sachs use the model in their daily operations.

Expected returns being normally distributed, with an unknown mean, is a central assumption in the Black-Litterman model. As such, the mean itself is a random variable with its own distribution. Consequently, uncertainty in the prior estimate of equilibrium excess returns is needed, and is achieved through the parameter τ. Determining its value is challenging, but in his paper *"The Factor Tau in the Black-Litterman Model"* 2013, Jay Walters attempt to develop an understanding for it and demonstrate proper calibration techniques.

With τ being difficult to estimate, the literature gives contradictory advice on its calibration. As a result, alternative models have been proposed, excluding the parameter entirely. Satchell and Scowcroft (2000) attempts to clarify the Black-Litterman model. Instead they alter it, using point estimates for the prior returns and the views, contrary to these being random variables, as in the original paper.

Building on Satchell and Scowcroft, Fusai and Meucci (2003) advocate for another model adjustment removing τ. Meucci has researched the model extensively by offering a method on how to incorporate non-normal views (Meucci, 2006), opening up for non-normality in every parameter Meucci (2008) and demonstrating how to conduct scenario analysis.

Specifying the covariance matrix, $\boldsymbol{\Omega}$, of private views around the unknown mean return is also challenging. In *"A step-by-step guide to the Black-Litterman model: Incorporating user-specified confidence levels"*, Idzorek suggest a simplifying approach, allowing investors to express confidence in a view as a number between 0 and 100 % . Based on the confidence

[7] See Portfolio Selection (Markowitz, 1952)

[8] For further details see Michaud (1989).

in each view, the posterior return estimates are tilted away from the prior estimates. A high degree of confidence results in a large tilt and vice versa.

The use of the Black-Litterman model in active portfolio management is investigated in Herold (2003). Focusing on generating positive alpha, the paper's contribution, in addition to showing how to apply the model in active management, is to propose measures to judge the validity of views. Herold applies a model using point estimates, similar to Fusai and Meucci (2003).

Applying the Black-Litterman model in active management poses challenges, as demonstrated by Da Silva et al. (2009). First, they argue that since active views are the essence of the model, it must be analyzed within the active management framework. Second, since it is originally derived under the mean-variance framework, implementing it as is, in active management results in suboptimal trades and unintentional risk taking. Lastly, they discuss possible remedies, including setting the equilibrium excess returns to zero, also proposed by Herold.

What becomes apparent when reviewing the literature on the Black-Litterman model is the difficulty in finding papers discussing how to generate views. One of these is *"U.S. Equity Mean-Reversion Examined"*, where Liew and Roberts apply the methodology of s-scores and conditional mean returns from the Ornstein-Uhlenbeck process , put forward by Avellaneda and Lee to generate views in the Black-Litterman model .

Their approach offers an elegant way of forming views on a set of securities and expressing confidence in these. Additionally, they adopt the zero equilibrium excess returns proposed by Da Silva et al.. Like Avellaneda and Lee (2010), the authors make use of PCA analysis and ETFs to investigate the source of profitability for their strategy. In contrast to Avellaneda and Lee (2010), Liew and Roberts find that the larger the number of components needed to explain a given level of variation is beneficial. This thesis is largely based on their paper.

