University of Bergen

Department of Mathematics

Portfolio Optimization and Diversification Benefits - A Local Gaussian Correlation Approach

Author:
Sverre Hauso Haugen

Supervisor:
Bård Støve

Master's Thesis in Statistics

Financial Theory and Insurance Mathematics

June 3 2019

Abstract

Volatility and correlation among asset returns are central to portfolio allocation and risk management. The idea is simple; low correlated assets are good for diversification, while highly correlated assets should be avoided. However, as shown by numerous empirical studies, it is well documented that there are asymmetries in the distribution of financial returns. This is especially true when the market is going down, which often leads to stronger dependence between assets, known as asymmetric dependence structures. This effect is important for risk management and the performance of international portfolios. An increased dependence in financial markets during a crisis period, will imply that the diversification effect could be less than anticipated.

In this work we will examine the dependence patterns over time between different asset classes, e.g. developed markets, emerging markets, commodity market, bonds etc., by using the theory of local Gaussian correlation (LGC). We will examine how the LGC and adjustments in the covariance matrix primarily can be used to hopefully improve the classical Markowitz portfolio theory, and in particular examine how inclusion of different asset classes to a portfolio could increase the potential benefits of diversification. This is of particular importance for large institutional investors with a long investment horizon.

Keywords · Asymmetric Dependence Structures · Local Gaussian correlation · Portfolio Optimization

Acknowledgement

I would like to thank my supervisor Bård Støve for his great guidance and excellent supervision during the writing of this thesis. Furthermore, I would like to express my gratitude to my fellow students for many interesting discussions and the faculty members and staff at the Department of Mathematics for making the study period enjoyable and smooth.

Finally, I would like to thank family and friends for their moral support and words of encouragement throughout the study period.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Financial time series are continuously brought to our attention, and as an investment manager, it is of particularly interest to be able to monitor the behaviour of these financial movements. There are two main objectives of investigating financial time series. First, it is important to understand how different time series behave. Tomorrow's prices are uncertain and it must therefore be described by a probability distribution. Hence, it is intuitive to use statistical methods to investigate these prices. The second objective is to use our knowledge of the financial behaviour to be able to reduce the risk, and therefore make even better investment decisions. Risks associated with such frequent fluctuation in prices can be summarized by the variances of future returns, either directly, or by their relationship with relevant covariance in the context of portfolio optimization.

In investment decisions creating portfolios, the optimization is an important process of diversification. The classical approach today known as modern portfolio theory (MPT), first introduced by Markowitz [1952], aim to allocate financial instruments by exactly maximizing the portfolio's expected return while minimizing its risk. Hence, Markowitz [1952] provides the foundation for the Mean-Variance (MV) optimization framework. A crucial assumption in the optimization framework is that asset returns follow a joint-Gaussian distribution, and the dependence between these asset returns are fully described by the linear correlation coefficient (see eq.(A.0.1)). In the financial world, the linear correlation coefficient, $\rho_{X,Y}$, has been widely used as a measure of dependence for several reasons, but mainly because it is very easy to compute (only depends on the second moments (σ_X^2, σ_Y^2)). In general however, linear correlation is not a good measure of dependence, as emphasized by Embrechts et al. [2001] and Embrechts et al. [2002], since it only is able to capture the symmetric linear dependence in the market. Therefore it does not have the desired properties for dependence measures, that is, invariance under non-linear strictly increasing transformation. It can therefore in many cases be very misleading to use linear correlation as a measure of dependence for non-elliptical distributions.

The study of dependence structures in international equity markets, however, has recently attracted increasing attention among theorists and empirical researchers. For instance, to control the risks that portfolio managers and regulators face, they have to take into account dependence between international equity markets when studying the returns across international financial markets. Therefore, the issue of asymmetric dependence structures, such as high dependence in a certain period, is of particularly interest for risk control and policy management. In addition, benefits from international diversification of asset allocation could be considerably affected by the asymmetric dependence structure. An implication of the asymmetric dependence is that the classical MV optimized portfolios are not efficient with respect to their effective risk profile, thus the benefit of diversification will erode if the correlations are asymmetric. In other words, if the market is going down and the correlation increases, as is often loosely stated as a fact, this effect is true being a consequence of the non-linearity of the dependence structure between asset returns. Several studies have sought to overcome this shortcoming by modeling the dependence structure by using copula theory, and employing it in the optimization of a portfolio. However, these procedures are in most cases quite complicated if done properly, and there are still no guarantee that portfolio allocations based on complicated models like these will perform better (see e.g. Yew-Low et al. [2016]).

In this thesis however, we propose a much simpler approach. Without making any assumptions on the nature of the underlying probability model, we present an adjustment to the correlation matrix of the asset returns that takes into account the current state of the market that the manager subjectively expects over the next investment period. Our aim is to try to tackle the asymmetric dependence challenge, thus we would like to minimize the problem caused by the estimation errors. A possible approach is to impose a short-sale constraint, i.e. all weights are required to be non-negative. This causes the expected return across assets to shrink towards their mean (see DeMiguel and Uppal [2009]), and at the same time shrinks the extreme elements of the covariance matrix (see Jagannathan and Ma [2003]), which in turn decreases the estimation errors and

hence the problem of extreme weights. The purpose of this thesis is to investigate the dependence structures in international equity markets, and isolate the effect of covariance information from expected returns. We will further show that the portfolio optimization is straightforward using this new approach, only relying on a tuning parameter, the bandwidth. In this way we quantify and estimate the "increased correlation" effect within the existing MV framework non-parametrically by the use of LGC. For monthly equity returns, this new method is shown to substantially outperform the equally weighted $1/N$ portfolio and the classical Markowitz portfolio, hopefully. The reason for considering the equally weighted $1/N$ portfolio is that DeMiguel and Uppal [2009] show that it performs quite well even though it does not rely on any optimization procedure. Our proposed procedure relates to other studies that uses a dynamic model for the correlation matrix to improve the asset allocation (see e.g. Engle and Colacito [2006]and Aslanidis and Casas [2013]), but, as will become clear later in the analysis, our procedure differs by using the theory of LGC. Despite the fact that the concept of local dependence was introduced over 30 years ago, it is only recently that it has been used in financial and econometric literature.

Furthermore, this thesis is divided into four parts. Following this introductory part is Part II, which presents background information necessary for the empirical analysis. The organization in Part II is structured as follows; Chapter 2 gives a brief introduction to Financial Returns, Chapter 3 describes the GARCH(p,q) model, Chapter 4 describes the importance of Risk Management, Chapter 5 gives a thorough introduction to Modern Portfolio Theory and Chapter 6 describes the Local Gaussian Correlation in details. After this, part III discusses Portfolio Optimization using LGC. This part is divided into three sub-chapters, and discusses computational challenges and approaches as well as conclusions. Section 7.1 describes the methodology used in this thesis in details. Section 7.2 briefly describes the asset allocation models investigated, and Section 7.3 applies the proposed analysis to stocks portfolio and shows the benefits that may accrue from this approach by using simulated series with the same characteristics as real data, assuming asymmetric dependence structure, in other words, an increased correlation as the market is going down. Part IV provides the empirical results. This part is divided into two sub-chapters. These begin with Chapter 8.1, which covers the descriptive statistics related to the dataset used in this thesis, and ends with Chapter 8.2, which give us an insight into the Empirical Analysis of the dependence structures and provides the empirical Results. The analysis presented in chapter 8 is performed by use of libraries, and the scripts are mainly written in R, a programming language for statistical computing and graphics. Chapter 9 is a summary of this thesis, including the conclusions we have drawn from our research and suggests several ideas for related future work. Following these concluding chapters are several appendices. Concluding this thesis is the bibliography.

Chapter 2

Financial Returns

In the financial world, the goal of investing is of course to make a profit, and the profit from investing, depends upon both the changes in prices and the amounts of assets being held. Return measure this, and in order to do a proper analysis of our data, we need to convert price data into values that can be modeled by a statistical distribution. If we can make the assumption that returns are stationary, that is, future returns will be similar to past returns, then we can apply the machinery of statistical inference and the probability distribution of P_t can be estimated from historical data. Hence, the uncertainty is unmeasurable if we are unwilling to make these assumptions, thus finance, as most of the sciences, is in a grey zone between measurable and unmeasurable uncertainty. In this section, we will introduce some of the most common used return measures in the context of financial analysis. The theory in this chapter is based on the textbooks Ruppert and Matteson [2015] and Anderson-Cook [Anderson-Cook], which offers a thorough introduction in financial returns and time series analysis. The texts also provide references for further reading.

2.1 Net-return

Consider discrete observations P_t of some asset price. The most common transformation yields arithmetic returns, defined as

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{P_t}{P_{t-1}} - 1, \tag{2.1}$$

where P_t is the price of the financial instrument at time t. The numerator, $P_t - P_{t-1}$, is the profit during the holding period, with negative profit meaning a loss. The denominator, P_{t-1}, is the initial investment at the start of the holding period, and therefore, the net-return can be viewed as the relative profit rate. Based on the returns at time t, the aggregation of daily returns over the period T can be expressed as the product of single-period net returns

$$r_T = \frac{P_T}{P_0} - 1 = \frac{P_t}{P_{t-1}} \times \frac{P_{t-1}}{P_{t-2}} \times \cdots \times \frac{P_1}{P_0} - 1 = \prod_{t=1}^{T} \frac{P_t}{P_{t-1}}. \tag{2.2}$$

2.1.1 Gross-return

Another common transformation is the simple gross return, defined as

$$1 + r_t = \frac{P_t}{P_{t-1}}. \tag{2.3}$$

The gross-returns are scale-free, meaning that they do not depend on units, although they depend on the unit of time t.

2.2 Log-return

Log-returns are denoted by r_t and defined as

$$r_t = log\Big(\frac{P_t}{P_{t-1}}\Big) = log(P_t) - log(P_{t-1}) = p_t - p_{t-1}, \tag{2.4}$$

where $p_t = log(P_t)$ is called the "log-price", and $r_t = log(P_t) - log(P_{t-1})$ correspond to the log-return differences of the asset prices at time t. Since returns are smaller in magnitude over shorter periods, we can

expect net-returns and log-returns to be similar for daily returns, less similar for monthly returns, and not necessarily similar for longer periods such as years. Assuming no dividends, the net-return over the holding period from time $t-1$ to time t is given by

$$r_t = (log(P_t) - log(P_{t-1})) \approx \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{P_t}{P_{t-1}} - 1.$$

The net-return and log-return have the same sign, though the magnitude of the log-return is smaller (larger) than that of the return if they are both positive (negative). One advantage of using log-returns is simplicity of multiperiod returns. A T period log-return is simply the sum of the single period log-returns, rather than the product as for net-returns. To see this, note that the T period log-return can be expressed as

$$r_T = log\left(\frac{P_t}{P_{t-1}}\right) + \cdots + log\left(\frac{P_1}{P_0}\right) = r_t + r_{t-1} + ... + r_1 = \sum_{t=1}^{T} r_t. \tag{2.5}$$

2.3 Adjustment for Dividends

Many stocks, especially those of mature companies, pay dividends that must be accounted for when computing returns. Similarly, bonds pay interest. If a dividend D_t is paid prior to time t, then the net-return at time t is defined as

$$r_t = \frac{P_t + D_t}{P_{t-1}} - 1, \tag{2.6}$$

and so the gross-return is

$$1 + r_t = \frac{P_t + D_t}{P_{t-1}}, \tag{2.7}$$

and the log-return

$$r_t = log\left(\frac{P_t + D_t}{P_{t-1}}\right). \tag{2.8}$$

$D_t = 0$ if there is no dividend between $t-1$ and t. Multiple-period net-return are defined as in eq.(2.2), that is

$$r_T = \frac{P_t + D_t}{P_{t-1}} \times \frac{P_{t-1} + D_{t-1}}{P_{t-2}} \times \cdots \times \frac{P_1 + D_1}{P_0} = \prod_{t=1}^{T} \frac{P_t}{P_{t-1}}. \tag{2.9}$$

Similarly, a multiple-period log-return are defined as in eq.(2.5), given by

$$r_T = log\left(\frac{P_t + D_t}{P_{t-1}}\right) + \cdots + log\left(\frac{P_1 + D_1}{P_0}\right) = \sum_{t=1}^{T} r_t. \tag{2.10}$$

2.4 Stylized Facts

At the heart of time series analysis is the identification of patterns within the stochastic processes that influences the data. These patterns are often referred to as the *stylized facts*, that is;

1. the return series are not identically independently distributed;
2. the absolute or squared valued returns are highly correlated;
3. the volatility appears to change randomly with time;
4. the distribution of the data is heavy-tailed and asymmetric, and therefore not Gaussian;
5. the extremes appear in clusters.

The following section focus on the technique used for modeling financial returns with time-varying volatility, that is, the widely used GARCH(p,q) model.

Chapter 3

GARCH

The GARCH models are mostly important in time series analysis, and particularly in financial applications where the goal is to analyze and forecast volatility. For that reason the variable of interest might be of the type as we described in Chapter 1. The number of these so-called heteroskedastic[1] models are immense, but the earliest models remain the most influential. In this section we will describe the classical GARCH(p,q) model, first introduced by Bollerslev [1986], and give a short introduction to the estimation and forecasting procedure of the more popular univariate GARCH(1,1) time-series model for volatility. The theory in this chapter is mostly based on the textbooks Aas and Dimakos [2004], Posedel [2005], Anderson-Cook [Anderson-Cook], Reider [2009], Ruppert and Matteson [2015] and Charpentier [2015] which offers a thorough introduction in financial returns and risk management. The texts also provide references for further reading.

3.1 The GARCH(p,q) process

In periods of stress, financial assets tend to fluctuate a lot. This means, statistically speaking, that the conditional variance for the given past is not constant over time, and the process x_t is therefore conditionally heteroskedastic. Econometricians usually say that volatility, which is expressed by

$$\sigma_t = \sqrt{Var(x_t|x_{t-1}, x_{t-2}, \dots)}, \tag{3.1}$$

changes over time. Understanding the nature of such time dependence is very important for many macroeconomic and financial applications, e.g. asset pricing. Models of conditional heteroscedasticity for time series have a very important role in today's financial risk management as it attempts to make financial decisions on the basis of the observed price asset data P_t in discrete time. A deficiency of the ARCH[2] models is that the conditional standard deviation process has high frequency oscillations with high volatility coming in short bursts. The GARCH models permit a wider range of behavior, in particular, it uses values of the past squared observations and past variances to model the variance at time t with more persistence into the volatility. We can describe the mean equation of a univariate time series x_t as expressed by the equation

$$x_t = E[x_t|\Omega_{t-1}] + \epsilon_t, \tag{3.2}$$

where $E[x_t|\Omega_{t-1}]$ denotes the conditional expectation operator, Ω_{t-1} denotes the information set at time $t-1$, and ϵ_t denotes the residuals, or the unpredictable part of the time series, and describes the uncorrelated disturbances with zero mean. Engle [1982] defined the residuals of the time series process as an auto-regressive conditional heteroskedastic process where all the ϵ_t terms takes the general form

$$\epsilon_t = z_t\sqrt{\sigma_t^2}, \tag{3.3}$$

where ϵ_t follows a strict white noise process with the constraint

$$z_t \sim iid\ \mathcal{D}_\theta(0, 1). \tag{3.4}$$

Here, θ represents additional distributional parameters. As for the Student t-distribution, this would be the degrees of freedom. The variance equation of the GARCH(p,q) process is defined as

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \tag{3.5}$$

[1] Recall that heteroscedasticity is just a fancy way of saying non-constant variance. Homoscedasticity means constant variance.
[2] A formal definition of the ARCH model, see Engle [1982].

The necessary and sufficient condition for stationarity with finite (non-negative) variance is given by

$$\sum_{i=1}^{r} \alpha_i + \sum_{j=1}^{s} \beta_j < 1, \tag{3.6}$$

where $\alpha_i \geq 0$ and $\beta_i \geq 0$. If all the coefficients β_j are equal to zero, the GARCH model is automatically reduced to the more general ARCH model. Conditions for strict stationarity and existence of the moments are given in Ling and Li [1997] and Ling and McAleer [2001].

Definition 3.1.1. *The more popular and most used GARCH(1,1)-model is defined as*

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2, \tag{3.7}$$

which follow the same constraints as the GARCH(p,q) model.

3.2 Standardized Student t-distribution

Bollerslev [1986] was the first to model financial time series for foreign exchange rates and stock indexes using the GARCH(1,1) models (see eq.(3.3)), extended by the use of the standardized Student t-distribution. When compared with conditionally normal errors, he found that the t-GARCH(1,1) errors much better captured the leptokurtosis seen in the data. Previous studies have shown that the Student t-distribution performs well in capturing the observed kurtosis in empirical log-return series. For more details, see Bollerslev [1986], Bollerslev et al. [1992] and Charpentier [2015].

Definition 3.2.1. *The density $f^*(z|\nu)$ of the standardized Student t-distribution can be expressed as follows;*

$$\begin{aligned} f^*(z|\nu) &= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi(\nu-2)}\Gamma\left(\frac{\nu}{2}\right)} \times \frac{1}{\left(1+\frac{z^2}{\nu-2}\right)^{\frac{\nu+1}{2}}} \\ &= \frac{1}{\sqrt{\nu-2}B\left(\frac{1}{2},\frac{\nu}{2}\right)} \times \frac{1}{\left(1+\frac{z^2}{\nu-2}\right)^{\frac{\nu+1}{2}}}, \end{aligned} \tag{3.8}$$

where $\nu > 2$ is the shape paraeter and $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the Beta function.

3.2.1 Maximum Likelihood Estimation

Given the model for conditional mean and variance, we use the maximum likelihood estimation approach to fit the parameters for the specified model of the return series as done in Charpentier [2015]. The procedure infers the process innovations or residuals by inverse filtering. This filtering transforms the observed process ϵ_t into an uncorrelated white noise process z_t to infer the corresponding conditional variances σ_t^2 via recursive substitution into the model-dependent conditional variance equations. Finally, the procedure uses the inferred innovations and conditional variances to evaluate the appropriate log-likelihood objective function. The maximum likelihood estimator concept interprets the density as a function of the parameter set, conditional on a set of sample outcomes. Using eq.(3.3), the log-likelihood function of the distribution $\mathcal{D}_\theta$ can be expressed as

$$l_N(\theta) = log \prod_t \mathcal{D}_\theta(x_t, E(x_t|\Omega_{t-1}), \sigma_t), \tag{3.9}$$

where $\mathcal{D}_\theta$ is the conditional distribution function. The second argument of $\mathcal{D}_\theta$ denotes the mean, and the third argument the volatility. For Bollerslev's Student-t GARCH(1,1)-model, the parameter set reduces to $\theta = (\mu, \omega, \alpha, \beta, \nu)$.

Chapter 4

Risk Management

Financial risks are often related to as unpredictable movements in financial variables, such as negative development in financial markets. A large part of risk management is the potential of measuring such future losses of a portfolio of assets, and in order to measure these potential losses, estimates must be made of future volatilities and correlations. The simplest approach to estimate volatility is to use historical standard deviation. As a recap, remember that volatility clustering is described as large changes that tends to be followed by large changes and small changes that tends to be followed by small changes, of either sign. In the context of finance, risks are classified into four main types to be considered, namely

1. market risk - risk associated with fluctuations in prices of traded financial assets;
2. credit risk - risk associated with uncertainty that debtors will honour their financial obligations;
3. liquidity risks - risk of an investment that cannot be bought or sold quickly enough to minimize loss;
4. operational risk - risk associated with possibility of human error, IT failure, dishonesty etc.

Value-at-Risk (VaR) and Expected Shortfall (ES) has become widely used because they can be applied to all these types of risks and securities, including complex portfolios. Value-at-Risk is today considered as the standard risk measure by financial investors. Let's denote the value of a portfolio at time t by $\mathcal{V}_t$, and assume that the random variable $\mathcal{V}_t$ is observable at time t. For a given time horizon Δ, the loss of the portfolio over the period $[\Delta_t, \Delta_{(t+\ell)}]$ can be expressed as

$$\mathcal{L}_{t+\ell} := \mathcal{L}_{[\Delta_t, \Delta_{(t+\ell)}]} = -(\mathcal{V}_{t+\ell} - \mathcal{V}_t). \tag{4.1}$$

While $\mathcal{L}_{[\Delta_t, \Delta_{(t+\ell)}]}$ is assumed to be observable at time $\Delta_{(t+\ell)}$, it is typically random from the viewpoint of Δ_t. The distribution of $\mathcal{L}_{[\Delta_t, \Delta_{(t+\ell)}]}$ is then termed as the loss distribution of our investment. In risk management, we are mainly concerned with the probability of large losses and hence with the upper tail of the loss distribution. The random variables $\mathcal{V}_t$ and $\mathcal{V}_{t+\ell}$ represents the portfolio value, and $\mathcal{L}_{t+\ell}$ is the loss from this given investment. In the following sections, particular attention will be given to Value-at-Risk and the related notion of the Expected Shortfall. The theory in this chapter is based on the textbooks McNeil [2005], Ruppert and Matteson [2015] and Charpentier [2015], which offers a thorough introduction in financial returns and time series analysis. The texts also provide references for further reading.

4.1 Value-at-Risk and Expected Shortfall

Definition 4.1.1. *Consider some portfolio of risky assets and a fixed time horizon Δ. Denote $F_{\mathcal{L}}(\ell) = P(\mathcal{L} \leq \ell)$ and the confidence level $\alpha \in (0, 1)$. Consider $\mathcal{L}$ as the loss over the holding period Δ, then for any loss distribution, VaR_α is the α^{th} upper quantile of $\mathcal{L}$. Hence,*

$$VaR_\alpha = inf\{\ell \in R : P(\mathcal{L} > \ell) \leq 1 - \alpha\} = inf\{\ell \in R : F_{\mathcal{L}}(\ell) \geq \alpha\} \tag{4.2}$$

is simply the maximum expected loss of a portfolio over the time horizon Δ with certain confidence level α.

Definition 4.1.2. *For any loss distribution $\mathcal{L}$, continuous or not, the expected shortfall at confidence level α is defined as*

$$ES_\alpha = \frac{1}{1-\alpha} \int_\alpha^1 q_u(F_{\mathcal{L}}) du, \tag{4.3}$$

where $q_u(F_{\mathcal{L}}) = F_{\mathcal{L}}^{\leftarrow}(u)$ is the quantile function of $F_{\mathcal{L}}$.

Hence, the expected shortfall is related to VaR_α by the equation

$$ES_\alpha = \frac{1}{1-\alpha}\int_\alpha^1 VaR_u(\mathcal{L})du, \tag{4.4}$$

which is the average of VaR over all levels of $u \geq \alpha$, and thus look further into the tail of the loss distribution.

4.1.1 Coherent risk properties

A serious deficiency of VaR is that it may discourage diversification, and for that reason it is being replaced by newer risk measures such as the Expected Shortfall. This problem was studied by Artzner [1996], who showed that VaR is not a coherent risk measure since it violates the property of subadditivity, which they believe that reasonable risk measures should have. Thus, a risk measure η is coherent if, for any two loss random variables X and Y satisfies the following axioms;

- Axiom 1 - Subadditivity
  For all $X, Y \in \mathcal{N}$, we have $\eta(X+Y) \leq \eta(X) + \eta(Y)$.

  This axiom is the most debated of the four characterizing coherent risk measures, probably because it rules out VaR as a risk measure in certain situations. An argument explaining why subadditivity indeed is a reasonable requirement, is that subadditivity reflects the idea that risk can be reduced by diversification.

- Axiom 2 - Monotonicity
  For all $X, Y \in \mathcal{N}$, such that $X \leq Y$ almost surely, then $\eta(X) \leq \eta(Y)$.

  From an economic point of view this axiom is obvious, since positions that lead to higher losses in every state of the world require more risk capital.

- Axiom 3 - Positive homogeneity
  For all $X \in \mathcal{N}$ and every $c > 0$, we have $\eta(cX) = c\eta(X)$.

  Axiom 3 is easily justified if we assume that axiom 1 holds. Subadditivity implies that, for $n \in \mathcal{N}$,

  $$\eta(nX) = \eta(X + \cdots + X) \leq n\eta(X). \tag{4.5}$$

  Since there is no diversification between the losses in the portfolio, it is natural to require that equality should hold in eq.(4.5), which leads to positive homogeneity.

- Axiom 4 - Translation invariance

For all $X \in \mathcal{N}$, and every $c \in \mathcal{M}$, we have that $\eta(X + c) = \eta(X) + c$.

This axiom states that by adding or subtracting a deterministic quantity c to a position leading to the loss X, we alter our capital requirements by exactly that amount. This axiom is in fact necessary for the risk-capital interpretation of η to make sense.

4.1.2 Estimation of VaR Based on POT and the GARCH Model

In the context of financial time series, modeling the lower tail of the return distributions is the primary focus of risk measures. A good example is the VaR_α. In this context, the extreme value theory (EVT) becomes of very high interest, as it consists of modeling the tail of the distributions. As introduced in the Appendix, one of the approaches used in EVT is the peak-over-threshold (POT) method. This approach finds its root in the EVT theorem, which states that when one selects a threshold high enough, the distribution of the values exceeding that threshold converges in distribution to the generalized Pareto distribution (GPD, pdf see eq.(A.3)). We follow the same procedure as in Charpentier [2015]. The estimation of VaR based on the POT method can then be implemented as follows; we first count the number of points that exceed the threshold u, and by default, we keep the default threshold at 5%. Second, we estimate the parameters of the GPD by means of the MLE, and calculate the Var based on eq.(A.4). We want to combine the POT method from EVT together with the GARCH(1,1)-model as done by McNeil and Frey [2000] (see Charpentier [2015]). Consider the univariate time series model as expressed by eq.(3.2). Then the one-step VaR forecast is represented by

$$VaR_\alpha = \mu_{t+1} + \sigma_t VaR_\alpha(Z). \tag{4.6}$$

Chapter 5

Modern Portfolio Theory

The Mean-Variance (MV) portfolio is a widely used method in the aspect of portfolio optimization framework and was first introduced by Markowitz [1952], which paved the way for modern portfolio optimization. It was further developed in the following decades by the economists William Sharpe and Merton Miller. The optimization approach, however, was neither used for making precise predictions, but rather used to understand how different asset classes generally behave over time. It establishes a range of reasonable expectations which can then be used to make slightly better informed investment decisions. The modern portfolio theory (MPT) is based upon two principles, that is,

1. maximize the expected return; and
2. minimize the risk, which we define to be the standard deviation of the return.

The optimization approach provides the fundamental basis for investors to have diversified portfolios, and it either assumes that portfolios are normally distributed or that investors exhibit quadratic utility preferences. Based on what we know about a normal distribution we can say that the standard deviation measures the magnitude of a certainty or a volatility of potential outcomes. The optimization goal is to determine the best trade-off between risk and return, subject to a set of constraints. Further, the MV model assumes the following; (1) prices of the instruments are exogenous and given; (2) returns follow stochastic processes that are elliptically distributed in probability space, meaning that a covariance matrix exists; (3) there are no tax or other costs to transaction; (4) markets for all assets are liquid; (5) assets are infinitely divisible; and (6) full investment is required. In this chapter we will focus on the portfolio optimization process, the diversification effect and potential benefits. The theory in this chapter is based on the textbooks Markowitz [1952], Ruppert and Matteson [2015], Anderson-Cook [Anderson-Cook] and Charpentier [2015], which offers a thorough introduction in financial analysis and portfolio theory. The texts also provide references for further reading.

5.1 Markowitz [1952]

The first way of managing portfolio selection was divided into two stages. The first stage started with observation and experience of the data, and ended with beliefs about the future performances of available securities. The second stage started with the relevant beliefs about future performances, and ended with the choice of the portfolios. We will concentrate on the second stage. We can see this analytically; let's suppose there are $N \in \mathbb{N}$ asset classes, where N is the number of assets. Then let the anticipated return on the risky assets, $\boldsymbol{r}_t \in \mathbb{R}^N$, which are expected to be normally distributed with expected value $\boldsymbol{\mu}_t \in \mathbb{R}^N$, and covariance matrix $\boldsymbol{\Sigma}_t \in \mathbb{R}^N \times \mathbb{R}^N$ of the portfolio of asset returns at time t. Further, let $\boldsymbol{\omega} \in \mathbb{R}^N$ be the vector of optimal portfolio weights. We know that $\boldsymbol{\Sigma}$ is semi positive definite by definition. We assume it to be positive definite.

Definition 5.1.1. $\boldsymbol{\Sigma}^{-1}$ *is positive definite.*

Proof. Let $z \in \mathbb{R}^N$, and $x = \Sigma^{-1}z$. Then

$$z^T\Sigma^{-1}z = (\Sigma\Sigma^{-1}z)^T\Sigma^{-1}z = (\Sigma^{-1}z)^T\Sigma^T\Sigma^{-1}z = x^T\Sigma^T x = x^T\Sigma x > 0. \tag{5.1}$$

□

In reality, no asset class is an exact linear combination of other asset classes, so this is a reasonable assumption. Furthermore, there is a rule in which implies that the investor should diversify and maximize the expected return. The rule states that the investor should diversify his funds among all those securities which

give maximum expected return. It assumes that there is a portfolio which gives both maximum expected return and minimum variance, and it commends this portfolio to the investor. Although diversification cannot eliminate all variance, there is a rate at which the investor can gain expected return by taking on variance, or reduce variance by giving up expected return. Before presenting the portfolio models, we first review different types of constraints that are implemented, and which correspond to the canonical optimization problems in the following sections. The first constraint is set by the goal to achieve the target reward measure. The **target reward** constraint, or the expected return μ_P for a portfolio with weights ω, is defined as

$$\mu_P = \sum_{i=1}^{N} \omega_i \mu_i = \boldsymbol{\omega}^T \boldsymbol{\mu}, \tag{5.2}$$

which is also known as the target value of the portfolio. The real portfolio return however, which we assume to be normally distributed with $r_P \sim \mathcal{N}(\mu_p, \sigma_P^2)$, is given by

$$r_P = \boldsymbol{\omega}^T r. \tag{5.3}$$

If we suppose there is a target value μ_P of the expected return on the portfolio, then for $N = 2$, μ_P is achieved by only one portfolio and its ω_1 value solves

$$\mu_P = \omega_1 \mu_1 + \omega_2 \mu_2 = \mu_2 + \omega_1(\mu_1 - \mu_2).$$

For $N \geq 3$, there will be an infinite number of portfolios achieving the target μ_P. The one with the smallest variance is called the "efficient" portfolio, which is our goal to achieve. The expected portfolio risk measure developed by Markowitz is an asset-weighted covariance matrix,

$$\sigma_P^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} \omega_i \omega_j \Sigma_{i,j} = \boldsymbol{\omega}^T \boldsymbol{\Sigma} \boldsymbol{\omega}, \tag{5.4}$$

where $\boldsymbol{\Sigma}$ is the covariance matrix and $\boldsymbol{\omega}$ are the given portfolio weights. Given a target value μ_P, an efficient portfolio minimizes the risk in eq.(5.4) subject to the constraints

$$\boldsymbol{\omega}^T \boldsymbol{\mu} = \mu_P, \tag{5.5}$$

and

$$\boldsymbol{\omega}^T \mathbf{1} = 1. \tag{5.6}$$

We assume that[1] $\omega_1 + \cdots + \omega_N = \mathbf{1}^T \boldsymbol{\omega} = \mathbf{1}$, where $\mathbf{1} \in \mathbb{R}^N$. This constraint states that all capital must be invested in the portfolio (**full investment** constraint), where the weights correspond to portions of the capital allocated to a given component. Another type of constraint is related to **long only** positions, which specify that we can only buy shares and therefore only have position-related weights in contrast to the case of short positions, in which the selling positions would be reflected as negative weights.

5.1.1 Optimization framework

In this section we adopt the general formulation for portfolio optimization, which consists of minimization of a risk measure given a target reward and operational constraints. The MV optimization problem is defined as follows; the weights of the chosen portfolio are given by a vector $\boldsymbol{\omega}_t$, that is invested in N risky assets; and the investor selects $\boldsymbol{\omega}_t$ to maximize the expected quadratic utility function at each time t, that is

$$\max_{\boldsymbol{\omega}_t} U = \boldsymbol{\omega}_t^T \boldsymbol{\mu}_t - \frac{\gamma}{2} \boldsymbol{\omega}_t^T \boldsymbol{\Sigma}_t \boldsymbol{\omega}_t, \tag{5.7}$$

where U is the investor's utility, and γ represents the investor's degree of risk aversion. Throughout this thesis, the risk aversion coefficient γ is fixed and equal to 1. Hence, for a range of different risk aversion levels, the MV optimization will produce corresponding optimal portfolios with a trade-of between expected volatility and expected return. There are three different ways of formulating the MV optimization explicit as in the following definitions. We will show how to minimize the risk subject to a lower bound on the expected return (def. 5.1.2), maximize the expected return subject to an upper bound on the risk (def. 5.1.3), and a third option which optimizes the corresponding ratio between risk and return subject to a given level of risk aversion (def. 5.1.4). One key concept that will be discussed is reduction of risk by diversifying the portfolio of assets held.

[1] Recall that for any two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, their inner product is $\boldsymbol{a}^T \boldsymbol{b} = \sum a_i b_i$. Therefore, $\boldsymbol{a}^T \mathbf{1} = \sum a_i$.

Definition 5.1.2. *The minimum risk formulation*

$$\begin{array}{lll} \min\limits_{\boldsymbol{\omega}\in\mathbb{R}^N} & \boldsymbol{\omega}^T\Sigma\boldsymbol{\omega} & \textit{Covariance Risk,} \\ \textit{subject to} & \boldsymbol{\omega}^T\boldsymbol{\mu} = \mu_P, & \textit{Target Return} \\ & \boldsymbol{\omega}^T\mathbf{1} = 1, & \textit{Full Investment} \\ & \boldsymbol{\omega} \geq 0, & \textit{Long Only Positions.} \end{array} \tag{5.8}$$

Definition 5.1.3. *The maximum return formulation*

$$\begin{array}{lll} \max\limits_{\boldsymbol{\omega}\in\mathbb{R}^N} & \boldsymbol{\omega}^T\boldsymbol{\mu} & \textit{Target return,} \\ \textit{subject to} & \boldsymbol{\omega}^T\boldsymbol{\Sigma}\boldsymbol{\omega} = \sigma_P^2, & \textit{Expected portfolio risk} \\ & \boldsymbol{\omega}^T\mathbf{1} = 1, & \textit{Full Investment} \\ & \boldsymbol{\omega} \geq 0, & \textit{Long Only Positions,} \end{array} \tag{5.9}$$

Definition 5.1.4. *Let $\gamma \in [0,\infty)$, the risk aversion formulation*

$$\begin{array}{lll} \min\limits_{\boldsymbol{\omega}\in\mathbb{R}^N} & \gamma\boldsymbol{\omega}^T\Sigma\boldsymbol{\omega} - \boldsymbol{\omega}^T\boldsymbol{\mu} & \textit{Risk aversion,} \\ \textit{subject to} & \boldsymbol{\omega}^T\mathbf{1} = 1, & \textit{Full Investment} \\ & \boldsymbol{\omega} \geq 0, & \textit{Long Only Positions,} \end{array} \tag{5.10}$$

The term $\omega \geq 0$ prevents the weights from getting negative values. In financial terms, negative portfolio weights are called *short-selling*. This is basically a risky way to profit from declining stock prices, that is, one sells the stocks without actually owning it. In other words, the trader sells to open a short position and expects to buy it back later at a lower price, and will keep the difference as a gain. This closes the short position[1].

Definition 5.1.5. *A weight vector $\boldsymbol{\omega}$ is called efficient, if $\boldsymbol{\omega}$ is a solution to*

- *(1) some $\mu_P \geq \mu_{min}$, or*
- *(2) some $\sigma_P^2 \geq \sigma_{min}^2$, or*
- *(3) some $\gamma \in [0,\infty)$,*

where μ_{min} is the expected return of the minimum variance portfolio.

The investor has a choice of various combinations of μ_P and σ_P depending on his choice of portfolio.

Definition 5.1.6. *The efficient frontier is represented by*

$$\{(\sigma_P^2, \mu_P) : \sigma_P^2 = \boldsymbol{\omega}^T\boldsymbol{\Sigma}\boldsymbol{\omega}, \mu_P = \boldsymbol{\omega}^T\boldsymbol{\mu}\}. \tag{5.11}$$

- *Vary μ_P along a grid, then for each value of μ_P on this grid, we compute σ_{μ_P} by*
  a) computing $\boldsymbol{\omega}_{\mu_P} = \boldsymbol{g} + \mu_P\boldsymbol{h}$; and
  b) then computing $\sigma_{\mu_P} = \sqrt{\boldsymbol{\omega}_{\mu_P}^T\boldsymbol{\Sigma}\boldsymbol{\omega}_{\mu_P}}$.
- *Plot the vales (μ_P, σ_{μ_P}). The values (μ_P, σ_{μ_P}) with $\mu_P \geq \mu_{min}$ as the efficient frontier. The other values of (μ_P, σ_{μ_P}) which lies below the efficient frontier are called inefficient portfolios.*

The efficient frontier, which is represented by the hyperbola in Figure 5.1, represents the portfolios with the greatest expected level of return given any level of risk, or their lowest level of risks given their level of return. All portfolio combinations on the efficient frontier are therefore called efficient portfolios. As an investor, we should only want the leftmost point on this locus, since this achieves the minimum value at risk and is called the global minimum variance portfolio. Every portfolio combination is illustrated on the right hand side of this hyperbola. A well known problem in the MV framework is that we can obtain optimal portfolios that are little diversified, and the main reason for wanting the MV method to be more robust is that we do not want to change a portfolio drastically due to small changes in estimated parameters. But why? In the following section we present the efficient-portfolio mathematics.

[1] In order to open a short position, a trader must have a margin account and will usually have to pay interest on the value of the borrowed shares while the position is open. A deitailed description of short-selling is given in any textbook on financial analysis.

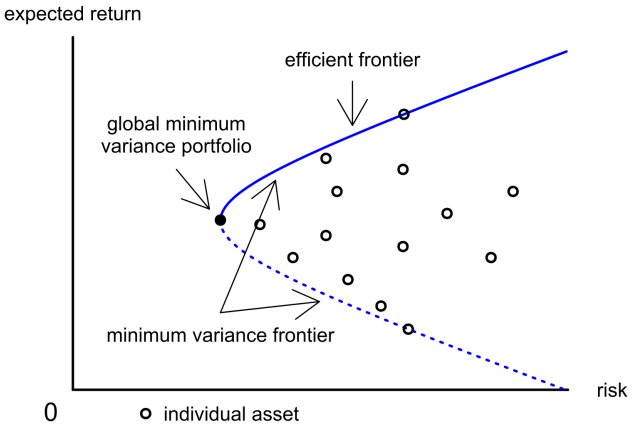


Figure 5.1: A graphical illustration of the efficient frontier, together with the global minimum-variance portfolio, minimum-variance frontier and the feasible set as defined above. Source; Munasca [2015].

5.2 Efficient-portfolio Mathematics

Let's denote the weights of the efficient portfolio by $\boldsymbol{\omega}_{\mu_P}$, and to find $\boldsymbol{\omega}_{\mu_P}$, we need to form the Lagrangian[2],

$$L(\boldsymbol{\omega}, \delta_1, \delta_2) = \boldsymbol{\omega}^T \boldsymbol{\Sigma} \boldsymbol{\omega} + \delta_1(\mu_P - \boldsymbol{\omega}_{\mu_P}^T \boldsymbol{\mu}) + \delta_2(1 - \boldsymbol{\omega}^T \mathbf{1}), \tag{5.12}$$

where λ_1 and λ_2 are the Lagrangian multipliers. Next, derive the first-order conditions,

$$\begin{aligned} 0 &= \frac{\delta}{\delta \boldsymbol{\omega}} L(\boldsymbol{\omega}, \delta_1, \delta_2) = 2\boldsymbol{\Sigma}\boldsymbol{\omega}_{\mu_P} - \delta_1 \boldsymbol{\mu} - \delta_2 \mathbf{1} \\ 0 &= \frac{\delta}{\delta \delta_1} L(\boldsymbol{\omega}, \delta_1, \delta_2) = \mu_P - \boldsymbol{\omega}^T \boldsymbol{\mu} \\ 0 &= \frac{\delta}{\delta \delta_2} L(\boldsymbol{\omega}, \delta_1, \delta_2) = 1 - \boldsymbol{\omega}^T \mathbf{1}, \end{aligned} \tag{5.13}$$

and solve for ω in terms of λ_1 and λ_2. The solution will be a function of λ_1 and λ_2, and these will be determined from the constraints in eq.(5.5) and eq.(5.6). In order to solve eq.(5.13), see definition 5.2.1.

Definition 5.2.1.

$$\frac{\delta}{\delta \boldsymbol{\omega}} L(\boldsymbol{\omega}, \delta_1, \delta_2) = \begin{pmatrix} (\boldsymbol{\omega}, \delta_1, \delta_2)/\delta\omega_1 \\ \vdots \\ (\boldsymbol{\omega}, \delta_1, \delta_2)/\delta\omega_N \end{pmatrix} \tag{5.14}$$

is the gradient of L with respect to $\boldsymbol{\omega}$ *with all other variables held fixed.*

In our analysis, we will use quadratic programming to find efficient portfolios with an arbitrary number of assets. An advantage of quadratic programming is that it allows one to impose constraints such as limiting short sales. In the following calculations we use the minimum risk formulation (5.1.2). Also note that we can invert the covariance matrix $\boldsymbol{\Sigma}$ because we assume that it is positive definite. The efficient portfolio weights, $\boldsymbol{\omega}$, produced by MV optimization satisfies the following equation

$$\boldsymbol{\omega}_{\mu_P} = \boldsymbol{g} + \boldsymbol{h}\mu_P. \tag{5.15}$$

Notice that $\boldsymbol{g}$ and $\boldsymbol{h}$ are fixed vectors, since they depend on the fixed vector $\boldsymbol{\mu}$ and the fixed matrix $\boldsymbol{\Sigma}$, but independent of μ_P, and hence, are given by

$$\begin{aligned} \boldsymbol{g} &= \frac{1}{d}(a\boldsymbol{\Sigma}^{-1}\mathbf{1} - b\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}) \\ \boldsymbol{h} &= \frac{1}{d}(c\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} - b\boldsymbol{\Sigma}^{-1}\mathbf{1}), \end{aligned} \tag{5.16}$$

[2]The method of Lagrange multipliers being used here is described in any textbook on multivariate calculus.

where

$$\begin{aligned} a &= \boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}, \\ b &= \boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \mathbf{1} = \mathbf{1}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}, \\ c &= \mathbf{1}^T \boldsymbol{\Sigma}^{-1} \mathbf{1}. \end{aligned} \tag{5.17}$$

Further, let $d = ac - b^2$ be the determinant of these linear equations. Note that the positive definiteness of $\boldsymbol{\Sigma}$ implies that $a, c > 0$. The scalars b, c and d are also fixed, since they are functions of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, that is, independent of μ_P.

Proof. For an $n \times n$ matrix $\boldsymbol{A}$, and an n-dimensional vector $\boldsymbol{x}$,

$$\frac{\delta}{\delta \boldsymbol{x}} \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} = (\boldsymbol{A} + \boldsymbol{A}^T) \boldsymbol{x}.$$

The solution to eq.(5.13) is given by

$$\boldsymbol{\omega}_{\mu_P} = \lambda_1 \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + \lambda_2 \boldsymbol{\Sigma}^{-1} \mathbf{1}, \tag{5.18}$$

where $\lambda_i = \frac{1}{2}\delta_i$, and are yet to be determined scalar quantities. We need to use constraints (5.5) and (5.6) to find λ_1 and λ_2. Using eq.(5.18), the constraints imply the equations

$$\begin{aligned} \mu_P &= \boldsymbol{\mu}^T \boldsymbol{\omega}_{\mu_P} = \lambda_1 (\boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}) + \lambda_2 (\boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \mathbf{1}), \\ 1 &= \mathbf{1}^T \boldsymbol{\omega}_{\mu_P} = \lambda_1 (\mathbf{1}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}) + \lambda_2 (\mathbf{1}^T \boldsymbol{\Sigma}^{-1} \mathbf{1}). \end{aligned} \tag{5.19}$$

Using a, b and c, as defined in eq.(5.17), we can rewrite the two equations in (5.19) in matrix form,

$$\begin{bmatrix} \mu_P \\ 1 \end{bmatrix} = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}. \tag{5.20}$$

Now eq.(5.20) can be written as

$$\begin{aligned} \mu_P &= a\lambda_1 + b\lambda_2 \\ 1 &= b\lambda_1 + c\lambda_2. \end{aligned} \tag{5.21}$$

The solution of eq.(5.21) is

$$\lambda_1 = \frac{-b + c\mu_P}{d} \quad \text{and } \lambda_2 = \frac{a - b\mu_P}{d},$$

and it follows after some algebra the result in eq.(5.15), where $\boldsymbol{g}$ and $\boldsymbol{h}$ are given as in eq.(5.16). □

5.2.1 The Minimum-Variance portfolio

We have just shown that efficient portfolios with expected return μ_P has weights equal to eq.(5.15), thus the variance of the return on the portfolio $\boldsymbol{\omega}_{\mu_P}^T \boldsymbol{r}$ is given by

$$\begin{aligned} Var(r_p) &= (\boldsymbol{g} + \boldsymbol{h}\mu_P)^T \boldsymbol{\Sigma} (\boldsymbol{g} + \boldsymbol{h}\mu_P) \\ &= \boldsymbol{g}^T \boldsymbol{\Sigma} \boldsymbol{g} + 2\boldsymbol{g}^T \boldsymbol{\Sigma} \boldsymbol{h} \mu_P + \boldsymbol{h}^T \boldsymbol{\Sigma} \boldsymbol{h} \mu_P^2. \end{aligned} \tag{5.22}$$

To find *the minimum variance portfolio*, we minimize this quantity over μ_P by solving the equation

$$0 = \frac{d}{d\mu_p} Var(R_p) = 2\boldsymbol{g}^T \boldsymbol{\Sigma} \boldsymbol{h} + 2\boldsymbol{h}^T \boldsymbol{\Sigma} \boldsymbol{h} \mu_p. \tag{5.23}$$

The solution is the expected return of the minimum variance portfolio given by

$$\mu_{min} = -\frac{\boldsymbol{g}^T \boldsymbol{\Sigma} \boldsymbol{h}}{\boldsymbol{h}^T \boldsymbol{\Sigma} \boldsymbol{h}}. \tag{5.24}$$

We plug μ_{min} into $Var(r_p)$, and find that the smallest possible variance of a portfolio is r_{min} given by

$$Var(r_{min}) = \boldsymbol{g}^T \boldsymbol{\Sigma} \boldsymbol{g} - \frac{(\boldsymbol{g}^T \boldsymbol{\Sigma} \boldsymbol{h})^2}{\boldsymbol{h}^T \boldsymbol{\Sigma} \boldsymbol{h}}. \tag{5.25}$$

The minimization problem with additional non-negativity constraints cannot be solved by the method of Lagrange multipliers because of the inequality constraints, and since it cannot be solved analytically and the solution requires optimization tools, it must be represented as a quadratic programming problem. Furthermore, the theory is unable to account for the presence of higher moments beyond the mean and variance in both the portfolio returns distributions or investor preferences, such as skewness and kurtosis.