Appendix E

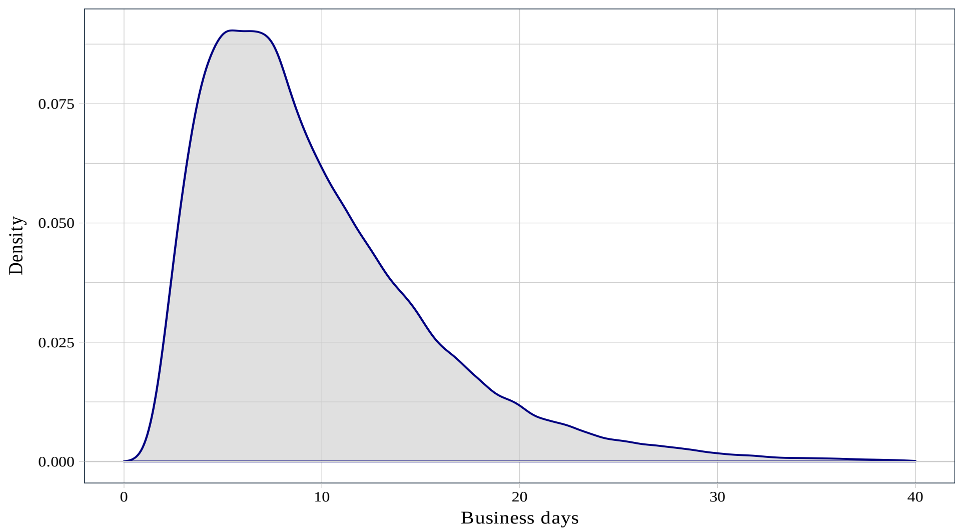
Additional Results

This appendix provides a review of the existing literature, mean-reversion results from the trading universe and from implementing all strategy varieties before costs. Furthermore, details about data corrections are outlined.

E.1 Mean-Reversion

Given that the residual process $dX_{i,t}$ is assumed to be cointegrated, the constituents of the trading universe should have a mean-reversion time of 35 days or less, corresponding to half the time window used to estimate the parameters of the Ornstein-Uhlenbeck process.

Figure E.1: Empirical Distribution of Mean-Reversion Time



This plot shows the empirical density of mean-reversion time, τ in days, for the all securities from 2005 to 2017. Note that the mean reversion time is estimated as $\frac{1}{\kappa_{i,t}}$.

With an expected mean-reversion time of less than ten days and a standard deviation of seven days across all sectors for the entire period, the constituents exhibit rapid mean-reversion. The literature has tried to explain the sources of deviation from the equilibrium relationship between pairs of financial assets. One theory, by Andrade et al., is that price discrepancies arise from uninformed demand shocks. Given that the constituents of the OBX is the 25 most liquid companies on OSE, such shocks should be absorbed faster than for the average stock.

Overall, the securities tend to mean revert quickly. The table below splits mean-reversion time by sector, to see if this behaviour is a common feature for all the sectors or not.

Table E.1: Descriptive Statistics- Mean-Reversion Time

Statistic	Maximum	Third quartile	Median	Mean	First quartile	Minimum	Fast days	Standard deviation
Energy	822.580	11.974	7.900	9.293	5.163	0.972	0.997	7.532
Materials	44.856	12.688	8.731	9.773	5.674	1.363	0.998	5.535
Industrials	48.736	13.767	8.763	10.429	5.891	1.593	0.998	6.117
Consumer Discretionary	63.794	14.314	9.313	11.278	6.448	1.353	0.990	7.165
Consumer Staples	78.646	13.463	9.508	10.630	6.250	1.162	0.994	6.135
Health Care	42.821	15.595	10.392	11.436	6.166	1.294	0.994	6.755
Financials	397.801	11.786	7.794	9.409	5.243	1.014	0.996	7.381
Technology	279.167	12.409	8.446	9.897	5.766	1.506	0.996	7.458
Telecom	137.100	11.151	7.543	8.746	5.327	1.787	0.999	5.286

This table provides insight into the nature of mean reversion for the constituents of the OBX index, based on a sector division. The row named *Fast days* are the percentage share of observations that have a mean reversion time lower than or equal to half a period, 35 days.

All sectors are close in means, with a span from 8.7 to 11.5 days. Furthermore, the share of fast mean-reversion observations is close to 100% for all industries. The largest difference is found intra-sector in the max values, where Energy, Financials and Technology have some observations that do not mean-revert, i.e. the speed of mean reversion, κ is low, indicating non-stationary time series with unit root. Consequently, for these instances, the model's assumptions do not hold, and gives unreliable results.

These findings indicate that the chosen trading universe is appropriate for a mean-reversion strategy. Additionally, there are no large differences between industries since all sector distributions are relatively homogeneous. This is a clear advantage considering the already limited number of stocks in the trading universe.