5.2.2 Effects of covariance

In statistics and financial theory, the covariance is used to describe the directional relationship between two risky assets. As for investing, we are highly interested in how the covariance is used to determine how different asset returns over a period of time changes to one another. As we know, a positive covariance means the assets return move together, which is bad because it increases the portfolio volatility. Conversely, negative covariance is good for diversification, which means they move in opposite directions, hence the portfolio volatility decreases. As will be discussed later in this thesis, by adjustments in the covariance matrix and inclusion of different asset classes to the data portfolio, then we will be able to determine if the portfolio is adequately diversified. Furthermore, the covariance between two random variables is defined as

$$Cov(r_{i,t}, r_{j,t}) = \frac{corr(r_{i,t}, r_{t,j})}{sd(r_i) \times sd(r_j)} \tag{5.26}$$

The one key benefit of diversification include minimizing risk of loss to your investment portfolio, which might be the core wisdom when investing in risky assets, that is, always keep a diversified portfolio in order to keep the level of risk as low as possible. Hence, by diversifying your investments, we can achieve smoother, more consistent investment returns over the medium to longer term. In order to give this principle of diversification a mathematical justification, see eq.(5.25) and eq.(5.26). Large, positive values of both these performance measures are an indication of superior portfolio performance, and that the portfolio is adequately diversified.

5.2.3 Out-of-sample performance measures

Following DeMiguel and Uppal [2009], we employ a number of performance measures, namely the Certainty Equivalent (CEQ) and Sharpe Ratio (SR).

Certainty Equivalent (CEQ)

The realized excess return on the portfolio is $r_{p,t+1} = w_{p,t} r_{t+1}$. We are interested in the average value of the $T - M$ realized returns, $\bar{r}_p$, and the standard deviation, $\hat{\sigma}_p$, of the portfolio's excess return. The certainty-equivalent (CEQ) return[2] is defined as "the amount of money that is equivalent to a given situation that involves uncertainty" and thus given by

$$CEQ = \bar{r}_p - \frac{\gamma}{2}\hat{\sigma}_p^2, \tag{5.27}$$

which can be interpreted as the risk-free rate of return that an investor is willing to accept instead of adopting the given risky portfolio rule p. Clearly the higher the CEQ, the better the rule. Similar to DeMiguel and Uppal [2009] and Tu and Zhou [2011], we assume that investors have quadratic utility preferences and a risk aversion (γ) value of 1. All the CEQ's have a common term of the average realized risk-free rate, which cancels out in their differences. We report the CEQ's by ignoring the risk-free rate term in our analysis.

Sharpe Ratio (SR)

Sharpe ratio is one of the most important tools to measure the performance of any investment. One way to evaluate if higher return came from taking more risk is by looking at the Sharpe Ratio. The estimate of the strategy's SR_p is calculated by the out-of-sample average portfolio's excess return divided by the return volatility $\hat{\sigma}_p$, and is therefore used to help investors understand the return of an investment compared to its risk. The Sharpe ratio is appropriate for well-diversified portfolios, just because it more accurately takes into account the volatility of the portfolio return. Throughout this thesis, we report the Sharpe Ratio by ignoring the risk-free rate term. Generally, the greater the value of the Sharpe Ratio, the more attractive the risk-adjusted return is compared to similar portfolios with lower returns. The formulation of the Sharpe Ratio is

$$SR = \frac{\bar{r}_p - r_f}{\hat{\sigma}_p}, \tag{5.28}$$

where r_f is the risk-free rate, $\hat{\sigma}_p$ is return volatility, and$\bar{r}_p$ is the average portfolio's excess return.

[2]The equations used to calculate these metrics (see eq.(5.27)) can be found in Tu and Zhou [2011] and Yew-Low et al. [2016].

Chapter 6

Local Gaussian Correlation

It is a well documented fact that there are asymmetries in the distribution of financial returns. These asymmetries appears when the market is going down, that is, there seem to be a stronger correlation between financial asset returns when the market falls rather than when it is stable or going up. This have the effect of destroying the benefit of diversification because of estimation error in the covariance matrix. In this chapter we present an approach based on the classical correlation concept which is capable of detecting and quantifying such asymmetries between financial returns such as bull and bear effects, but from a very different angle than that of the conditional correlation (see eq.(A.0.2)) and which result in a dependence measure that is quite easy to interpret.

We first return to the classical conditional correlation. Similar localization is crucial in this new approach. The central idea is simple; let's consider a point $x = (x, y)$ on the plane $\mathbb{R}^2$. We are then interested in approximating the general bivariate return density function f in the neighborhood of A around x, and we denote this bivariate Gaussian approximation with a return density ψ_x. We do this procedure on the entire region where f is defined, and take the correlation coefficient $\rho(x)$ of that Gaussian density as our measure of local dependence, and call it the local Gaussian correlation[1] (LGC). There are several precise interpretations of the correlation coefficient $\rho(x)$, but most importantly, it completely describes the dependence structure of a pair of random variables $v = (v_1, v_2)^T$ having ψ_x as its density, and since ψ_x is close to f_x in the neighborhood of x, it completely approximates the dependence structure of f around x. We explain this approach in more fine details in a return distribution context and demonstrate that it has a number of advantages compared to the conditional correlation. The main references used in this chapter; Tjøstheim and Hufthammer [2012], Støve and Tjøstheim [2014], Berentsen [2014] and Otneim [2017]. The texts also provide references for further reading.

6.1 Local Gaussian approximation and local correlation

Let the return at each point $x = (x, y)$ be a random variable, then the general bivariate return density function f_x is approximated by a Gaussian bivariate density function

$$\begin{aligned}\psi_x &= \psi(v, \theta(x)) \\ &= \frac{1}{2\pi\sigma_1(x)\sigma_2(x)\sqrt{1-\rho(x)^2}}\Big\{ -\frac{1}{2(1-\rho(x)^2)}\Big[\Big(\frac{v_1-\mu_1(x)}{\sigma_1(x)}\Big)^2 \\ &+ \Big(\frac{v_2-\mu_2(x)}{\sigma_2(x)}\Big)^2 - 2\rho(x)\Big(\frac{v_1-\mu_1(x)}{\sigma_1(x)}\Big)\Big(\frac{v_2-\mu_2(x)}{\sigma_2(x)}\Big)\Big]\Big\},\end{aligned} \tag{6.1}$$

where $v = (v_1, v_2)^T$ is the running variable in the Gaussian distribution, $\mu_i(x)$ are the local means, $\sigma_i(x)$ are the local standard deviations and $\rho(x)$ is the local correlation coefficient at point x. The population value of the local parameters depend on x such that ψ_x is close to f_x in that specific neighborhood of A, and is defined by the five-dimensional parameter vector

$$\theta(x) = \theta_b(x) = (\mu_1(x), \mu_2(x), \sigma_1(x), \sigma_2(x), \rho(x)). \tag{6.2}$$

But, the representation in eq.(6.1) is not well-defined unless it is the result of a minimization of a penalty function that measures the distance between ψ_x and f_x in the neighborhood of x. Such a penalty function can be described by a kernel function $K_b(v-x)$ (see eq.(6.4)), where the bandwidths are represented by $b = (b_1, b_2)$. Hence, the population parameter $\theta_b(x)$ are obtained by minimizing the local penalty function given by

$$q = \int K_b(v-x)[\psi_x(v, \theta(x)) - log\psi_x(v, \theta(x))f(v)]dv, \tag{6.3}$$

[1] Some of its theoretical properties are given in Tjøstheim and Hufthammer [2012].

where

$$K_b(v-x) = \frac{K(b_1^{-1}(v_1-x_1))K(b_2^{-1}(v_2-x_2))}{b_1 b_2}. \tag{6.4}$$

The penalty function in eq.(6.3) is measuring a sort of Kullback-Leibler distance between $\psi(\cdot,\theta(x))$ and $f(\cdot)$, that is in mathematical statistics, simplified as the distance of one "true" probability distribution to a second reference "target" probability distribution. As new information arrives, we move to another point x' of f, and another Gaussian density $\psi_{x'}$ is required to approximate $f_{x'}$ in the neighborhood A' of x'. Again, q can be used to obtain a new Gaussian approximation density $\psi_{x'}$ of $f_{x'}$ around x'. In theory, there are several properties and advantages of applying this model as a description of the asymmetric dependence within financial data;

1. The dependence measure is based on a family of Gaussian distributions, and describes the dependence relation for ψ_x and hence approximately for f_x in the neighborhood of A around x, since ψ_x approximates f around that point.

2. There are several ways of choosing the bandwidths $b = (b_1, b_2)$ of interest, for example, we can use the "plugin" or the "cross-validation" method[2], or we can simply choose a fixed size of 1 standard deviation from the point x. Although, in this thesis we have chosen the adaptive method of cross-validation, and hence, the bandwidths will vary as we move to another point x'. Hence, the size of b will be larger as if there are few observations in the neighborhood of x, and made smaller if there are many observations.

3. As $b \to \infty$, the measure of local correlation is steadily increasing to the global correlation coefficients. Hence, properties that are true for global dependence can be transferred locally to a neighborhood A.

4. In the case that f itself is Gaussian, $\rho(x) = \rho$ everywhere, where ρ is the ordinary correlation of f. This follows from the definition of ρ_b via the minimization of q by using local likelihood theory, and thus $\rho(x)$ does not suffer from the bias problem of the conditional correlation. However, there is some boundary bias in the estimation, but this is similar to boundary bias in ordinary kernel estimation.

5. As will be seen in Section 8.2, the local correlation $\rho(x)$ is capable of detecting and quantifying asymmetries in financial returns such as bull and bear effects.

6. It is possible to generalize the local approach to a set of d variables, having a joint density function f. The localized correlation $\rho(x)$ is then replaced by a local covariance matrix $\Sigma = \sigma_{ij}(x_i, x_j)$, for $i, j = 1, \dots d$. To avoid the curse of dimensionality in the estimation procedure, then for each pair of variables, the local covariance at point $(X_1 = x_1, \dots, X_d = x_d)$ has been restricted to depend only on (x_i, x_j).

7. Using local Gaussian likelihood theory (summarized in section 6.2) we can construct asymptotic confidence intervals for $\rho(x)$, this allow us to judge whether an observed asymmetry for financial returns measured by $\hat{\rho}(x)$ is statistically significant. However, this is beyond our field of research, and is left for future works.

To find an estimate of $\theta_b(x)$, we require the method for fitting a Gaussian density ψ_x to f in the neighborhood of A around x by minimizing the penalty function given in eq.(6.3). One such method is the local likelihood estimation process, which is briefly explained in the following section.

6.2 Local likelihood theory

Given the observations $X_t = (X_t, Y_t)$, $t = 1, \dots, T$, the estimates of $\hat{\theta}_b(x)$ from f are obtained by maximizing the local log-likelihood function (see Hjort and Jones [1996] and Tjøstheim and Hufthammer [2012]) defined as

$$l = log(L(X_1, \dots, X_T, \theta(x))) = \frac{1}{T}\sum_{t=1}^{T} K_b(X_t - x) log\psi_x(X_t, \theta(x)) - \int K_b(v-x)\psi_x(v,\theta(x))dv, \tag{6.5}$$

where $b = (b_1, b_2)$ and the Kernel function is as described in eq.(6.4). As in the preceding section, we introduce the Kernel function $K_b(X_t - x)$ to describe the neighborhood of A around x. Notice that, by using the law of large numbers, and letting $T \to \infty$, it is easily seen that the local log-likelihood function, $-log(L)$, converges towards the penalty function q defined in eq.(6.3). By using the notation for the derivative

$$w_j(v,\theta(x)) = \frac{\delta}{\delta\theta_j} log\psi_x(v,\theta(x)),$$

[2]For more formal methods for choosing the bandwidth, see Tjøstheim and Hufthammer [2012].

it is easily seen that it satisfies the set of equations

$$\frac{\delta l}{\delta \theta_j} = \frac{1}{T}\sum_{t=1}^{T} K_b(X_t - x) w_j(X_t, \theta(x)) - \int K_b(v-x) w_j(v, \theta(x)) \psi_x(v, \theta(x)) dv. \tag{6.6}$$

We must first observe that when $\theta(x)$ is chosen to minimize q, and again, by letting T $\to \infty$ and using the law of large numbers on the first therm on the r.h.s of eq.(6.6), the expression for $\frac{\delta log(L)}{\delta \theta_j}$ converges towards

$$\to \int K_b(v-x) w_j(v, \theta(x))[f(v) - \psi_x(v, \theta(x))] dv,$$

and $\hat{\theta}_b(x)$ converges in distribution to $\theta_b(x)$ satisfying $\frac{\delta q}{\delta \theta} = 0$, with q as defined in eq.(6.3). Hence, θ_b is defined by the previous equation defined above. For small bandwidths b_1, b_2 and under appropriate smoothness conditions, requiring $\frac{\delta log(L)}{\delta \theta_j} = 0$ for all j , the following condition is satisfied,

$$w_j(x, \theta(x))[f(x) - \psi_x(x, \theta(x))] + O(b^T b) = 0,$$

and the local likelihood estimates constraints $\psi(v, \theta(x))$ to be close to $f(x)$ when v is close to x. This is the sense in which the family ψ_x approximates f as the neighborhood defined by the bandwidth b_1, b_2 shrinks, which is also obtained by differentiating the penalty function q of eq.(6.3). In practice, we obtain the estimates $\hat{\theta}_{b,T}$ by requiring $\frac{\delta log(L)}{\delta \theta_j} = 0$, and solving the resulting five-dimensional set of equations numerically. Note that we do not only obtain a local correlation estimate of $\hat{\rho}_b(x)$, but also local estimates of the mean $\hat{\mu}_{1,b}(x), \hat{\mu}_{2,b}(x)$, and local variances $\hat{\sigma}^2_{1,b}, \hat{\sigma}^2_{2,b}(x)$, where the latter can be used to obtain local covariance estimates[3].

For further analysis, we can check the quality of the Gaussian approximation by comparing $\psi(x, \hat{\theta}_b(x))$ to $\tilde{f}(x)$, the kernel estimate of f. Note that as $b \to \infty$, eq.(6.5) reduces to the ordinary log-likelihood for a Gaussian distribution ψ plus a constant, and hence $\rho(x)$ reduces to the ordinary global Gaussian correlation.

6.2.1 Distribution of the Parameters

In Hjort and Jones [1996], the main point of the local likelihood analysis is to derive alternatives to the kernel estimate of f. Arguments in Hjort and Jones [1996] and Tjøstheim and Hufthammer [2012] demonstrate that $\hat{\theta}_b(x)$ is asymptotically normal such that

$$(Tb_1b_2)^{1/2}[\hat{\theta}_{T,b}(x) - \theta_b(x)] \to \mathcal{N}(0, J_b^{-1} M_b (J_b^{-1})^T),$$

where

$$\begin{aligned} J_b = &\int K_b(v-x) w(v, \theta_b(x)) \\ &\times w^T(v, \theta_b(x)) \psi(v, \theta_b(x)) dv \\ &- \int K_b(v-x) \nabla w(v, \theta_b(x)) \\ &\times [f(v) - \psi(v, \theta_b(x))] dv \end{aligned} \tag{6.7}$$

and

$$\begin{aligned} M_b = b_1 b_2 &\int K_b^2(v-x) w(v, \theta_b(x)) \\ &\times w^T(v, \theta_b(x)) f(v) dv \\ -b_1 b_2 &\int K_b^2(v-x) w(v, \theta_b(x)) f(v) dv \\ &\times \int K_b^2(v-x) w^T(v, \theta_b(x)) f(v) dv \end{aligned} \tag{6.8}$$

where $v = (v_1, v_2)^T$ and $x = (x_1, x_2)$. These expressions are somewhat deceptive, since $J_b^{-1} M_b J_b^{-1}$ is of order $(b_1 b_2)^{-2}$, so that $Var[\hat{\theta}_{T,b}(x)]$ is of order $(Tb_1^3 b_2^3)^{-1}$, a considerably slower convergence rate than in the traditional non-parametric rate of $(Tb_1 b_2)^{-1}$, when $b_1, b_2 \to 0$. This limiting process is needed when the asymptotic distribution of $\hat{\theta}_{T,b}$ is considered as $T \to 0$ and $b_1, b_2 \to 0$. The leading term in the covariance expression $J_b^{-1} M_b J_b^{-1}$ is quite problematic to evaluate, and in practice we have used the bootstrap, but then we have to assume i.i.d. observations.

[3]For more details about the local parameters and estimation using local likelihood, including limiting behavior as $b_i \to 0$, $i = 1, 2$, and estimation of standard errors, we refer to Tjøstheim and Hufthammer [2012].

Chapter 7

Portfolio Optimization using LGC

As emphasized in the previous section, the main idea is that the global dependence structure is better described by a portfolio of local dependence measures compared to a one-constant measure of dependence, and thereby, the local dependence will better describe the asymmetric dependence structure within specific regions. The main goal of this chapter is to extend the classical MV framework (Markowitz [1952]) by exactly using this non-parametric estimation procedure. The chief objective is to demonstrate the usefulness through empirical examples using financial returns and simulations. For monthly equity returns data, this new method is chosen to substantially outperform the equally weighted $1/N$ portfolio and the classical Markowitz portfolio, hopefully. We will show that the portfolio optimization is straightforward using this new approach, only relying on a tuning parameter, the bandwidth. The main references used in this chapter; Støve and Tjøstheim [2014], Yew-Low et al. [2016], Otneim [2017] and Berentsen [2014]. The texts also provide references for further reading.

7.1 Methodology

First, let's describe the portfolio allocation problem in practice. We typically use monthly return data, but of course, shorter or longer time horizons are possible. In fact, this mimics the typical scenario faced by a portfolio manager. In the case of historical data, we follow the same approach as DeMiguel and Uppal [2009], Tu and Zhou [2011] and Yew-Low et al. [2016] where rolling sampling windows of historical returns are used to estimate the expected return vector $\boldsymbol{\mu}_t$ and covariance matrix $\boldsymbol{\Sigma}_t$ required as inputs into the Markowitz model. More specifically, the process is given as follows;

1. At time t, a rolling sampling window of M trading months is selected.
2. During each month t, starting from $t = M + 1$, the return data for the M previous months are used to calculate the one month ahead expected return vector $\boldsymbol{\mu}_t$ (see eq.(5.2)) and covariance matrix $\boldsymbol{\Sigma}_t$. As new information arrives at month $t + 1$, we update these estimates. This process is repeated by incorporating the return for each month going forward and ignoring the earliest one, until the end of the sample.

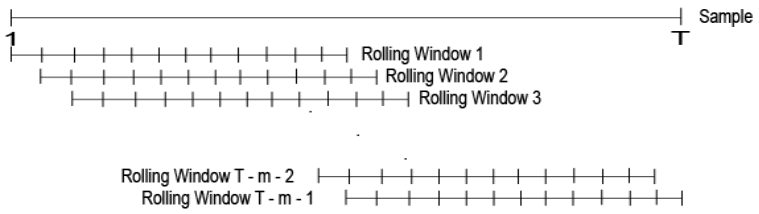


Figure 7.1: Rolling sampling window.

3. Based upon these estimates, we solve the various optimization problems and calculate the updated portfolio weights at every first trading day of each month, and do the rebalancing to construct a portfolio that achieves the desired investment objective.
4. The estimates of $\boldsymbol{\omega}$ are then used to calculate out-of-sample realized returns $\hat{r}_{t+1}$ (see eq.(5.3)) and portfolio performance over the next month. A total of $T - M$ out-of-sample returns are produced for each model.
5. These out-of-sample returns and portfolio weights are analyzed using a range of performance metrics and statistical measures that are reported for each model, respectively. For example, we can examine the cumulative returns resulting from a one dollar initial investment after a specified end date.

To ensure the robustness of our results, we use the same strategy as Tu and Zhou [2011] and Yew-Low et al. [2016], who find that longer sampling windows generally results in improved portfolio performances. Hence, we use rolling sampling windows of 120 and 240 trading months. As the optimization problem is presented in this manner, where the mean vector $\boldsymbol{\mu}_t$ and covariance matrix $\boldsymbol{\Sigma_t}$ are calculated globally, no explicit consideration is taken of asymmetries in the return distribution. In contrast, the method of local Gaussian correlation calculates a portfolio of local dependence measures to describe the asymmetric return distributions. To examine the impact of different levels of risk aversion and different maximum desired portfolio volatilites, we distinguish between mean-variance and minimum-variance investors and analyze the benefits (and downside) by the use of LGC.

Our method is identical to the above approach, except that in Step 2, the rolling sampling windows of historical returns are first used to estimate the LGC, then a local covariance matrix $\boldsymbol{\Sigma_t}$ are estimated based on the correlation estimates from this model. More specifically, step 2 is replaced by the following process;

- During each month t, starting from $t = M+1$, returns from the M previous months are used to calculate the one month ahead local covariance matrix $\boldsymbol{\Sigma}_t$ directly from the LGC estimates, $\hat{\rho}_b(x)$ and $\hat{\sigma_{i,b}}(x)$. As new information arrives at month $t+1$, we update these estimates. This process is repeated by incorporating the return for each month going forward and ignoring the earliest one, as previously mentioned.

Thus, our model is specified to account for asymmetries within the marginal distribution, and thus to hopefully reduce estimation error and provide a more reliable estimate of the efficient frontier.

7.1.1 Grid-map

In the optimization procedure when applying the LGC model, we end up with exactly one covariance matrix in each grid-point. A key question is then; what covariance matrix should be used for solving the optimization problem in finding the optimal weights for the different asset classes. In practice, a regular $N \times N$ grid is placed across the area of interest, then an investor is able to pick any grid-point of her own interest, respectively. For instance, a subjective meaning of where she thinks the market of a particular asset is going to be in the following trading month. Our strategy is to rely only on the covariance matrices estimated on the negative diagonal of the grid-map. This strategy is chosen on the basis of potential diversification benefits and performance gains when relying on the asymmetric dependence structure, especially under bear market conditions. An alternative strategy is that we don't rely on any particular assumption, but based upon the last recorded return observations, we estimate the one-step-ahead local covariance in that specific area using a "moving-grid" map. More specifically, the "moving-grid" point is defined as

$$x = moving - grid - point = (r^i_{t-1}, r^j_{t-1}). \tag{7.1}$$

Then, we approximate the general bivariate return density function f in the neighborhood of A around x, and take the correlation coefficient, $\rho(x)$, of that Gaussian density as our measure of the dependence structure within the specific area. In this way we are able to pick the "most likely" covariance matrix in the following trading month, and thus hopefully use this to produce even smoother portfolio target weights. In order to solve the problem of extreme correlations and estimation errors, we have restricted the moving-grid to only operate under normal market conditions, that is, around grid-point 0. I will explain this further in section 8.2.

7.1.2 Explaining implementation in R

In the analysis we have mainly used the **lg**[1] package in the procedure of estimating the local correlation. This package provides a very powerful collection of functions to describe the dependence structure within financial returns, and to analyze them from different points of view. The entire R code, which contains of approximately 8000 lines and 9 scripts, can be sent upon request.

The local Gaussian correlation in *R*

An $N \times N$ grid is placed across the area of interest, this is implemented in the next code snippet;

```
z <- 1/100
grid <- cbind(seq(-N, N, z), seq(-N, N, z))
```

whereas a moving-grid for bivariate random variables are implemented, after some restrictions on the grid (for example that the grid cannot pass the boundaries beyond N, N), as;

```
if (r1 < r2) {
  grid <- cbind(seq(r1, r2, r), seq(r1, r2, r))
} else {
  grid <- cbind(seq(r2, r1, r), seq(r2, r1, r))
}
```

[1]The packages used in this thesis can be found at the CRAN webpage.

The equation $\frac{\delta l}{\delta \theta_j} = 0$ (see eq.(6.5)) do not in general have closed-form solutions, therefore it must be maximized directly. In the R-package **lg**, the *lg_main* function lets the user supply a dataset and set a number of options, which is then used to maximize eq.(6.5) for different values of x. Hence we create an $lg - object$ that can be supplied to other functions in the **lg** package, such as the *dlg* function. The *dlg* function takes an $lg - object$ as produced by the *lg_main* function (where all the running parameters are specified), and a grid of points where the density estimate should be estimated. The *dlg* function does multivariate density estimation using the locally Gaussian density estimator[2], which is then used to obtain a set of local parameters (see eq.(6.3)). The estimation procedure for bivariate random variables are implemented in the next code snippet;

```
lg_object <- lg_main(x, bw_method="cv", est_method = "5par",
                     transform_to_marginal_normality = FALSE)
density_estimate <- dlg(lg_object, grid = grid)
lgc[2,1] <- density_estimate1$loc_cor[i,]
lgc[1,2] <- lgc[2,1]

lgcCov[2,1] <- lgc[2,1]*(sd(x[,1])*sd(x[,2]))
lgcCov[1,2] <- lgcCov[2,1]
```

This procedure is performed for each pair-wise asset return series. Since eq.(6.5) has to be optimized for every point on the grid-map, the computational time will increase proportionally with the size of the grid and as the number of grid-point within the increases. The method used for selecting the bandwidths is left to the user. In this way, the LGC can be estimated with different bandwidths, reflecting the dependence structure on different scales of locality. In this analysis, the bandwidths are chosen by the Cross-Validation method, which is very slow, but accurate. An alternative is to use the plugin method, which is fast, but crude compared to the cross-validation method. Note that if the bandwidths are very large, the local correlation estimates can be shown to converge to the global correlation in all grid-points. Further, we use the estimation method "5par", which is a full non-parametric locally Gaussian fit of a bivariate density as laid out and used by Tjøstheim and Hufthammer [2012], and others[3]. The estimated covariance matrix is then fed into the Markowitz machinery.

Portfolio optimization in R

As mentioned earlier, portfolio allocation is a quadratic programming (QP) problem, which contain a quadratic term $\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x}$ in the objective function. The R-package we have used providing quadratic solvers is **quadprog** (see Charpentier [2015]). The quadratic formulation is presented in the next code snippet;

```
QP_solver <- function(c, Q, cstr = list(), trace = FALSE) {--}
```

and since a wrapper function has been defined for the QP solver, implementation of the MV portfolio is straightforward, according to

```
MV_QP <- function(x, target, Sigma=lgcCov, ...,
                  cstr=c(fullInvest(x),
                         targetReturn(x, target),
                         longOnly(x), ...),
                  trace=FALSE) {--}
```

The function can be tested in a variety of ways. For example, we can optimize the weights that minimize the risk using the equally weighted portfolio return $mean(x)$ as the target return (see eq.(5.8));

```
w1[i,] <- MV_QP(x, mean(x), Sigma=lgcCov)
w2[i,] <- MV_QP(x, mean(x), Sigma=lgcCov, cstr = c(fullInvest(x), targetReturn(x, mean(x))))

w3[i,] <- MV_QP(x, Sigma=lgcCov, cstr = c(fullInvest(x), longOnly(x)))
w4[i,] <- MV_QP(x, Sigma=lgcCov, cstr = c(fullInvest(x)))
```

Note that the only difference between the Mean-Variance and Minimum-Variance portfolio is that we remove the target return $mean(x)$. Furthermore, if we produce portfolio weights of NAs, we are forced to re-estimate the covariance matrix at grid-point $x + 1$ on the diagonal of the grid-map, and then fit the new covariance matrix into the portfolio allocation model. Therefore, in order to fix the problem of estimation error, and to remain as risk-averse as possible, we follow a slightly different approach. For every optimization rule producing target weights of NAs, we loop one step towards the center of the grid-map until we produce valid portfolio weights. See the following code snippet;

```
if(any(is.na(w[i,]))) {
  for (j in 1:length(grid[,i])) {
    if (any(is.na(w[i,]))) {--}
    else { break } } }
```

[2]The locally Gaussian density estimator was introduced by Otneim and Tjøstheim [2017].
[3]Other references of importance cited in this chapter; Berentsen [2014] and Støve [2017]

7.2 Asset allocation models

The classical $1/N$ weighted portfolio strategy will be used as the benchmark model in our simulation study and throughout this thesis. It is widely used to evaluate different portfolio strategies, and in contrast to the sophisticated MV strategies, the naive $1/N$ diversification rule doesn't rely on any theory or data. The model distributes weights equally across the portfolio at the start of the sampling period, and is left unadjusted for the rest of the investment horizon. On the other hand, we can also use a re-balanced $1/N$ portfolio, that is, the profit (or loss) gained from one investment will be spread equally across the other assets within the portfolio. This benchmark model provides a guideline for periodically re-balancing your portfolio to help manage risk, and we may discover we are taking too much or too little risk. Not only can the naive $1/N$ investment strategy perform better than those sophisticated rules recommended from investment theory, but also as shown elsewhere and below, most of these rules do not perform well in real data sets. These findings raise a serious doubt on the usefulness of the investment theory, and to address this problem, we examine whether some or all of the optimization rules can be sufficiently improved so that they can outperform the naive $1/N$ rule. Our strategy for answering this question is to focus on the idea that optimal portfolio diversification is dependent upon the quality of the sample inputs into the portfolio allocation model. Of particular interest are the asymmetries within the distribution of stock returns. These asymmetries manifest in the form of asymmetric volatility clustering, skewness within the distribution or as asymmetric correlation.

Table 7.1: Mean-variance optimization rules considered.

# Model	Abbreviation
Benchmark models	
1. 1/N with re-balancing	EWR
Classic approach that ignores estimation error	
2. Mean-Variance without shortsale constraints	MVS
Moment restriction approaches	
3. Minimum Variance without shortsale constraints	MIN
Portfolio constraint approaches	
4. Mean-Variance with shortsale constraints	MVSC
5. Minimum Variance with shortsale constraints	MINC

This table reports the portfolio optimization rules implemented within our study.

The MVS strategy is the classic approach where historical mean returns and the covariance matrix are used to determine the weights for each out-of-sample period, where no consideration is given within the optimization rule to adjust for estimation error in any form. One frequently cited drawback is the use of the covariance matrix to estimate risk, where the problem resides in the fact that the sample covariance estimator is sensitive to outliers. As such, the theory is unable to account for the presence of higher moments beyond the mean and variance in both the portfolio returns distributions or investor preferences. Thus, the MVS optimization is often criticized for having little practical use as it maximizes estimation error, produces inappropriate portfolios on account of the error introduced by potential outliers that influence the estimation of the mean. Another popular benchmark model that will be used in our simulation study and empirical analysis, is the MIN portfolio strategy. This portfolio strategy can be easily implemented by removing the target return condition from the constraints of the portfolio. Thus, the objective of this portfolio strategy is to minimize portfolio risk measured as variance of portfolio returns. One advantage is that it does not require any return estimates, and it therefore benefits from the higher premium per unit of volatility of low-risk assets. Traditionally, as will be done in our analysis, many papers have used the MVS and MIN models as benchmarks to compare their proposed model improvements. DeMiguel and Uppal [2009] document that the best performing MV rules are often variants of the MIN. Ignoring expected returns successfully reduces the occurrence of extreme out-of-sample portfolio weights compared to the MVS (Jagannathan and Ma [2003]).

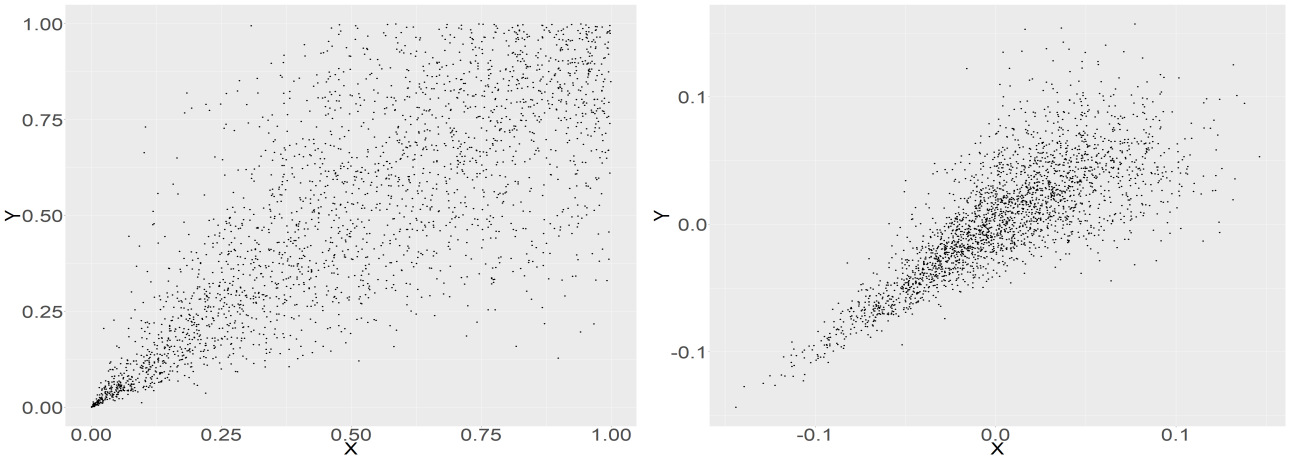
Models with portfolio constraints

The implementation of the MVC and MINC rules are the same as for their non-constrained counterparts except all weights are non-negative. Intuitively, all positions in the portfolio are long only. This approach accords with the fact that many funds in practice have short sales restrictions. Furthermore, empirical evidence shows that short-sales constrained portfolios usually exhibit better performance as they shrink the elements of the covariance matrix (see Jagannathan and Ma [2003]).

7.3 Simulation study

This section applies the proposed analysis to stocks portfolio and shows the benefits that may accrue from our proposed approach by using simulated series with the same characteristics as real data, that is, assuming asymmetric dependence structure within the market, in other words, an increased correlation as the market is going down. Hence, to demonstrate the asymmetric correlation, we create a two-dimensional clayton copula using the *claytonCopula* function with α=3. Adjustments in the α-coefficient determines the degree of dependence between the asset returns, respectively. We generate n=464 random variates for each asset using the *rCopula* function (See Figure 7.2 (a)). Now, let's generate two random variables (X, Y), which are the returns from two different assets and follow the dependence structure as defined by the *claytonCopula* and *rCopula* function, all having the properties of $r_t \sim \mathcal{N}(\mu = 0.00665, \sigma^2 = 0.045^2)$ (see Figure 7.2 (b)). A visual representation of the diagonal local Gaussian correlation values are given in Figure 7.3.

Figure 7.2: Asymmetric copula with lower tail dependence.



(a) 2-dimensional Clayton copula with α=3

(b) $\mathcal{N}(\mu = 0.665, \sigma^2 = 0.045^2)$

Furthermore, in financial analysis, portfolio performance measures are a key aspect of the investment decision process. These tools provide investors with the necessary information to assess how effectively their money has been invested (or may be invested). Remember that portfolio returns are only part of the story. Without evaluating risk-adjusted returns, an investor cannot possibly see the whole investment picture, which may inadvertently lead to clouded decisions. Hence, we follow the procedure used by Tu and Zhou [2011] and report the Sharpe Ratio (SR) and Certainty Equivalent (CEQ) return values to evaluate the out-of-sample performance across the portfolio optimization strategies investigated, e.g., the performance of a portfolio using covariances estimated under bear market conditions are defined as SR^{bear}_{LGC} and CEQ^{bear}_{LGC}, etc. The performance of a moving-grid estimated portfolio is defined as SR_{LGC} and CEQ_{LGC}, and global estimated portfolios as SR_M and CEQ_M, respectively.

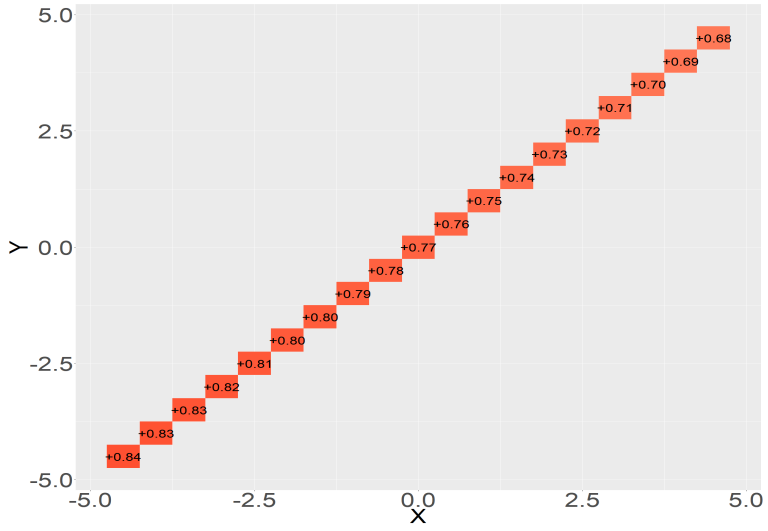


Figure 7.3: Diagonal local Gaussian correlation values for two simulated asset returns.

Before we delve into the performance results of our simulation study, I will clarify that the main goal of this simulation study is to look at the behavior of the different performance measures when adjustments in the covariance matrix (e.g. using the estimated covariance matrix under bear market conditions etc) are made with the goal of manipulating the risk-and-return characteristics in our favor. If we are able to better predict the one-step-ahead covariance matrix and more precisely capture the asymmetry within the asset returns, then we might be able to even further reduce the volatility within our estimated target portfolio weights, and thus perform better with our portfolio strategies. An even better outcome would be to reduce the volatility within our portfolio targets, and at the same time increase the average portfolio's excess return. The results are given in Table 7.2 and 7.3.

7.3.1 Out-of-sample performance of a rolling window

Table 7.2 reports the Sharpe Ratio and CEQ return values for the simulated data across the MVS- and MIN optimization strategies against the benchmark $1/N$ strategy. Panel A in Table 7.2 reports the Sharpe Ratio and CEQ values for the benchmark $1/N$ model. Overall, the $1/N$ portfolio performs quite well in the sense of Sharpe Ratio and CEQ compared to the more sophisticated MV models.

Table 7.2: Portfolio performance results across the optimization rules with fixed- and moving grid-points.

Panel A: *Benchmark model*			Sharpe Ratio (SR)			Certainty-equivalent (CEQ)		
$\text{EWR}_{M=120}$ (1/N)			0.12733			0.81056		
$\text{EWR}_{M=240}$ (1/N)			0.12208			0.53936		
Portfolio Strategy	SR_M	SR_{LGC}^{bear}	$SR_{LGC}^{norm.}$	SR_{LGC}^{bull}	CEQ_M	CEQ_{LGC}^{bear}	$CEQ_{LGC}^{norm.}$	CEQ_{LGC}^{bull}
Panel B1: *alternative portfolio rules, Window size M=120, global s.d.*								
Classic approach that ignores estimation error								
MVS	0.12733	0.12733	0.12733	0.12733	$0.81056^{\#}$	$0.81056^{\#}$	$0.81056^{\#}$	$0.81056^{\#}$
Moment restrictions								
MIN	0.13165	$0.13248^{\uparrow}$	0.13165	0.13090	$0.86610^{\#}$	$0.87719^{\uparrow\#}$	$0.86616^{\uparrow\#}$	$0.85627^{\#}$
Panel C1: *alternative portfolio rules, Window size M=240, global s.d.*								
Classic approach that ignores estimation error								
MVS	0.12208	0.12208	0.12208	0.12208	0.53936	0.53936	0.53936	0.53936
Moment restrictions								
MIN	0.12395	$0.12429^{\uparrow}$	$0.12396^{\uparrow}$	0.12361	0.55894	$0.56380^{\uparrow}$	$0.55896^{\uparrow}$	0.55479
Moving-grid	SR_M		$SR_{LGC}^{global-s.d.}$	$SR_{LGC}^{local-s.d.}$	CEQ_M	$CEQ_{LGC}^{global-s.d.}$		$CEQ_{LGC}^{local-s.d.}$
Panel B2: *alternative portfolio rules, Window size M=120*								
Classic approach that ignores estimation error								
MVS	–		0.12733	0.12733	–	$0.81056^{\#}$		$0.81056^{\#}$
Moment restrictions								
MIN	–		0.13165	0.13156	–	$0.86612^{\uparrow\#}$		$0.86493^{\#}$
Panel C2: *alternative portfolio rules, Window size M=240*								
Classic approach that ignores estimation error								
MVS	–		0.12208	0.12208	–	0.53936		0.53936
Moment restrictions								
MIN	–		$0.12396^{\uparrow}$	0.12394	–	0.55894		0.55875

Table 7.2 reports the Sharpe Ratio and CEQ metric for various portfolio rules (see Table 7.1) applied to simulated data when global (e.g., SR_M, CEQ_M) or local estimates (e.g., SR_{LGC}, CEQ_{LGC}) are applied. Panel A contains the benchmark models. Panel B and C reports the out-of-sample portfolio performance using sampling windows length of both 120 and 240 months, respectively. $^{\#}$ indicates higher CEQ when shorter historical sampling windows are applied.

As can be seen from the results in Table 7.2, Panel B1 indicates better performance for 1 (2) strategies when relying on local covariances estimated under bear- normal, and bull market conditions, rather than global estimates. The CEQ return report that the best performing strategies are local estimates under bear market conditions. Overall, the MIN rule seems to outperform the MVS rule across all strategies. In Panel C1 ($M = 240$), we find that longer sampling windows result in increased portfolio performance as oppose to $M = 120$. However, Panel C1 report increased performance for 2 (2) strategies when we rely on a longer sampling window. As for the moving-grid approach in Panel B2 and C2, note that estimates using global

standard deviations to estimate covariances produce higher Sharpe Ratio across all strategies, more specifically, higher Sharpe Ratio are produce for 1 (0) strategies when global (local) s.d. are used to estimate covariances, whereas 1 (0) strategies report higher CEQ return values. Summarizing Table 7.2, by using simulated data, we show that the extensions to the various mean-variance models that have been designed to deal with estimation error reduce only moderately the estimation error in a few models.

7.3.2 Coefficients of the optimization rules

The analysis in Section 7.3.1 indicates that the best performing strategies are variants of the MIN model in the sense of both the Sharpe Ratio and CEQ values. Further, in Table 7.3 we report the average- and standard deviation of the portfolio's excess return, and to assess whether the use of local estimates are able to simultaneously improve portfolio performance and increase (decrease) the coefficients of the strategies, we identify cases where the portfolio performance is superior to that based on the global estimates by "#", in the sense of CEQ return. In Table 7.3, Panel B1 ($M = 120$) and C1 ($M = 240$) reports the local- and global estimated average accumulated portfolio's excess return and volatility. Panel B1 report increased average return

Table 7.3: Coefficients applied to the different optimization rules investigated.

Portfolio Strategy			$\mu_{1/N}$			$\sigma_{1/N}$		
Panel A: *Benchmark model*								
$\text{EWR}_{M=120}$			1.6034			0.12593		
$\text{EWR}_{M=240}$			0.70713			0.057925		
Portfolio Strategy	μ_M	μ^{-4}_{LGC}	μ^{0}_{LGC}	μ^{4}_{LGC}	σ_M	σ^{-4}_{LGC}	σ^{0}_{LGC}	σ^{4}_{LGC}
Panel B1: *alternative portfolio rules, Window size M=120, global s.d.*								
Classic approach that ignores estimation error								
MVS	1.6034	1.6034	1.6034	1.6034	0.12593	0.12593	0.12593	0.12593
Moment restrictions								
MIN	1.6925	$1.7208^{\uparrow\#}$	$1.6926^{\uparrow\#}$	1.6726	0.12856	0.12989	0.12857	$0.12778^{\downarrow}$
Panel C1: *alternative portfolio rules, Window size M=240, global s.d.*								
Classic approach that ignores estimation error								
MVS	0.70713	0.70713	0.70713	0.70713	0.057925	0.057925	0.057925	0.057925
Moment restrictions								
MIN	0.73450	$0.74198^{\uparrow\#}$	$0.73453^{\uparrow\#}$	0.72843	0.059256	0.059697	0.059258	$0.058930^{\downarrow}$
Moving-grid	μ_M	$\mu^{global-s.d.}_{LGC}$		$\mu^{local-s.d.}_{LGC}$	σ_M	$\sigma^{global-s.d.}_{LGC}$		$\sigma^{local-s.d.}_{LGC}$
Panel B2: *alternative portfolio rules, Window size M=120*								
Classic approach that ignores estimation error								
MVS	–	1.6034		1.6034	–	0.12593		0.12593
Moment restrictions								
MIN	–	$1.6925^{\#}$		1.6897	–	0.12856		$0.12843^{\downarrow}$
Panel C2: *alternative portfolio rules, Window size M=240*								
Classic approach that ignores estimation error								
MVS	–	0.70713		0.70713	–	0.057925		0.057925
Moment restrictions								
MIN	–	0.73450		0.73421	–	$0.059255^{\downarrow}$		$0.059239^{\downarrow}$

Table 7.3 report the volatility- and the average portfolio's excess return when global (μ_M, σ_M) or local (μ_{LGC}, σ_{LGC}) estimates are applied. The mean of the weights are presented as percentage points. Panel A contains the benchmark models. Panel B and C report the portfolio performance using sampling windows of 120 and 240 trading months, respectively. $^{\#}$ indicates higher CEQ value when local estimates are applied as opposed global estimates.

and decreased standard deviations for 2 (1) strategies, whereas Panel C1 report increased return and decreased return volatility for 2 (1) strategies. Accordingly, this leads to increased performance in the sense of Sharpe Ratio and CEQ as can be seen in the previous section. For the longer sampling window of 240 trading months, Panel C2 report increased average return and decreased volatility for 0 (2) moving-grid strategies, whereas only 0 (1) strategies indicate improvements in Panel B2. Note that the strategies which exhibit lower standard deviations, are not improved in terms of their CEQ returns. The dependence structure within the simulated variates can be easily manipilated by adjustments in the α parameter in order to demonstrate in even

more fine details the possible benefits by applying the LGC estimates. However, in the following section we do the simulation 50 times, and calculate the performance by averaging the results from each simulation.