E.2 Results for all Portfolios

The essence here is the strategy's hypothetical performance between 2005 and 2017. To illustrate it, a selection of performance measures for the benchmark portfolio and reference portfolios [1] is reported along with gross results for all strategy variations.

Results are split into three categories: Return measures, risk measures and risk adjusted measures. Furthermore, each of the strategy portfolios have their own names: Statistical Arbitrage - Black-Litterman (SABL), a code A or B, indicating if the covariance matrices are estimated using equal weights or with exponential weighted moving averages. Lastly, the numeric value from one to four indicates whether the drift is included and/or if centred means are used. The different codes are summarised below.

Table E.2: Portfolio Variations

Portfolio	Strategy name	Covariance estimation technique	Included drift	Estimator means
SABL A1	Statistical Arbitrage - Black-Litterman	Equal weighted	No	Uncentred
SABL A2	Statistical Arbitrage - Black-Litterman	Equal weighted	No	Centred
SABL A3	Statistical Arbitrage - Black-Litterman	Equal weighted	Yes	Uncentred
SABL A4	Statistical Arbitrage - Black-Litterman	Equal weighted	Yes	Centred
SABL B1	Statistical Arbitrage - Black-Litterman	Exponentailly weighted moving averages	No	Uncentred
SABL B2	Statistical Arbitrage - Black-Litterman	Exponentailly weighted moving averages	No	Centred
SABL B3	Statistical Arbitrage - Black-Litterman	Exponentailly weighted moving averages	Yes	Uncentred
SABL B4	Statistical Arbitrage - Black-Litterman	Exponentailly weighted moving averages	Yes	Centred

This table describes the different portfolios, and the estimation techniques that they are based on.

[1]Equal weighted, GMVP A and GMVP B

Table E.3: Gross Portfolio Performance

Portfolio	OBX	Equal weight	GMVP A	GMVP B	SABL A1	SABL A2	SABL A3	SABL A4	SABL B1	SABL B2	SABL B3	SABL B4
Returns												
Annual HPR	0.102	0.039	0.018	0.015	-0.004	0.155	0.227	0.123	0.002	0.162	0.227	0.119
Daily max	0.116	0.147	0.100	0.113	0.298	0.215	0.250	0.515	0.285	0.241	0.244	0.515
Daily min	-0.107	-0.146	-0.089	-0.102	-0.332	-0.312	-0.164	-0.250	-0.194	-0.177	-0.159	-0.247
Yearly mean	0.142	0.092	0.047	0.046	0.108	0.270	0.352	0.245	0.111	0.274	0.352	0.237
Yearly median	0.277	0.332	0.102	0.085	0.263	0.327	0.389	0.359	0.314	0.441	0.329	0.348
Risk												
Ann.Volatility	0.259	0.311	0.234	0.243	0.460	0.425	0.429	0.452	0.454	0.420	0.429	0.446
Skewness	-0.325	-0.449	-0.048	0.088	-0.153	-0.516	0.110	1.538	0.286	-0.086	0.219	1.553
Kurtosis	6.655	8.831	4.153	5.206	14.627	11.167	7.071	39.753	11.738	7.772	6.615	41.018
MaxDD	0.648	0.676	0.624	0.628	0.734	0.699	0.642	0.731	0.779	0.674	0.581	0.738
Risk adjusted measures												
Ann.SR	0.371	0.078	-0.014	-0.026	-0.070	0.393	0.563	0.280	-0.054	0.415	0.568	0.272
Alpha	0	-0.050	-0.096	-0.096	-0.039	0.057	0.192	0.104	-0.039	0.073	0.192	0.075
Tracking Error	0	0.246	0.286	0.294	0.429	0.407	0.407	0.425	0.421	0.386	0.403	0.415
Information Ratio	0	-0.205	-0.334	-0.327	-0.092	0.142	0.471	0.245	-0.095	0.188	0.477	0.180