7.3.3 Moving-grid with N=50 simulations

In this section we do the same as in the previous section, but now the simulation is executed 50 times. We then calculate the portfolio performance by the average portfolio's excess returns $\frac{1}{N}\sum_{n=1}^{N}\bar{r}_n$ and return volatility from each simulation. As can be seen, the results in Panel B and C of Table 7.4 indicates that the only local portfolio

Table 7.4: Portfolio performance results across the optimization rules with fixed- and moving grid-points.

Panel A: *Benchmark model*			Sharpe Ratio (SR)			Certainty-equivalent (CEQ)		
$\text{EWR}_{M=120}$ (1/N)			0.11666			0.50768		
$\text{EWR}_{M=240}$ (1/N)			0.12988			0.81402		
Portfolio Strategy	SR_M	SR_{LGC}^{bear}	$SR_{LGC}^{norm.}$	SR_{LGC}^{bull}	CEQ_M	CEQ_{LGC}^{bear}	$CEQ_{LGC}^{norm.}$	CEQ_{LGC}^{bull}
Panel B1: *alternative portfolio rules, Window size M=120, global s.d.*								
Classic approach that ignores estimation error								
MVS	0.11666	0.11666	0.11666	0.11666	0.50768	0.50768	0.50768	0.50768
Moment restrictions								
MIN	0.11638	0.11590	0.11620	0.11638	0.51216	0.48785	0.49581	0.50037
Panel C1: *alternative portfolio rules, Window size M=240, global s.d.*								
Classic approach that ignores estimation error								
MVS	0.12988	0.12988	0.12988	0.12988	0.81402	0.81402	0.81402	0.81402
Moment restrictions								
MIN	0.12735	0.12559	0.12670	$0.12988^{\uparrow}$	0.78198	0.76412	0.77637	$0.81402^{\uparrow}$
Moving-grid	SR_M		$SR_{LGC}^{global-s.d.}$		CEQ_M		$CEQ_{LGC}^{global-s.d.}$	
Panel B2: *alternative portfolio rules, Window size M=120*								
Classic approach that ignores estimation error								
MVS	–		0.11666		–		0.50768	
Moment restrictions								
MIN	–		0.11638		–		0.51212	
Panel C2: *alternative portfolio rules, Window size M=240*								
Classic approach that ignores estimation error								
MVS	–		0.12988		–		0.81402	
Moment restrictions								
MIN	–		0.12734		–		0.78196	

Table 7.2 shows the Sharpe Ratio and CEQ metric for various portfolio rules (see Table 7.1) applied to simulated data when global (e.g., SR_M, CEQ_M) or local estimates (e.g., SR_{LGC}, CEQ_{LGC}) are applied. Panel A contains the benchmark models. Panel B and Panel C show the performance of the portfolio rules for sample windows length of 120 and 240 months, respectively. $^{\#}$ indicates higher CEQ when longer historical sampling windows are applied.

outperforming the global estimates is the MIN strategy under bull market conditions. In Panel C, when we use a longer sampling window of M=240 trading months, our results show the same as Tu and Zhou [2011], who find that longer sampling windows result in improved portfolio strategy performance, especially in the sense of CEQ. In addition, the benchmark $1/N$ portfolio performs equally or better than most local portfolio estimates, which is supported by both sampling window sizes of 120 and 240 trading months, respectively. For more information on the portfolio performance, see Table 7.5.

Discussion

An implication of the asymmetric dependence structure within financial returns is that the classical MV optimized portfolios are not efficient with respect to their effective risk profile, thus the benefit of diversification will erode if the correlations are asymmetric. This effect is barely illustrated in Table 7.2 and 7.4, as we take advantage of the asymmetric dependence when relying only on covariances from a falling market. In other words, when the market is going down and the correlation increases, this effect is true and a consequence of

the non-linearity of the dependence structure between the asset returns. As we discussed earlier, the main goal of this simulation is only to demonstrate how adjustments in the covariance matrix can help manipulate the risk-and-return characteristics to improve portfolio performance. As for the empirical analysis, we will further investigate potential benefits of diversification when adding bonds and commodity futures into our portfolio of financial returns. The idea is simple; low correlated assets are good for diversification, while highly correlated assets should be avoided. In addition, benefits from international diversification of asset allocation could be considerably affected by the asymmetric dependence structure, and this will be highlighted as we delve into the empirical analysis in Chapter 8. Our goal in the analysis is to manage these asymmetric dependence in larger portfolios and alternate asset classes, and improve portfolio performance. The asymmetric dependence

Table 7.5: Coefficients applied to the different optimization rules investigated.

Portfolio Strategy			$\mu_{1/N}$			$\sigma_{1/N}$		
Panel A: *Benchmark model*								
$EWR_{M=120}$			2.047			0.17546		
$EWR_{M=240}$			1.3717			0.10561		
Portfolio Strategy	μ_M	μ_{LGC}^{bear}	$\mu_{LGC}^{norm.}$	μ_{LGC}^{bear}	σ_M	σ_{LGC}^{bear}	$\sigma_{LGC}^{norm.}$	σ_{LGC}^{bull}
Panel B1: *alternative portfolio rules, Window size M=120, global s.d.*								
Classic approach that ignores estimation error								
MVS	2.047	2.047	2.047	2.047	0.17546	0.17546	0.17546	0.17546
Moment restrictions								
MIN	2.0233	2.0458	2.0463	2.0465	0.17384	0.17652	0.17610	0.17585
Panel C1: *alternative portfolio rules, Window size M=240, global s.d.*								
Classic approach that ignores estimation error								
MVS	1.3717	1.3717	1.3717	1.3717	0.10561	0.10561	0.10561	0.10561
Moment restrictions								
MIN	1.3158	1.2991	1.3148	1.3717	0.10332	0.10343	0.10377	$0.10561^{\#}$
Moving-grid	μ_M		$\mu_{LGC}^{global-s.d.}$		σ_M		$\sigma_{LGC}^{global-s.d.}$	
Panel B2: *alternative portfolio rules, Window size M=120*								
Classic approach that ignores estimation error								
MVS	–		2.047		–		0.17546	
Moment restrictions								
MIN	–		2.0233		–		0.17385	
Panel C2: *alternative portfolio rules, Window size M=240*								
Classic approach that ignores estimation error								
MVS	–		1.3717		–		0.10561	
Moment restrictions								
MIN	–		1.3157		–		0.10332	

Table 7.3 shows the standard deviation- and the average portfolio's excess return when global (μ_M) or local (σ_{LGC}) estimates are used for the data set investigated. The mean of the weights are presented as percentage points. Panel A contains the benchmark models. Panel B and C show the performance of the portfolio rules for sample windows length of 120 and 240 months, respectively. $^{\#}$ indicates a higher CEQ value when local estimates are applied as opposed global estimates.

structure in finance examines the risks and benefits of asset correlation, and provides effective strategies for more profitable portfolio management. Beginning with a thorough explanation of the extent and nature of asymmetric dependence in the financial markets, this thesis delves into the LGC model, and how it may implement to boost portfolio performance.

Chapter 8

Empirical Results

The empirical analysis employ a sample data in total of 464 trading months, and consists of monthly closing prices of a number of indices and commodity futures provided by Thomson Reuters Datastream, they are all denominated in US dollar. The data sets analyzed in this thesis are listed in Table 8.1, respectively.

1. The FTSE All-Share Index, originally known as the FTSE Actuaries All Share Index, is a capitalisation-weighted index, comprising around 600 of more than 2,000 companies traded on the London Stock Exchange. As of 29 December 2017, the constituents of this index totalled 641 companies (source FTSE Russell All-Share Index fact sheet).

2. The S&P 500 Comp. Index is an American stock market index based on the market capitalization of 500 large companies having common stock listed on the NYSE, NASDAQ, or the Cboe BZX Exchange.

3. The UK-and-US Benchmark 10 Year DS Govt. bond Indices represents yields and interest rates. In general, government bonds are issued by national governments and denominated in local currency.

4. The Thomson Reuters Equal Weight Commodity Index is recognized as a major barometer of commodity prices. The index comprises 17 commodity futures that are continuously re-balanced.

5. The S&P GSCI Gold serves as a benchmark for investment in the commodity markets and as a measure of commodity performance over time.

Table 8.1: Data sets considered.

Name	Source	N	Time period
(1) FTSE All Share	Datastream	668	01.01/1963-01.08/2018
(2) S&P 500 Comp.	Datastream	668	01.01/1963-01.08/2018
(3) UK Benchmark 10 Year DS GOVT. Ind.	Datastream	668	01.01/1963-01.08/2018
(4) US Benchmark 10 Year DS GOVT. Ind.	Datastream	668	01.01/1963-01.08/2018
(5) TR Equal Weight CCI	Datastream	668	01.01/1963-01.08/2018
(6) S&P GSCI Gold	Datastream	487	01.02/1978-01.08/2018

Table 8.1 lists the data and its sources, where N denotes the total number of monthly observations within the portfolio data set. The last column represents the time line available for each data index.

We restrict the timeline for all assets to the period from 01.01.1980 to 01.08.2018, due to data availability constraints. The presentation of our empirical analysis is structured as follows; Chapter 8.1 provides descriptive statistics related to the dataset and briefly discusses the t-GARCH model, chapter 8.2 give us an insight into the empirical analysis of the asymmetric dependence structures within our data set and provides empirical results, and provide us with the empirical results and performance using a rolling sampling window of 120 and 240 trading months, respectively.

8.1 Descriptive statistics

The analysis aim in the first section to identify statistical properties of our data, and therefore we need to transform the data into so-called returns. The returns are defined as the change in the natural logarithm of each market's index price (see eq.(2.4)), where P_t is the index closing price from Datastream. Market prices normalized to 100 (see Figure A.2). The monthly returns for the entire period however, are shown in Figure 8.3,

and the descriptive statistics are shown in Table 8.2, whereas Panel B, C, D and E show the pairwise correlation matrices over the entire sampling period. Again, as mentioned in Chapter 6, these results should be looked at as a sort of benchmark statistics, because stationarity cannot be expected to hold over such a long period.

Table 8.2: Summary index statistics for monthly data from January 1980 to August 2018.

	FTSE-All-Share	S&P-500-Comp.	BMUK10Y	BMUS10Y	TR-Equal-Weight-CCI	S&P-GSCI-Gold-Spot
Panel A: *Statistics*						
Observations	464	464	464	464	464	464
Arithmetic Mean	0.6280	0.7042	0.7690	0.5830	0.0794	0.1770
Standard deviation	4.5880	4.4061	2.3756	2.4165	3.5108	5.2114
Variance	21.0498	19.4135	5.6434	5.8393	12.3260	27.1592
Sharpe Ratio	0.1372	0.1587	0.3235	0.2396	0.0228	0.0345
Skewness	-1.3047	-0.9708	-0.1287	0.4541	-0.5941	0.0265
Kurtosis	6.3282	3.6939	1.3442	1.9810	3.8039	3.0623
Maximum	12.5225	14.6116	8.8507	12.6603	13.3842	26.3363
3 Quartile	3.5590	3.2647	2.1508	1.8531	1.9975	2.8990
Median	1.1755	1.2424	0.8425	0.4970	0.1508	-0.1609
1 Quartile	-1.4738	-1.6942	-0.5854	-0.9222	-1.7936	-2.6684
Minimum	-32.7109	-24.6768	-7.8244	-7.5998	-20.0500	-21.8865
Panel B: *Global correlation matrix*						
FTSE All Share	1.00					
S&P 500 Comp.	0.76029	1.00				
BMUK10Y	0.18438	0.01735	1.00			
BMUS10Y	-0.06722	-0.02941	0.48916	1.00		
TR Equal Weight CCI	0.24614	0.28836	-0.09398	-0.18451	1.00	
S&P GSCI Gold Spot	0.03755	0.03057	0.07953	0.07651	0.48343	1.00
Panel C: *Local correlation matrix estimated under bear market conditions*						
FTSE All Share	1.00					
S&P 500 Comp.	0.85634	1.00				
BMUK10Y	0.09001	-0.04800	1.00			
BMUS10Y	-0.13516	-0.07979	0.47074	1.00		
TR Equal Weight CCI	0.30035	0.41501	-0.07170	-0.17192	1.00	
S&P GSCI Gold Spot	0.03550	0.09964	0.09138	0.07887	0.58126	1.00
Panel D: *Local correlation matrix estimated under normal market conditions*						
FTSE All Share	1.00					
S&P 500 Comp.	0.75868	1.00				
BMUK10Y	0.186657	0.01817	1.00			
BMUS10Y	-0.065931	-0.02839	0.48930	1.00		
TR Equal Weight CCI	0.24577	0.28675	-0.09413	-0.18468	1.00	
S&P GSCI Gold Spot	0.03688	0.02903	0.07929	0.07710	0.48242	1.00
Panel D: *Local correlation matrix estimated under bull market conditions*						
FTSE All Share	1.00					
S&P 500 Comp.	0.70471	1.00				
BMUK10Y	0.25488	0.08565	1.00			
BMUS10Y	-0.01718	0.02071	0.50988	1.00		
TR Equal Weight CCI	0.20561	0.22987	-0.11417	-0.19620	1.00	
S&P GSCI Gold Spot	0.06930	0.04141	0.08752	0.07346	0.43842	1.00

Table 8.2 report descriptive statistics for the alternative asset classes used in this study. Panel A reports the summary statistics: annualized mean returns, standard deviations, Sharpe ratios as well as skewness and kurtosis figures which represents the third and fourth moment of the return distribution. Panel B reports the global correlation matrix, whereas Panel C, D and E reports the local correlation matrix estimated on the diagonal at grid-points -4, 0 and 4, respectively.

From Panel A in Table 8.2, notice that the arithmetic means are lower for the commodity indices than for stocks and bonds, and also exhibits a very high standard deviation. As a result, the Sharpe ratio is considerably

lower for the commodity indices. This suggest that commodity indices are not very attractive as a stand-alone investment for longer investment horizons. As for the stock indices, they follow a trend with clusters of high volatility, hence a higher standard deviation. The bond indices follow a rather strictly smooth curve with low volatility, and relatively high arithmetic mean, hence a higher Sharpe ratio. Thus, the bond indices may give best investment opportunities. Further, it is often valuable when evaluating an investment to know whether the instrument we examine follow a normal distribution. From our summary statistics in Table 8.2, we notice that the skewness deviates from 0, which represents the excess value used to describe the distribution. In general, if the skewness is outside the range of $[-1, 1]$, the distribution is highly skewed. If the skewness lies in between $[-1, -0.5]$ or $[0.5, 1]$, the distribution is moderately skewed, and if skewness is inside the $[-0.5, 0.5]$ interval, the distribution is approximately symmetric. Hence, skewness measures the degree of asymmetry in the return distribution, whereas positive skewness indicates that more of the returns are positive, and negative skewness indicates that more of the returns are negative. Furthermore, as can be seen from the histograms in Figure 8.1, the data seems to be either left or right skewed. In other words, the distribution is asymmetrical, that is, the distribution's peak is off center toward the limit, and a tail stretches away from it. Hence, all monthly returns are non-normally distributed, especially in the form of leptokurtosis. Thus the normal distribution is not

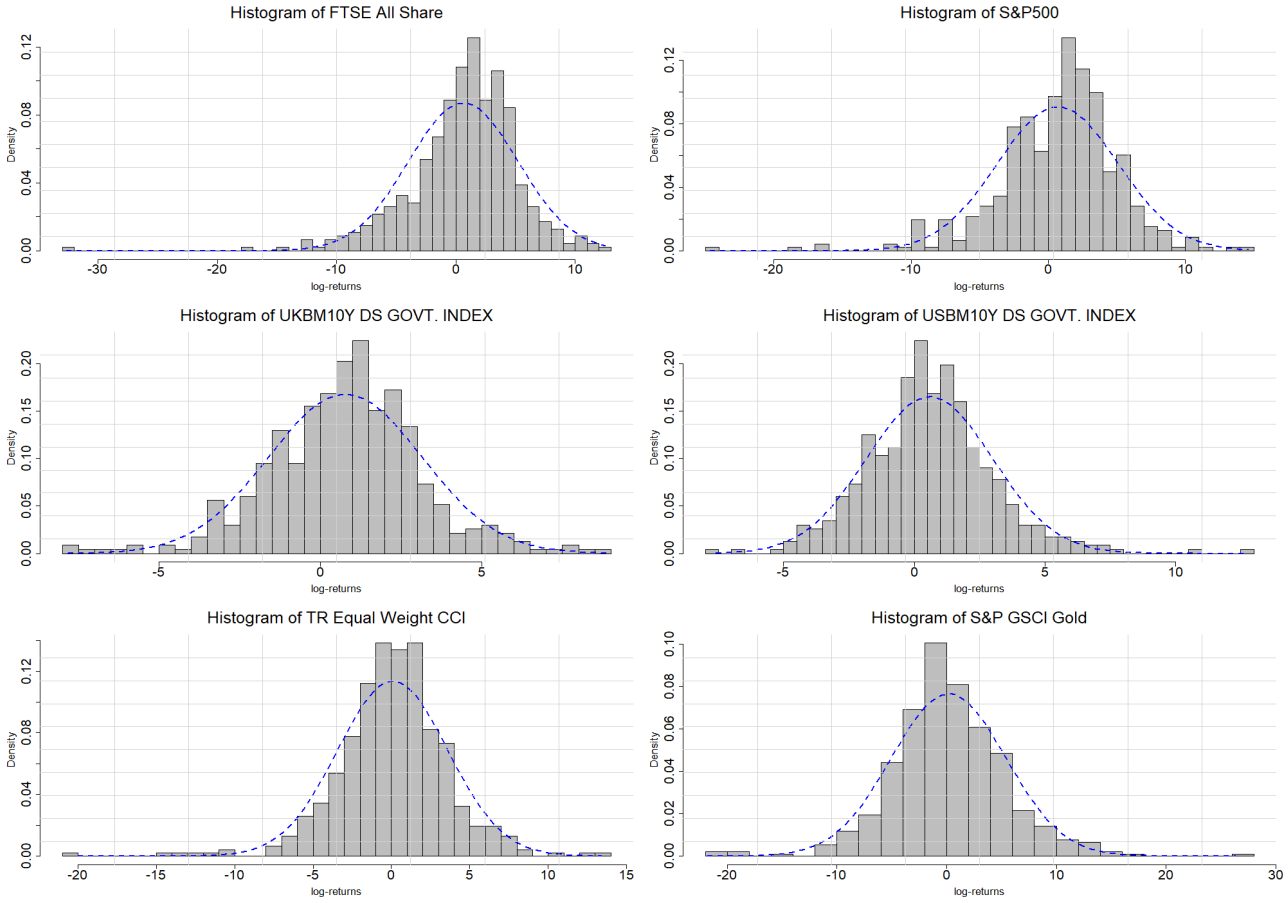


Figure 8.1: Histogram of monthly compound log-returns.

adequately gonna capture the probability that we will end up in the area where the losses occurs with low frequency and high severity, and therefore underestimate the likelihood and magnitude of these extreme tail losses. As we can see, the extent and direction of the skewness differs across sectors and equity markets. An investor should in most cases prefer a positively skewed asset to a similar asset that has a negative skewness. Furthermore, the kurtosis is in general high, that is, between 1.3442 and 6.3282. The kurtosis measures the concentration of the return in any given part of the distribution. We use the kurtosis definition where a normal distribution has kurtosis equal to 3. The high kurtosis values indicate a departure from the Gaussian distribution, and in these cases the global correlation may not be a good measure of the dependence structure. These observations indicates that all of our return series are more heavy-tailed. In general however, a rational investor should prefer an asset with a low to negative excess kurtosis, as this will indicate more predictable return. The major exception is generally a combination of high positive skewness and high excess kurtosis.

We further tested our data for normality by examining the Quantile-Quantile (QQ) plot, as shown in Figure 8.2. The QQ-plot relates the empirical and theoretical quantiles, and the points should fall approximately along

the reference line inside the confidence interval. The larger the departure from the reference line, the greater the evidence that the data come from a population with a different distribution. We can tell from Figure 8.2 that the observations do not lie inside the reference lines, especially in the tails. This confirms that the tails of the empirical distributions are not well described by the normal distribution, and again, are rather heavy-tailed.

Covariance matrix and diversification benefits

To obtain some insights in terms of potential diversification benefits, we report the pairwise global and local correlation matrices for the dataset investigated, given in Panel B, C, D and E in Table 8.2. As can be seen from these panels, the pairwise global correlation between each asset class is 0.76029, 0.48916 and 0.48343, respectively. We estimate the pairwise local correlation within each asset class at grid-point -4, 0 and 4, more specifically, under bear-, normal- and bull markets conditions, respectively. Notice how the local correlation

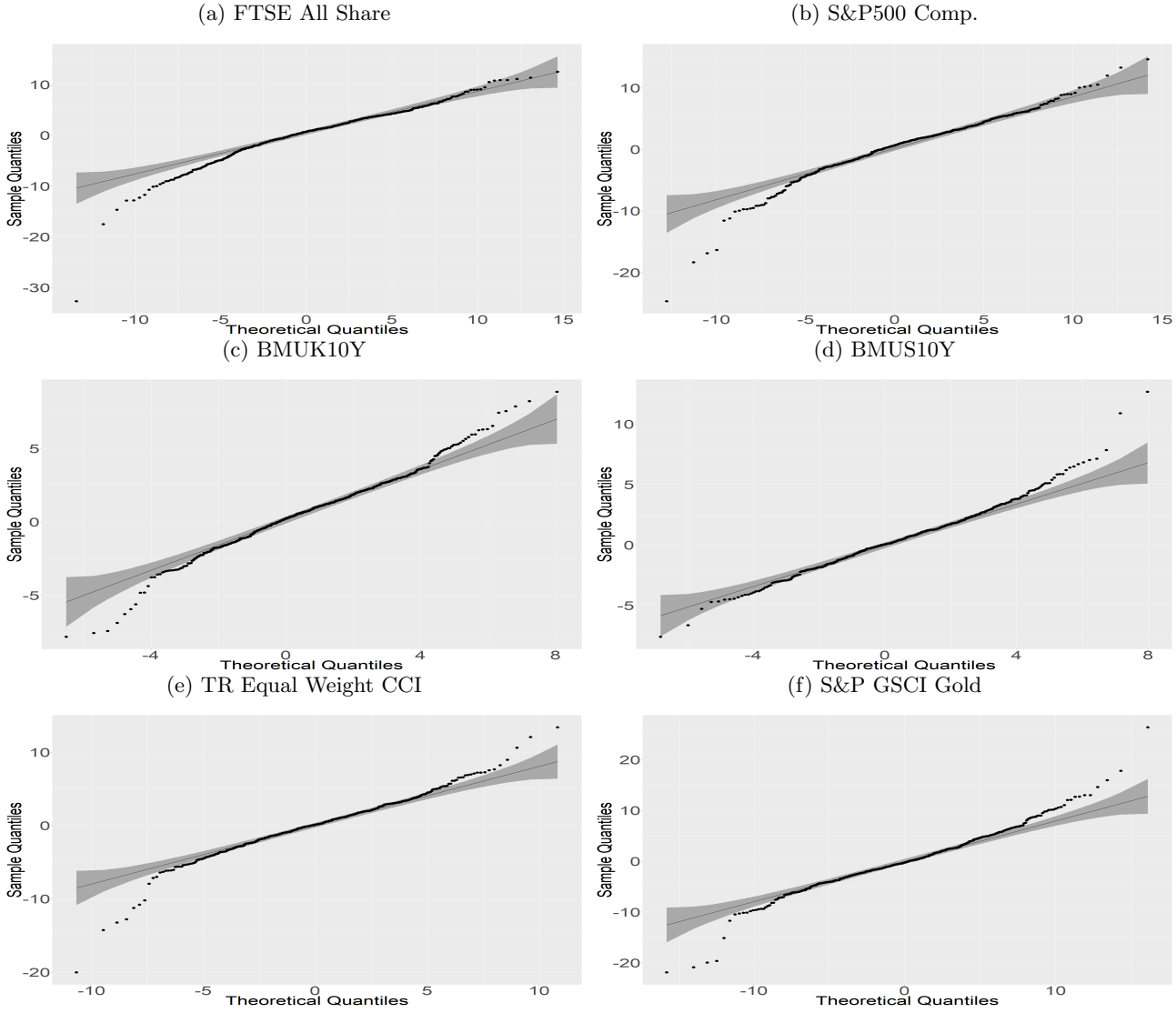


Figure 8.2: QQ-plots of monthly log-return indices.

adequately detect the asymmetric dependence structure within the market. For example, the pairwise local correlations at grid-point -4 (bear market); 0.85634, 0.47074, 0.58126, at grid-point 0 (normal market); 0.75868, 0.48930, 0.48242, and at grid-point 4 (bull market); 0.70471, 0.50988 and 0.43842. Hence, the LGC model is able to capture potential stronger correlations in a falling market, as opposed to normal- and bull markets. We will delve even further into this later in this thesis. As can be seen from the results above, the pairwise correlations between the asset classes are relatively high, this is an indication that we may not want to build a portfolio that only consists of assets within the same classes, but rather want to build a portfolio that is more widely spread among the different asset classes. As mentioned earlier, even if commodities do not appear to be an attractive asset class as a stand-alone investment, they might add value in a portfolio context by enhancing the risk-return profile. Hence, to add value, the correlations with other asset classes must be low, or even negative.

As we notice from Table 8.2, the pairwise correlations of the commodity futures and the traditional asset classes are low, and in some cases they are even negative, with some exceptions, for example, the commodity index TR Equal Weight CCI show high and positive correlation with the stock indices. Thus, by adding commodities to our portfolios we may benefit by diversification. As we have discussed, there are different approaches that can be used in order to select the best assets for long-term investing. One way is choosing the assets with the highest Sharpe Ratio. Another way is choosing the assets that are less risky, while fixing a constant level of expected return and vice versa. We can say for example, that it is much better investing in the 'BMUK10Y' index than in the 'S&P GSCI Gold Spot' index, because the first asset has a bigger expected return and a lower risk represented by the standard deviation.

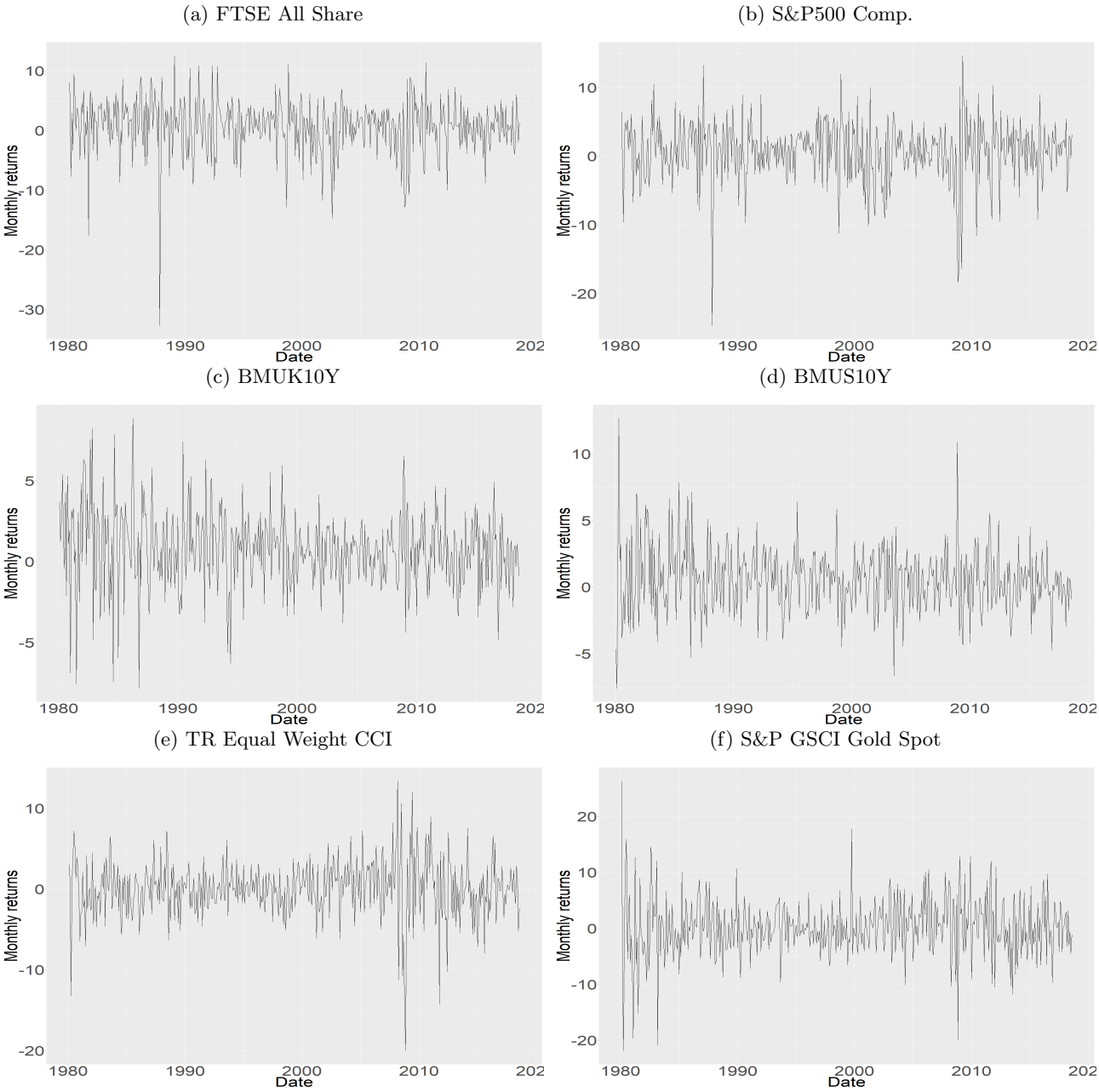


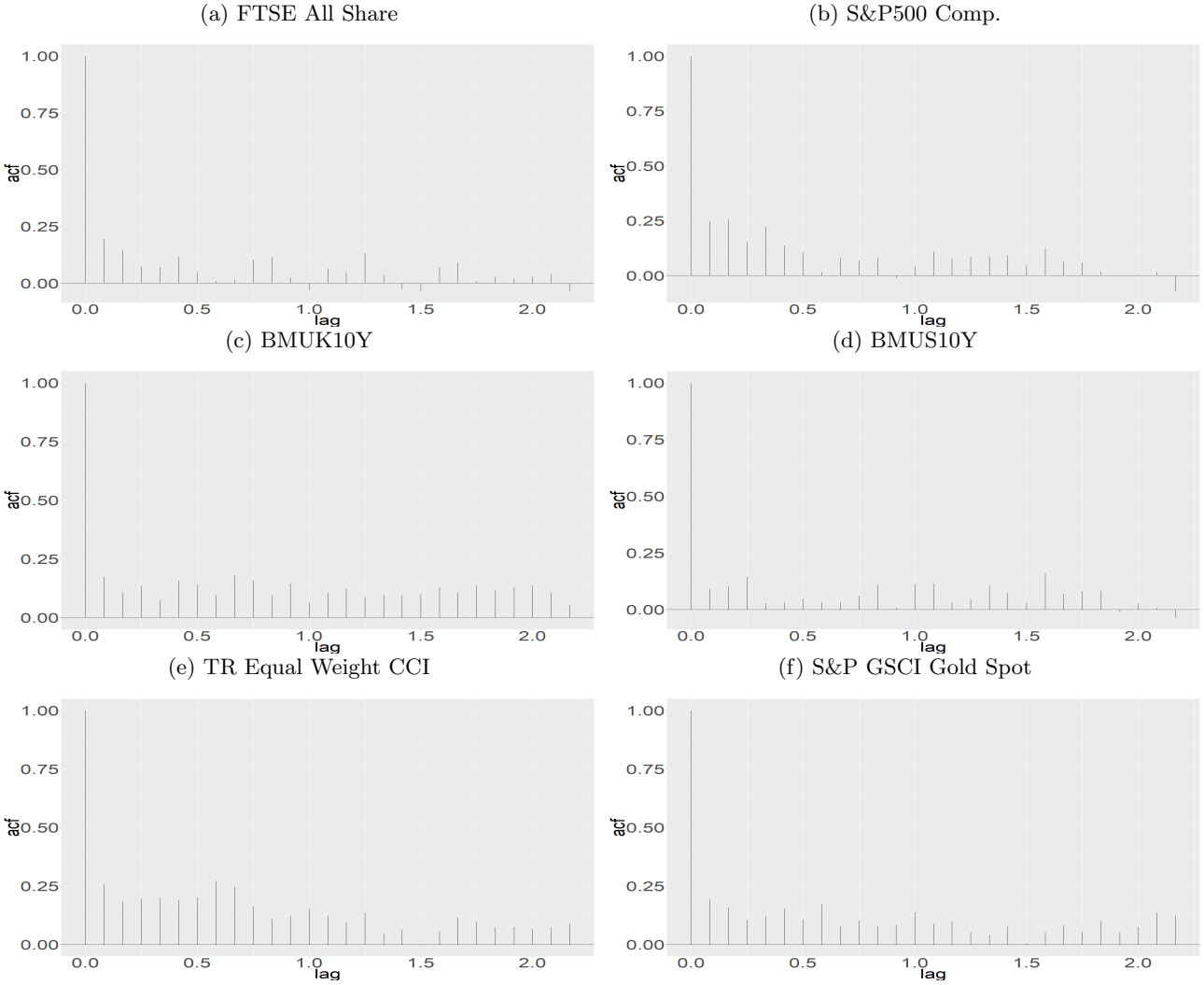
Figure 8.3: Non-standardized log-returns.

The covariance matrix is similar to the correlation matrix, but when the covariance is calculated, the data are not standardized, that is, we use the standard deviation for the whole sample period and multiply with the correlation. As a reminder, the co-variance equation is defined in eq.(5.26). Since the goal is to produce adequately diversified portfolios, we must have combination of the right assets in the right proportions with the goal of manipulating risk and return to work in our favour as much as possible. This secure that when one return falls, another one will rise to offset the loss. This is the core wisdom when investing in risky assets, that is, always keep a diversified portfolio in order to keep the level of risk as low as possible. Therefore, relying on the asymmetric dependence structure within the market, and by adjustments in the covariance matrix, we will

be able to manipulate the risk-and-return characteristics in our favor, hopefully. Note that local correlation- and covariances estimated at grid-point 0 are approximately equal to the global estimates.

When analyzing financial time series, we expect the return series to display time-conditional structure of volatility, that is, periods of high or low volatility are most probably followed by high or low volatility in the following month (see Figure 8.3). Thus, strong evidence of volatility clustering supports the stylized fact that there is far more predictability in conditional volatility than in returns. However, we want to advantage of these volatility clusters by using the local Gaussian correlation model. We will get back to this later in the analysis.

Figure 8.4: Auto-correlation plot of the absolute monthly log-return indices.



Another typical stylized fact is the autocorrelation, also known as *serial correlation* of price changes (*returns*), which in general is largely insignificant, but some small and interesting anomalies do exists, as illustrated in Figure 8.4. Given that the absolute values of returns can be used as a proxy for volatility, the autocorrelation of the absolute returns indicate the time dependency of the volatility. The autocorrelation describes the correlation between observations at two different times, hence if the return series are independent over time, the absolute values of the returns should be uncorrelated. The presence of volatility clustering can therefore be detected by a strong autocorrelation in the absolute values of the returns. As can be seen by the plots in Figure 8.4, there are positive values for a relatively large number of lags, which is a good indication of volatility clustering, and that a GARCH model might fit the data well. The autocorrelation function of the volatility in Figure 8.4 exhibits long-range dependence of approximately 20 trading months.

8.1.1 GARCH Filtration and Forecasting

With an understanding of how prices behave we will be able to make even better decisions, therefore we want tools such as financial models to generate forecasts for both future value and volatility of the returns. Moreover, it is important to have knowledge of the uncertainty of such forecasts. The sampling-window is one important aspect and will be a trade-off between having enough data for the parameter estimation to be stable, and too

much data so that the forecast do not reflect the current market conditions. For example, extreme fluctuation in the past can have great influence on the long-term volatility forecasts made today, hence a very long time series with several outliers is unlikely to be suitable. It is therefore very important to decide whether major market events from several years ago should be influencing forecasts today. On the other side, if the data period is too small, the estimates of the GARCH model might lack stability as the data window is slightly altered.

Table 8.3: Estimated coefficients using the t-GARCH(1,1) model.

Market	μ	ω	α	β	θ
FTSE All Share	0.0102	7.35e-04	0.3186	0.3653	4.9911
	(1.69e-03)	(1.81e-04)	(0.1053)	(4.24e-02)	(1.0710)
S&P500 Comp.	9.67e-03	1.22e-04	0.1713	0.7723	6.5442
	(1.62e-03)	(4.42e-05)	(4.93e-02)	(2.86e-02)	(1.73)
BMUK10Y Index	6.49e-03	7.52e-06	5.48e-02	0.9300	9.64
	(9.76e-04)	(4.42e-06)	(2.24e-02)	(1.82e-02)	(4.20)
BMUS10Y Index	5.32e-04	7.89e-05	9.00e-02	0.8780	7.22
	(1.98e-03)	(3.15e-05)	(3.18e-02)	(2.21e-02)	(2.19)
TR Equal Weight CCI	5.06e-03	6.79e-06	3.42e-02	0.9510	7.73
	(9.98e-04)	(3.66e-06)	(1.45e-02)	(1.24e-02)	(2.33)
S&P GSCI Gold Spot	6.94e-04	2.53e-05	7.16e-02	0.9040	8.75
	(1.33e-03)	(1.18e-05)	(2.22e-02)	(1.66e-02)	(3.24)

Table 8.3 provides the fitted values for the parameter set, which are obtained by minimizing the "negative" log-likelihood function given in eq.(3.9). As for the GARCH(1,1) model, we start with $\alpha = 0.1$ and $\beta = 0.8$ as typical values of financial time series, and set ω as the variance adjusted by the persistence $\omega = Var(x)(1 - \alpha - \beta)$.

A major objective of the investigation of heteroskedastic time series is forecasting. Expressions for forecasts of both the conditional mean and the conditional variance can be derived, with the two properties capable of being forecast independently of each other. For a GARCH(1,1) process, the n-step-ahead forecast of the conditional variance $\hat{\sigma}^2_{t+n|t}$ is computed recursively from

$$\hat{\sigma}^2_{t+n|t} = \hat{\omega} + \hat{\alpha}\epsilon^2_{t+n-1|t} + \hat{\beta}\sigma^2_{t+n-1|t}, \text{where } \epsilon^2_{t+1|t} = \sigma^2_{t+1|t}.$$

The goodness-of-fit is evaluated by checking the significance of the parameter estimates and measuring how well it models the volatility of the process. If a GARCH-model adequately captures volatility clustering, the absolute values of the standardized returns, given by

$$\epsilon^*_t = \frac{\hat{\epsilon}_t}{\hat{\sigma}_t}$$

where $\hat{\sigma}_t$ is the estimate of the volatility, should have no autocorrelation. This can be tested by a graphical inspection of the autocorrelation function, or by more formal statistical tests, such as the Ljung-Box statistic. Little or no auto-correlation implies an adequately good fit of the t-GARCH(1,1) model.

For additional information on the t-GARCH(1,1) process, see Charpentier [2015], which provides R scripts and additional restrictions- and/or extensions to the model introduced in this thesis.

8.2 Empirical analysis of dependence structures

In this section we apply the LGC model to describe monthly return data from financial markets. We treat these return data as coming from the same bivariate density f, and we assume stationary, which of course, depends upon the time scale and length of the time interval investigated, but assuming stationarity for such a long time period is hardly realistic. As mentioned earlier, the estimator of local correlation only depends upon two smoothing parameters, which will be crucial in the process of describing the dependence structure, especially within restricted areas with few observations. We will not delve into the adaptability and other properties of the smoothing parameters in this thesis, however this could be more closely investigated in future works. For greater clarity, we may restrict ourselves by only looking at the local correlation values on the diagonal.[1]

Figure 8.5 shows the diagonal values of the estimated pairwise local correlations between the stock returns. It is seen that the bivariate return distribution is not Gaussian, since in particular there are large local correlation values for both large negative and large positive returns. Although, the estimates for extreme values are subject to relatively large estimation errors due to the limited number of observations, hence we restrict ourselves by

[1]This main reference in this sub-section is Støve and Tjøstheim [2014]. The paper also include references for further reading.

only looking at the grid-map of size 4×4. This caused a lot of problems when we first introduced the moving-grid strategy, in fact, the moving-grid resulted in unstable estimates, and which did not perform quite well. Hence, after applying many restrictions, we came to the conclusion that the best performing moving-grid strategy was in fact the one that only operated under normal market conditions, that is, around grid-point 0 on the diagonal.

Furthermore, Figure 8.6 shows the estimated pairwise local correlation diagonal values within each asset class from the three markets investigated, all using rolling sampling windows of 120 and 240 trading months, again computed as if stationarity holds within the entire sampling period. Clearly, we find evidence of correlation asymmetry in all three markets, and in particular larger local correlations in bear markets (red line) for the stock indices. We also see that the local correlation estimates decreases in normal and bull markets, compared to bear market correlations. This implies that diversification opportunities would be largest in a bull market – or rather erode in bear markets (e.g., a downward price trend), which is when they are needed the most. Also note that the LGC curves might be quite different for different pairs of indices. Hence, from the plots given in Figure 8.6, we can tell that the LGC curves give a much more detailed picture than ordinary global correlation measures, and interestingly, by relying on the higher correlation that is present in a falling market, the results provide evidence that accounting for correlation asymmetries enhance performance gains.

Figure 8.7, 8.8 and 8.9 show the pairwise diagonal local correlation values across the three markets. The same pattern as before is present, that is, asymmetric dependence structure. In Figure 8.7 however, note how the local correlations are higher within bull markets as opposed to normal- and bear market conditions, and how the rolling correlation values are falling with time. Also note how the local correlation between each asset return series behave in 2007-2008, when the financial crisis took place. In Figure 8.8, we report the rolling local correlation between stock and commodity indices. Again, a large degree of correlation asymmetry is present, but as opposed to the plots in Figure 8.7, the correlation seems to increase with time. Note that the absolute value of the diagonal correlation always seem to be higher under bear market conditions, that is, stronger

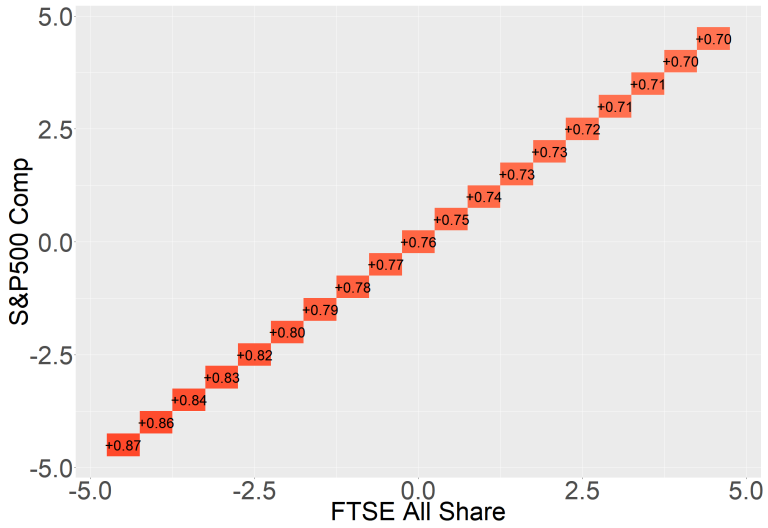


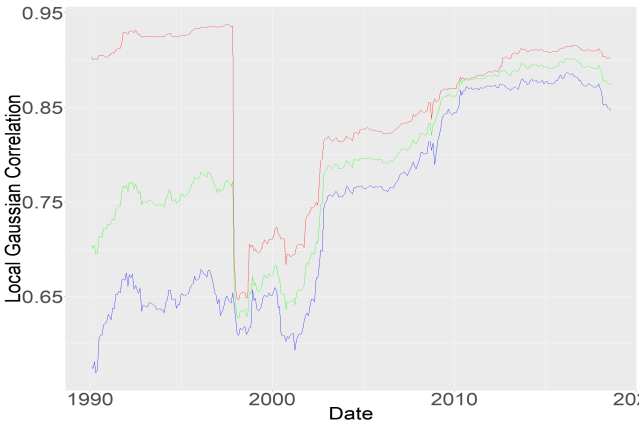
Figure 8.5: Out-of-sample diagonal local Gaussian correlation values for FTSE All Share vs S&P500 Comp. returns, from January 1980 to August 2018.

correlation values for negative values as well as for positive values. The rolling local correlation between bond and commodity indices are reported in Figure 8.9, and again, asymmetric correlation are reported across the markets. Hence, the LGC properties seems in fact to be asymmetric as far as bear and bull markets are concerned, and therefore, the change in prices are non-normally distributed as stated previously in the analysis. Much of the variation in the LGC however, could be expected to be due to volatility, and as previously mentioned, models such as the Student t-GARCH(1,1) process could be fitted to the data before analyzing them. This have been done in earlier studies (see e.g. Støve and Tjøstheim [2014]), where they do find similar asymmetric effects in the correlation for standardized returns as for non-standardized returns, but in less extent after filtrating the residuals and removing the volatility. We believe our results to be consistent across such models, but clearly this should be more closely investigated as we use different data sets.