This table shows a range of performance measures for the trading strategy, including all approaches to apply it, including the benchmark (OBX) and an equal weighted portfolio in addition to global minimum variance portfolios. Results are reported *before* trading costs

Evidently, the difference between using EWMA and equal weights when estimating the covariance matrix is not big, which is reasonable given the short estimation window. More surprising is that the EWMA technique underperforms when the drift is included [2]. Compared to the benchmark, the reference portfolios yield negative alphas, and lower Sharpe ratios. The minimum variance portfolios also generate lower means and medians than the benchmark, but the equal weighted covariance approach is slightly better than the EWMA. Furthermore, the equal weighted yields a higher median return than the benchmark, indicating larger or more frequent drawdowns than the benchmark.

For the sake of brevity, the discussion of performance differences will be based on the B-class portfolios. Comparing these, the effect of centring means is significant when the drift is excluded, with a spread in holding period returns of 16 percentage points. While yielding lower returns, the centring of means stabilises the return process since the daily min and max are less extreme for B2 compared to B1. The latter portfolio's close-to-zero holding period return, high mean and median return and positive skew is likely the result of acute events, which can be seen from its relatively high kurtosis and max drawdown compared to B2. Risk adjusted measures indicate that, despite B2 outperforming the index based on holding period returns, it generates an information ratio below what is top tier, 0.5, according to Grinold and Kahn (2000, p.114).

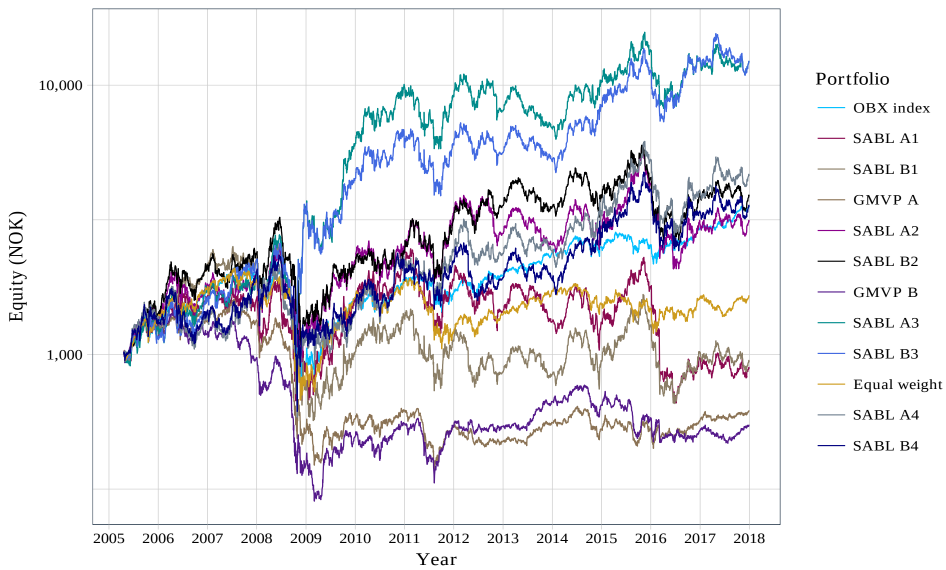
Including the drift, which is equivalent to incorporating a momentum strategy, further

[2] When comparing HPR for A4 and B4.

improves performance. The largest increase is between non-centred and centred means (B1 and B3). Moreover, when the drift is included, non-centred means yield the best results, while the opposite is true when the drift is excluded. This is surprising, since Avellaneda and Lee (2010) find that centred means is the superior approach. Although, that paper does not report results with an included drift, making it difficult to confirm if they see the same shift.

Including the drift and not centring the means (B3) performs best, with the same holding period return as A3, but slightly (possibly insignificant) better risk adjusted measures and lower volatility. Compared to B4, it yields a higher holding period return, lower volatility, is less skewed with a smaller kurtosis, generates a higher Sharpe ratio and alpha, lower tracking error and higher information ratio. Having a clear view of the performance measured by different metrics, the figure below, plots the cumulative returns before costs for all portfolios.