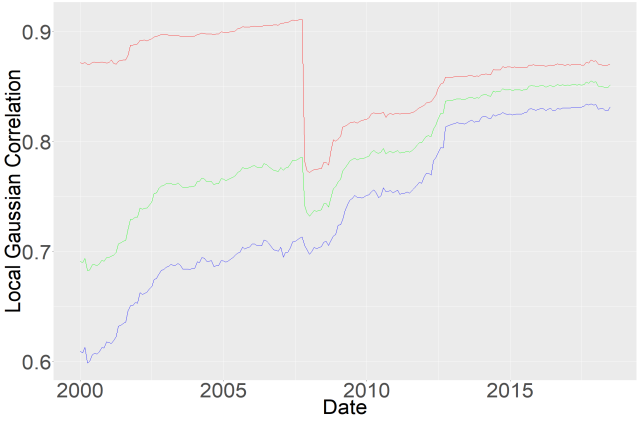
As we already know, an increased dependence in financial markets during a crisis period will imply that the diversification effect could be less than anticipated. In the following sections we will therefore examine how the LGC and adjustments in the covariance matrix by accounting for asymmetries primarily can be used to hopefully improve the classical Markowitz portfolio theory, and in particular examine how inclusion of different asset classes to a portfolio could increase the potential benefits of diversification. As a reminder, low correlated assets are good for diversification, while highly correlated assets should be avoided, hence we could probably

say that bonds will be a major part of our investment as it is very low and even negatively correlated with the other asset classes. This is of particular importance for large institutional investors with a long investment horizon.

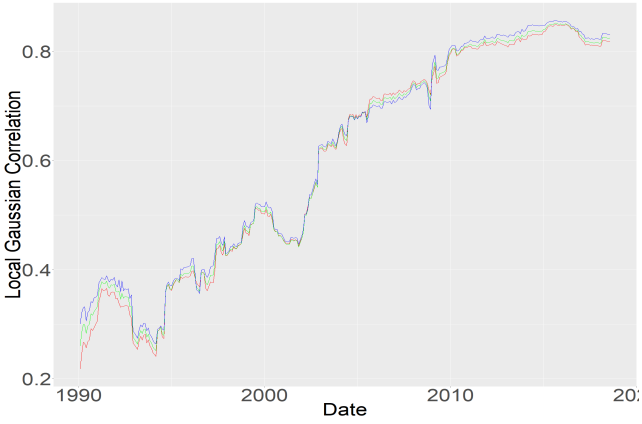
Figure 8.6: Diagonal local Gaussian correlation values of monthly returns within each market, using rolling sampling windows of 120 and 240 trading month. Color code; - bear-, - normal-, - bull market conditions.



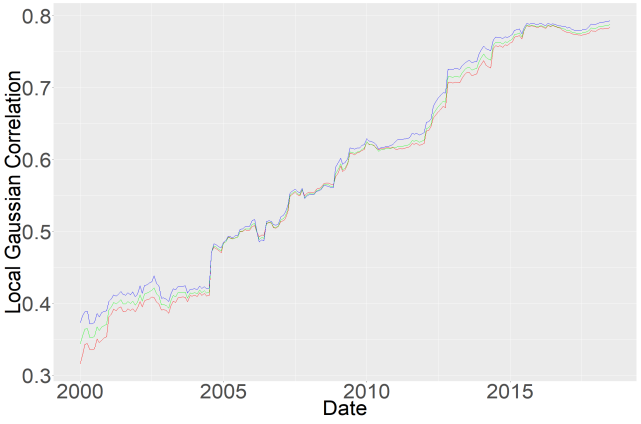
(a) FTSE All Share vs S&P500 Comp.



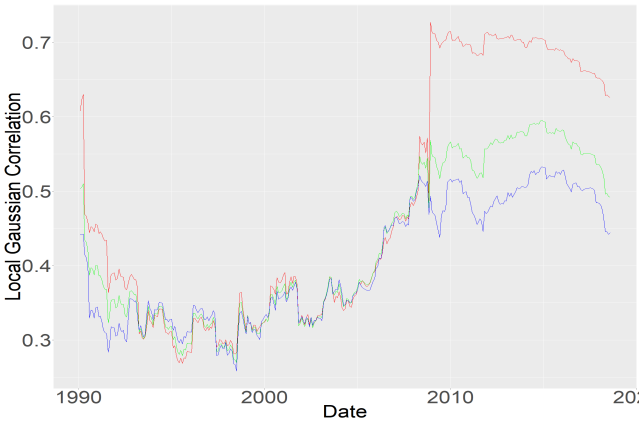
(b) FTSE All Share vs S&P500 Comp.



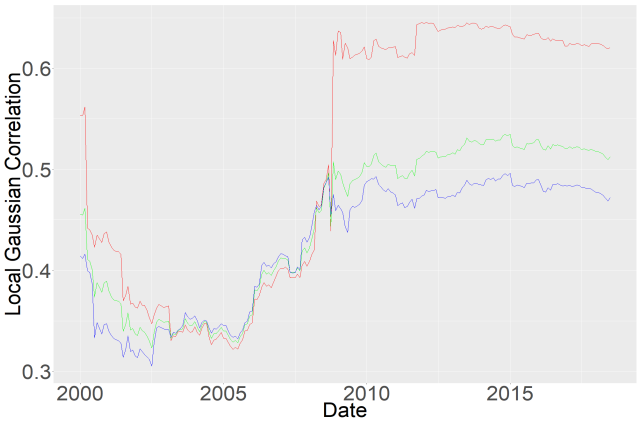
(c) BMUK10Y vs BMUS10Y



(d) BMUK10Y vs BMUS10Y



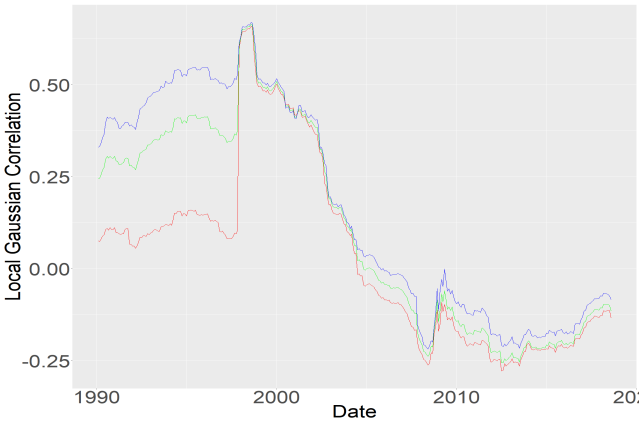
(e) TR Equal Weight CCI vs S&P GSCI Gold



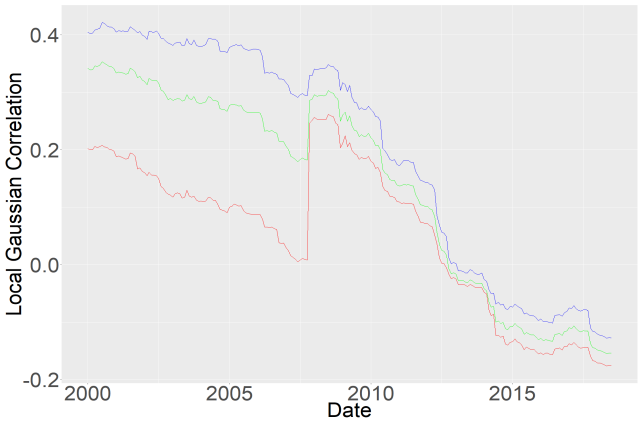
(f) TR Equal Weight CCI vs S&P GSCI Gold

Note how the plots in Figure 8.6 corresponds to the plots in Figure 8.3, that is, periods of high volatility follow strong asymmetric correlation between the pairwise asset returns. For example, plot (a) and (b) in Figure 8.6. The local Gaussian correlation clearly captures the asymmetries when we are in periods of high volatility and volatile clusters, as for example in the late 1990's. See Figre 8.3 for more information. Similarly, plot (e) and (f) in Figure 8.6, the strong asymmetric dependence within the market happening around 2008 are explained by period of high volatility, especially for negative returns. These observations are also consistent with the results discussed earlier (see Figure 8.4), there are positive values for a relatively large number of lags, which is a good indication of volatility clustering and asymmetric dependence structures within the markets.

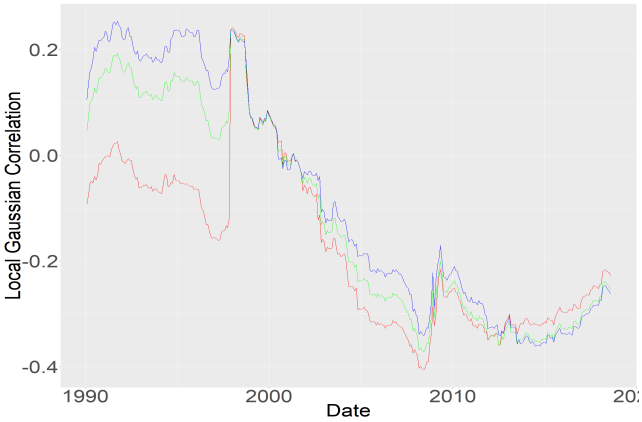
Figure 8.7: Diagonal local Gaussian correlation values between pairwise stock vs bond indices, using rolling sampling windows of 120 and 240 trading month. Color code; - bear-, - normal-, - bull market conditions.



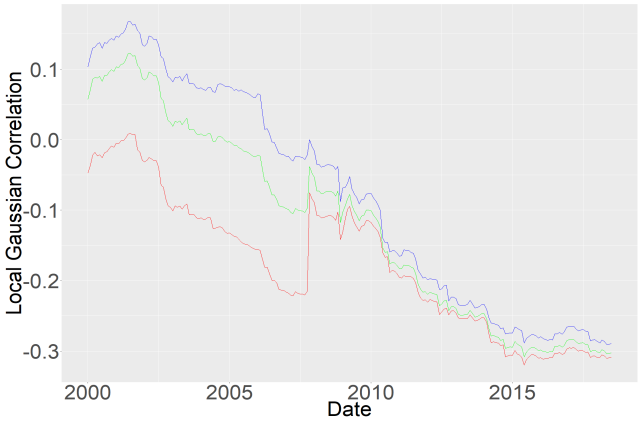
(a) FTSE All Share vs BMUK10Y



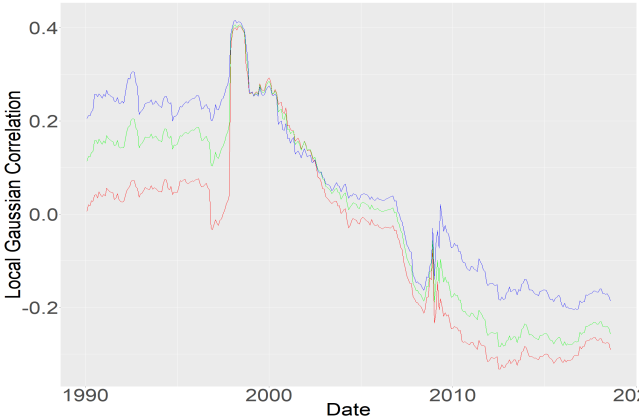
(b) FTSE All Share vs BMUK10Y



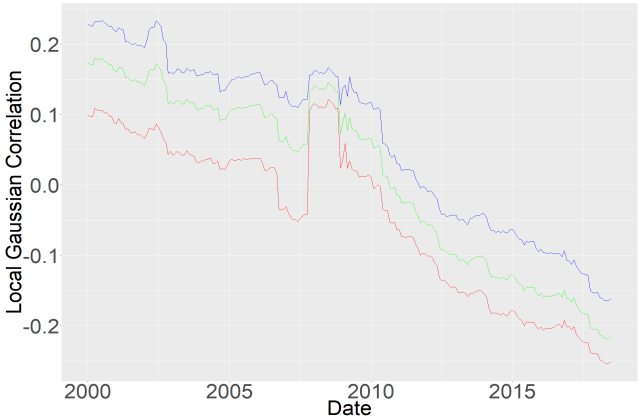
(c) FTSE All Share vs BMUS10Y



(d) FTSE All Share vs BMUS10Y

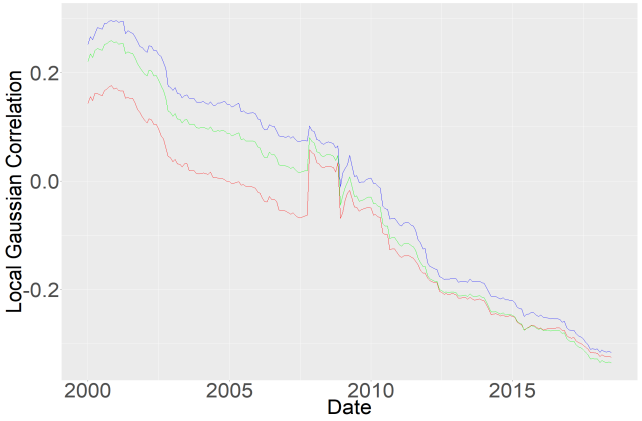


(e) S&P500 Comp. vs BMUK10Y



(f) S&P500 Comp. vs S&P BMUK10Y

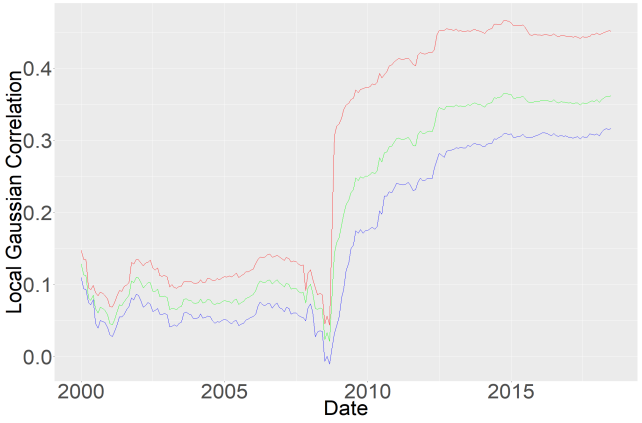
(g) S&P500 Comp. vs BMUS10Y



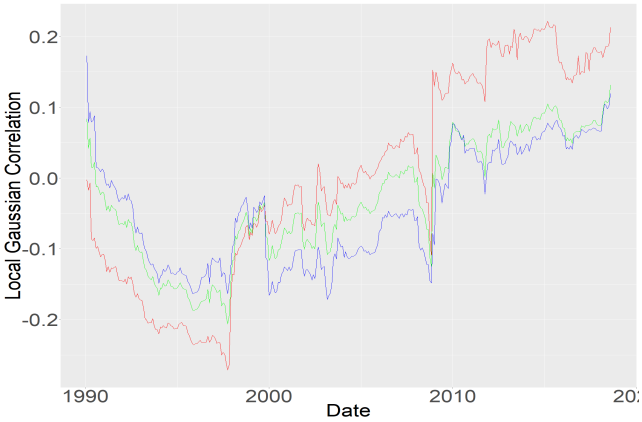
(h) S&P500 Comp. vs BMUS10Y

Figure 8.8: Diagonal local Gaussian correlation values between pairwise stock vs commodity indices, using rolling sampling windows of 120 and 240 trading month. Color code; - bear-, - normal-, - bull market conditions.

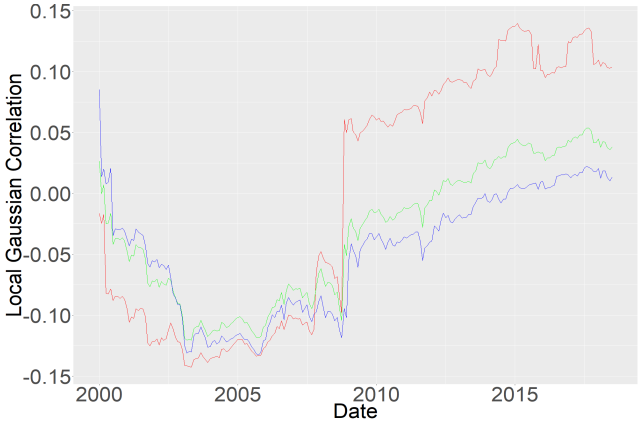
(a) FTSE All Share vs TR Equal Weight CCI



(b) FTSE All Share vs TR Equal Weight CCI

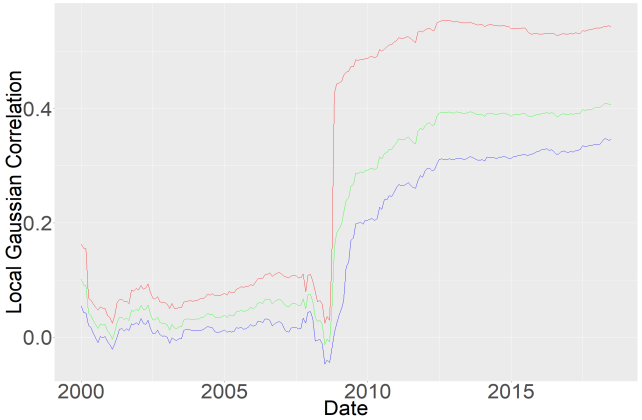


(c) FTSE All Share vs S&P GSCI Gold



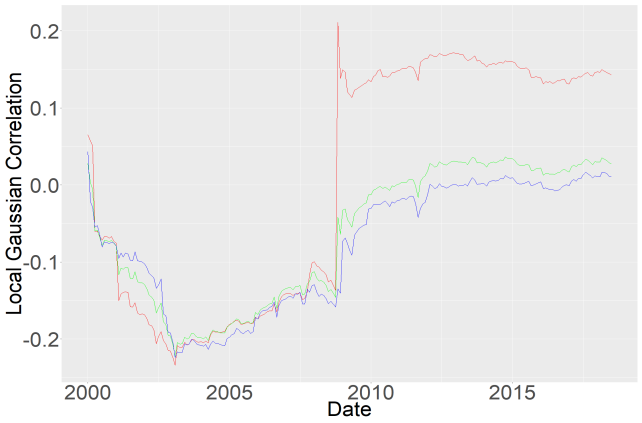
(d) FTSE All Share vs S&P GSCI Gold

(e) S&P500 Comp. vs TR Equal Weight CCI



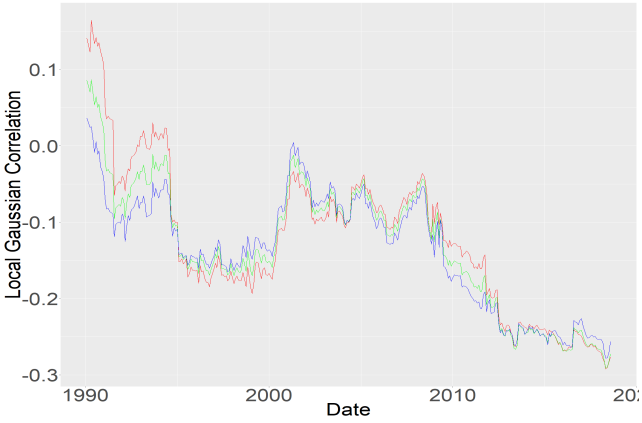
(f) S&P500 Comp. vs TR Equal Weight CCI

(g) S&P500 Comp. vs S&P GSCI Gold

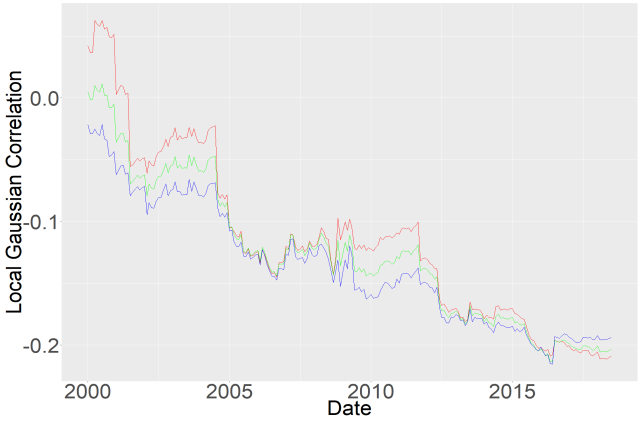


(h) S&P500 Comp. vs S&P GSCI Gold

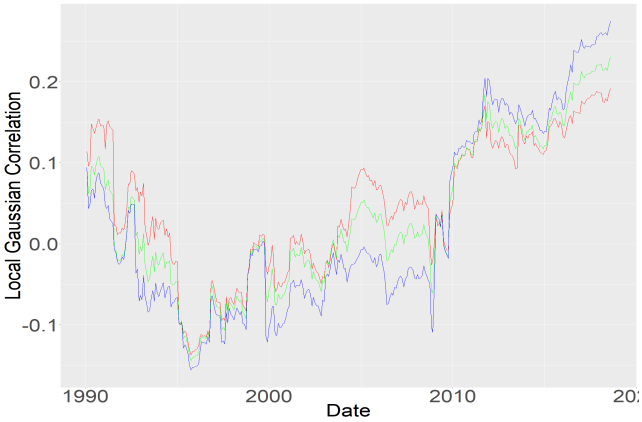
Figure 8.9: Diagonal local Gaussian correlation values between pairwise bond vs commodity indices, using rolling sampling windows of 120 and 240 trading month. Color code; - bear-, - normal-, - bull market conditions.