Figure E.2: The Portfolios Cumulative Returns



This panel plots the cumulative returns for all varieties of the SABL strategy, the reference portfolios and the benchmark portfolio. Cumulative returns are plotted before trading costs. The assumed initial capital is NOK 1000. Note that the y-axis is log-scale with base 10.

In figure E.2 it is visible how the SABL A/B3 portfolios generate high returns during the last part of 2008 and 2009, before growing at a slower rate. Hence, using non-centred means result in positions that mitigates the large market drawdowns during 2008, leaving A/B3 at a higher value by year-end, which is unique for these two portfolios. The root cause of the steep increase during these years is explored in section 5.

A reason for the divergence between centred and non-centres means might be that since the latter is a shrinkage estimator aimed at reducing variability at the expense of introducing bias, it has resulted in significant bias in the expected return estimates, which does not reflect the actual state for the securities under consideration. Alternatively, the reduction of variance could have yielded less extreme estimates, giving smaller positions and reducing returns when these positions turned out to be smart bets. Or it could simply be due to randomness.

In summary, the SABL B3 delivered the best performance measured by annualised holding period return and risk adjusted performance measures. Plotting the cumulative returns for all portfolios before costs, it is evident that a significant part of its value was accumulated between 2008 and 2011. The aim of the next section is to focus on this portfolio and the benchmark, by breaking performance down by year and sector, see if strong trading signals result in active positions and test for significance of returns.

E.3 Constituent Corrections

Table E.4: Irregular Changes to OBX Constituent List

Date	Company	Event	Adjustment
30.01.2006	Smedvig Serie A	Excluded from index, delisted as of 2006-06-18	Company removed
30.06.2006	Teekay Petrojarl	Listed on OSE and fast track entry into the OBX	Company included
20.10.2006	Teekay Petrojarl	Excluded from index	Company removed
20.03.2007	Tandberg Television	Excluded from index, sold to Ericsson	Company removed
15.08.2008	Awilco Offshore	Excluded from index, delisted	Company removed
02.12.2009	Tandberg ASA	Excluded from index, delisted as of 2010-04-28	Company removed
10.01.2011	Subsea 7	Acergy and Subsea 7 INC merged to form Subsea 7	Company and number of stocks changed
20.05.2012	Statoil Fuel and Retail	Excluded from index, delisted as of 2012-07-12	Company removed
30.08.2012	Golar LNG	Excluded from index, delisted at the same date	Company removed
24.02.2014	Algeta	Excluded from index, purchased by Aviator Acquisition	Company removed
29.09.2014	Aker	Demerger of Aker Solutions into Akastor and Aker Solutions Holdings	Company and number of stocks changed

This table document the irregular constituent changes in the OBX index between 2005 and 2017. What is shown is the date of the change, the companies that were affected, a description of the event that took place and the changes made.

Table E.5: Stock Splits and Reversals

Date	Company	Event	Ratio (new:old)
09.06.2005	Petroleum Geo Services	Stock split	3:1
23.06.2005	DNO	Stock split	4:1
10.05.2006	Norsk Hydro	Stock split	5:1
15.06.2006	TGS NOPEC Geophysical Company	Stock split	4:1
16.06.2006	DNO	Stock split	4:1
18.12.2006	Petroleum Geo Services	Stock split	3:1
27.12.2006	Prosafe	Stock split	5:1
30.03.2007	Aker Solutions	Stock split	5:1
20.04.2007	Orkla	Stock split	5:1
21.01.2014	Marine Harvest	Stock reversal	1:10
24.05.2017	Lerøy Seafood Group	Stock split	10:1

This table document all the companies performing either stock splits or reversals while being a constituent in the OBX index. The table states the event dates, the companies involved, if event was a split or reversal in addition to the conversion ratio with the number of new shares in the numerator and the number of old shares in the denominator.