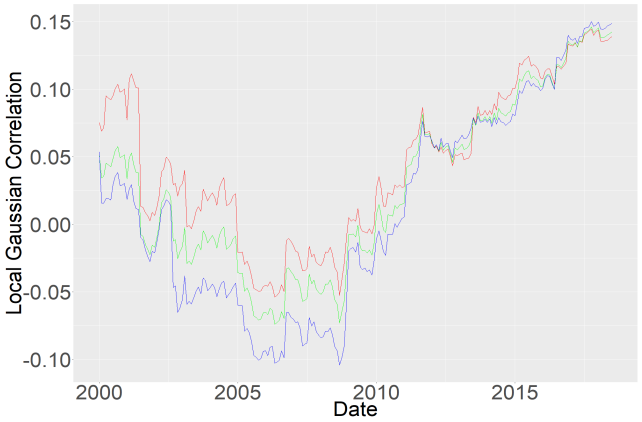
(a) BMUK10Y vs TR Equal Weight CCI



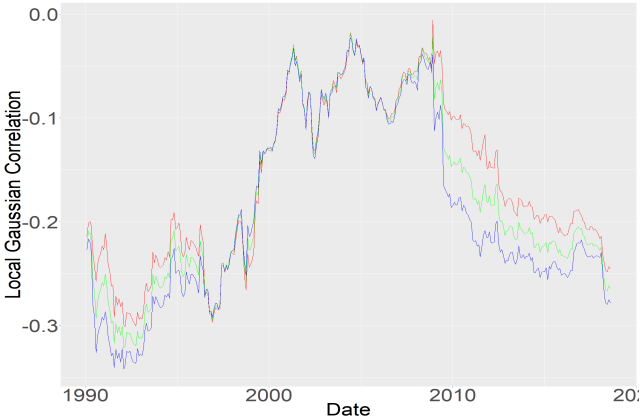
(b) BMUK10Y vs TR Equal Weight CCI



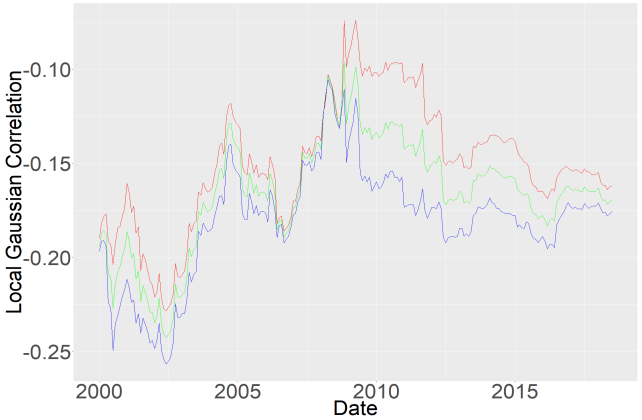
(c) BMUK10Y vs S&P GSCI Gold



(d) BMUK10Y vs S&P GSCI Gold

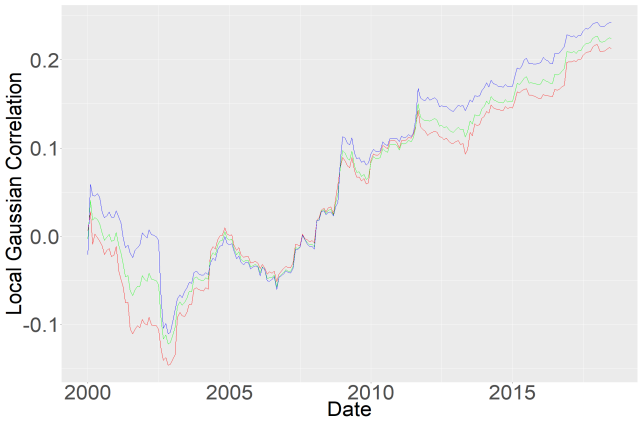


(e) BMUS10Y vs TR Equal Weight CCI



(f) BMUS10Y vs TR Equal Weight CCI

(g) BMUS10Y vs S&P GSCI Gold



(h) BMUS10Y vs S&P GSCI Gold

8.2.1 Buy-and-Hold performance

In this section we apply the Buy-and-Hold investment strategy on our data set across the portfolio optimization strategies, using sampling windows of 120 and 240 trading months, respectively. The Buy-and-Hold strategy is simply a "do-nothing" strategy, in which we as an investor buy and holds a portfolio of assets according to an initial combination of weights over a long period regardless of short-term fluctuations. This is based on the view that in long term conditions, financial markets give a good rate of return even while taking into account a large degree of volatility. Hence, the $1/N$ strategy is similar to the Buy-and-Hold strategy where the investor exhibits inertia as the default asset allocation is accepted and avoids future reinvestment and rebalancing decisions and associated costs. In the Buy-and-Hold investment analysis we therefore compare our local portfolio estimates to the global portfolio estimates, relative to the benchmark $1/N$ portfolio. More precisely, we compare local correlated estimates from bear-, normal- and bull market conditions to global estimates, and report the out-of-sample cumulative return performance across all portfolio optimization rules investigated.

In Table 8.5 (8.10) we report the global correlation- and covariances, and the corresponding local correlation- and covariance estimates in Table 8.6 (8.11), 8.7 (8.12) and 8.8 (8.13) under bear-, normal- and bull market conditions, for the given sampling windows of 120 (240) trading months, respectively. We then fit the estimated covariances into the portfolio allocation models and analyzes the cumulative performance graphs given in Figure 8.10 and 8.11, across the portfolio strategies with an initial wealth of \$1. In Table 8.4 (8.9) we report the estimated portfolio target weights for each portfolio allocation model. In Panel X1, A1 and B1 we report the portfolio target weights for the MVS portfolio strategy. In Panel X2, A2 and B2, we report the MIN portfolio target weights. In Panel Y1, C1 and D1 we report the portfolio target weights for the MVSC portfolio strategy, and in Panel Y2, C2 and D2 we report the portfolio target weights for the MINC portfolio strategy, respectively. Note that in Panel B2 for both Table 8.4 and 8.9, we are heavy on short-selling and produce extreme portfolio target weights when we use local s.d. to estimate the covariances, especially under bear market conditions. Further, note how local estimates from normal market conditions corresponds to the results from global estimates. For more details, see Table 8.4 and 8.9. Furthermore, we do not report the portfolio target weights for "5par-marginals-fixed" estimates, but we do report the portfolio performance in the sense of Sharpe Ratio given in Table 8.14 and 8.15, and the accumulated performance graphs in Figure A.3 and A.4 back in the Appendix. Furthermore, we also fit weighted portfolio target weights, estimated from bear- and normal market conditions, that is, 80% from bear market conditions, and 20% from normal market conditions.

When we fit local portfolio estimates using global standard deviation, we use local correlation estimates from grid-point -4 (bear market), 0 (normal market) and 4 (bull market), respectively. As for estimates using local standard deviation, we also use local correlation estimates from grid-point 0 (normal market) and 4 (bull market), but in the process of finding covariances estimated under bear market conditions to produce portfolio weights, we first created a grid-map of approximately 801 grid-points on the diagonal. We start with the first observation at the leftmost grid-point, and loop towards the center of the grid-map until we find a covariance matrix that fit the portfolio allocation models. For a sampling window of $M = 120$ trading months, we find a "suitable" covariance matrix after 239 steps, and for the longer sampling window of $M = 240$ trading months, after 193 steps. The same procedure is implemented when using a rolling sampling window, as done in section 8.2.2. Hence, we implement a sort of "moving" grid-point for local covariances under bear market conditions.

From plot (a) in Figure 8.10, note that we get quite stable estimates when we use global standard deviation to calculate the covariances, and as can be seen, the local portfolios estimated under normal- and bull market conditions outperform the benchmarks. As for the estimates using local standard deviation, in plot (b), we note that we get similar performance using local estimates under normal- and bull market conditions, but as for bear market estimates, the performance is outstanding and outperform the benchmarks by far. If we look at the corresponding target portfolio weights in Table 8.4 (Panel B1), the model is short-selling heavy on the FTSE All Share Index, and invest a large amount in S&P500 Comp., BMUK10Y and TR Equal Weight CCI Index, which might explain the outstanding performance. Further, plot (c) and (d) are quite similar to the benchmark portfolios, but in plot (c) we do get the same results as earlier, that is, local estimates from normal- and bull markets are outperforming, whereas in plot (d), the local estimates under bear market conditions are outperforming. The corresponding plots in Figure 8.11 are consistent with these results. Also note that across the strategies investigated, most of the portfolio allocation models are outperforming the benchmark $1/N$ portfolio. Furthermore, in plot (e) and (f) in Figure 8.10, the MIN model using global standard deviations are outperformed by the benchmark 1/N portfolio, whereas local estimates using local standard deviations outperforms, but as we can see, the returns do fluctuates a lot. As for plot (g) and (h), the local portfolios are outperformed by the benchmark $1/N$ portfolio. For more information, see page 42-49.

Overall, as we can see in Panel A in Table 8.14, the local portfolios estimated under normal- and bull market conditions, and using global s.d. to calculate covariances, outperforms the benchmarks models, whereas local portfolios estimated under bear market conditions and using local s.d. outperforms. The "5par-marginals-fixed" estimates provide us with the same results as the last strategy discussed, more or less.

Sampling window size of 120 trading months from January 1980 to January 1990.

Table 8.4: Portfolio target weights across the optimization rules investigated ($M = 120$).

Portfolio strategy	FTSE-All-Share	S&P500 Comp.	BMUK10Y Index	BMUS10Y Index	TR-Equal-Weight-CCI	S&P-GSCI-Gold
Panel X1: *MVS portfolio target weights estimated with global correlation estimates*						
Markowitz	0.03107	0.05650	0.24970	0.35568	0.34841	-0.04137
Panel A1: *MVS portfolio target weights estimated with local correlation and global s.d.*						
Bear	0.06030	0.04111	0.21649	0.36973	0.37901	-0.06664
Norm.	0.00378	0.09237	0.26702	0.33536	0.34246	-0.04099
Bull	-0.00585	0.10028	0.29087	0.31239	0.33011	-0.02780
80-20%	0.04707	0.0529	0.23136	0.35826	0.36923	-0.05888
Panel B1: *MVS portfolio target weights estimated with local correlation and local s.d.*						
Bear	-0.2875	0.4349	0.38465	0.24020	0.29213	-0.06438
Norm.	0.03619	0.04873	0.24879	0.35793	0.34710	-0.03874
Bull	0.05357	0.04652	0.24457	0.34190	0.34053	-0.02709
80-20%	-0.21928	0.35722	0.35663	0.26054	0.30181	-0.05692
Panel X2: *MIN portfolio target weights estimates with global correlation estimates*						
Markowitz	0.01068	0.05594	0.20304	0.35516	0.41413	-0.03896
Panel A2: *MIN portfolio target weights estimated with local correlation and global s.d.*						
Bear	0.02556	0.06674	0.19091	0.36266	0.42684	-0.07272
Norm.	-0.00847	0.08236	0.22182	0.33914	0.40381	-0.03867
Bull	-0.01728	0.08333	0.23817	0.32084	0.39475	-0.01980
80-20%	0.01699	0.07006	0.20036	0.35430	0.42042	-0.06214
Panel B2: *MIN portfolio target weights estimated with local correlation and local s.d.*						
Bear	-4.84339	5.77303	1.77235	-1.49056	0.04512	-0.25655
Norm.	0.01813	0.04532	0.20303	0.35828	0.41135	-0.03610
Bull	0.03517	0.03695	0.19653	0.34629	0.40386	-0.01880
80-20%	-3.86768	4.62581	1.45718	-1.12319	0.11687	-0.20900
Panel Y1: *MVSC portfolio target weights estimated with global correlation estimates*						
Markowitz	0.034706	0.053253	0.25241	0.35340	0.30623	0.00
Panel C1: *MVSC portfolio target weights estimated with local correlation and global s.d.*						
Bear	0.09843	0.00	0.20728	0.37492	0.31937	0.00
Norm.	0.00632	0.09096	0.26991	0.33243	0.30038	0.00
Bull	0.00	0.09409	0.28856	0.31529	0.30206	0.00
80-20%	0.07874	0.01882	0.22353	0.36299	0.31591	0.00
Panel D1: *MVSC portfolio target weights estimated with local correlation and local s.d.*						
Bear	0.00	0.08692	0.26823	0.34749	0.29737	0.00
Norm.	0.03870	0.04713	0.25168	0.35510	0.30740	0.00
Bull	0.05289	0.04481	0.24749	0.34284	0.31198	0.00
80-20%	0.01058	0.07850	0.26408	0.34656	0.30029	0.00
Panel Y2: *MINC portfolio target weights estimats with global correlation estimates*						
Markowitz	0.01326	0.05287	0.20369	0.35299	0.37718	0.00
Panel C2: *MINC portfolio target weights estimated with local correlation and global s.d.*						
Bear	0.08862	0.00074	0.18788	0.37458	0.34817	0.00
Norm.	0.00	0.07473	0.21906	0.33870	0.36751	0.00
Bull	0.00	0.06806	0.22484	0.32767	0.37942	0.00
80-20%	0.07089	0.01421	0.19527	0.36520	0.35442	0.00
Panel D2: *MINC portfolio target weights estimated with local correlation and local s.d.*						
Bear	0.00	0.06591	0.21086	0.35259	0.37064	0.00
Norm.	0.01967	0.04368	0.20370	0.35566	0.37729	0.00
Bull	0.03353	0.03518	0.19543	0.34720	0.38864	0.00
80-20%	0.00671	0.05977	0.20777	0.35151	0.37424	0.00

Table 8.5: Global correlation- and covariance matrix ($M = 120$).

	FTSE-All-Share	S&P500 Comp.	BMUK10Y Index	BMUS10Y Index	TR-Equal-Weight-CCI	S&P-GSCI-Gold
Panel A: *Correlation matrix*						
FTSE All Share	1.00					
S&P500 Comp.	0.71008	1.00				
BMUK10Y Index	0.23992	0.11128	1.00			
BMUS10Y Index	0.04227	0.19304	0.26001	1.00		
TR Equal Weight CCI	0.14180	0.12358	0.08598	-0.22063	1.00	
S&P GSCI Gold	0.08338	0.11798	0.09311	-0.00554	0.50418	1.00
Panel B: *Covariance matrix*						
FTSE All Share	3.0209e-03					
S&P500 Comp.	1.8106e-03	2.1523e-03				
BMUK10Y Index	4.1536e-04	1.6263e-04	9.9219e-04			
BMUS10Y Index	7.3875e-05	2.8480e-04	2.6045e-04	1.0113e-03		
TR Equal Weight CCI	2.5056e-04	1.8431e-04	8.7070e-05	-2.2556e-04	1.03347e-03	
S&P GSCI Gold	3.1002e-04	3.7026e-04	1.9839e-04	-1.1918e-05	1.0964e-03	4.5757e-03

Correlation- and covariances used to estimate benchmark mean-variance and minimum-variance portfolios.

Table 8.6: Local correlation- and covariance estimates under bear market conditions ($M = 120$).

	FTSE-All-Share	S&P500 Comp.	BMUK10Y Index	BMUS10Y Index	TR-Equal-Weight-CCI	S&P-GSCI-Gold
Panel A1: *Correlation matrix*						
FTSE All Share	1.00					
S&P500 Comp.	0.90322	1.00				
BMUK10Y Index	0.07575	0.00583	1.00			
BMUS10Y Index	-0.09276	0.07418	0.21753	1.00		
TR Equal Weight CCI	0.15601	0.17290	0.14084	-0.21389	1.00	
S&P GSCI Gold	-0.00266	0.12095	0.11407	0.01667	0.60745	1.00
Panel B1: *Covariance matrix w/ global s.d.*						
FTSE All Share	3.0209e-03					
S&P500 Comp.	2.3031e-03	2.1523e-03				
BMUK10Y Index	1.3115e-04	8.5192e-06	9.9219e-04			
BMUS10Y Index	-1.6213e-04	1.0944e-04	2.1789e-04	1.0113e-03		
TR Equal Weight CCI	2.7565e-04	2.3267e-04	1.4262e-04	-2.1866e-04	1.0335e-03	
S&P GSCI Gold	-9.8937e-06	3.7956e-04	2.4305e-04	3.5868e-05	1.3210e-03	4.5757e-03
Panel A2: *Correlation matrix used to estimate covariances in Panel A2*						
FTSE All Share	1.00					
S&P500 Comp.	0.768749	1.00				
BMUK10Y Index	0.205056	0.083167	1.00			
BMUS10Y Index	0.014215	0.166754	0.247482	1.00		
TR Equal Weight CCI	0.147983	0.137714	0.101767	-0.2187874	1.00	
S&P GSCI Gold	0.059578	0.114190	0.096839	0.0043184	0.53040	1.00
Panel B2: *Covariance matrix w/ local s.d.*						
FTSE All Share	3.02089e-03					
S&P500 Comp.	2.47111e-03	2.15235e-03				
BMUK10Y Index	3.8128e-04	1.2543e-04	9.9219e-04			
BMUS10Y Index	2.612e-05	2.4908e-04	2.4395e-04	1.0113e-03		
TR Equal Weight CCI	2.8450e-04	2.3234e-04	1.0457e-04	-2.2169e-04	1.03347e-03	
S&P GSCI Gold	2.3392e-04	3.6822e-04	2.0635e-04	9.1084e-06	1.17586e-03	4.5757e-03

Note: Different correlation matrices are used to estimate covariances with global- and local s.d. (see Panel A1 and A2), respectively. This is based on NA produced portfolio weights, see section 7.1.2 for details.

Table 8.7: Local correlation- and covariance estimates under normal market conditions ($M = 120$).

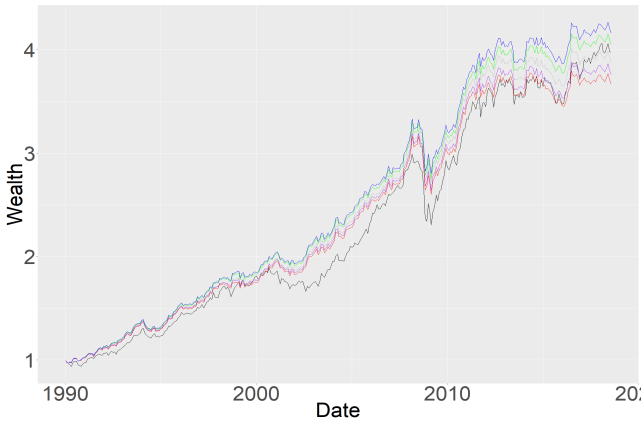
	FTSE-All-Share	S&P500 Comp.	BMUK10Y Index	BMUS10Y Index	TR-Equal-Weight-CCI	S&P-GSCI-Gold
Panel A: *Correlation matrix*						
FTSE All Share	1.00					
S&P500 Comp.	0.70401	1.00				
BMUK10Y Index	0.24579	0.11323	1.00			
BMUS10Y Index	0.04597	0.19571	0.26003	1.00		
TR Equal Weight CCI	0.14168	0.12295	0.08585	-0.22090	1.00	
S&P GSCI Gold	0.08345	0.11706	0.09256	-0.00432	0.50304	1.00
Panel B: *Covariance matrix w/ global s.d.*						
FTSE All Share	3.0209e-03					
S&P500 Comp.	1.7952e-03	2.1523e-03				
BMUK10Y Index	4.2553e-04	1.6547e-04	9.9219e-04			
BMUS10Y Index	8.0346e-05	2.8873e-04	2.6047e-04	1.0113e-03		
TR Equal Weight CCI	2.5033e-04	2.1130e-04	8.6934e-05	-2.2583e-04	1.0335e-03	
S&P GSCI Gold	3.1026e-04	3.6735e-04	1.9722e-04	-9.2919e-06	1.0939e-03	4.5757e-03
Panel C: *Covariance matrix w/ local s.d.*						
FTSE All Share	3.0209e-03					
S&P500 Comp.	1.7429e-03	2.15235e-03				
BMUK10Y Index	4.1851e-04	1.6346e-04	9.9219e-04			
BMUS10Y Index	7.9034e-05	2.8531e-04	2.5818e-04	1.0113e-03		
TR Equal Weight CCI	2.4627e-04	2.0876e-04	8.6165e-05	-2.2385e-04	1.0335e-03	
S&P GSCI Gold	3.0459e-04	3.6200e-04	1.9502e-04	-9.1873e-06	1.0801e-03	4.5757e-03

Table 8.8: Local correlation- and covariance estimates under bull market conditions ($M = 120$).

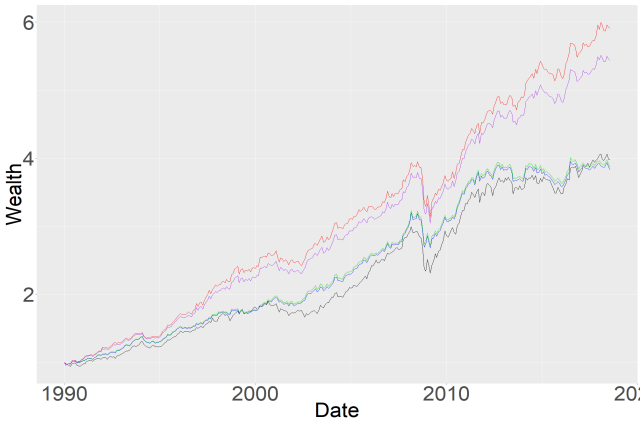
	FTSE-All-Share	S&P500 Comp.	BMUK10Y Index	BMUS10Y Index	TR-Equal-Weight-CCI	S&P-GSCI-Gold
Panel A: *Correlation matrix*						
FTSE All Share	1.00					
S&P500 Comp.	0.57614	1.00				
BMUK10Y Index	0.33040	0.204917	1.00			
BMUS10Y Index	0.10533	0.251826	0.300294	1.00		
TR Equal Weight CCI	0.11675	0.081831	0.036417	-0.22703	1.00	
S&P GSCI Gold	0.17272	0.156204	0.094382	-0.04606	0.441479	1.00
Panel B: *Covariance matrix w/ global s.d.*						
FTSE All Share	3.0209e-03					
S&P500 Comp.	1.4691e-03	2.1523e-03				
BMUK10Y Index	5.7201e-04	2.9946e-04	9.9219e-04			
BMUS10Y Index	1.8410e-04	3.7153e-04	3.0080e-04	1.0113e-03		
TR Equal Weight CCI	2.0628e-04	1.7412e-04	3.6877e-05	-2.3209e-04	1.0335e-03	
S&P GSCI Gold	6.4214e-04	4.9021e-04	2.0110e-04	-9.9081e-05	9.6004e-04	4.5757e-03
Panel C: *Covariance matrix w/ local s.d.*						
FTSE All Share	3.0209e-03					
S&P500 Comp.	9.4508e-04	2.1523e-03				
BMUK10Y Index	4.4979e-04	2.6854e-04	9.9219e-04			
BMUS10Y Index	1.5461e-04	3.4739e-04	3.0724e-04	1.0113e-03		
TR Equal Weight CCI	1.6022e-04	1.4895e-04	3.4205e-05	-2.3198e-04	1.0335e-03	
S&P GSCI Gold	5.6837e-04	4.7206e-04	2.0329e-04	-1.0452e-04	9.4574e-04	4.5757e-03

The plots on left side of Figure 8.10 use global s.d.-, whereas those on right side use local s.d. to estimate covariances. Color code; - benchmark $1/N$-, - global-, - bear-, - normal-, - and bull-market conditions. The - 80-20% weighted portfolio are estimated from bear- and normal market conditions, respectively.

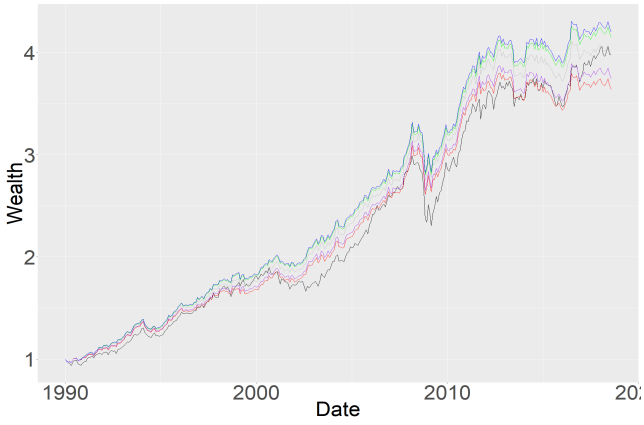
Figure 8.10: Buy-and-Hold portfolio performance using a sampling window of 120 trading months.



(a) Mean-Variance with short-selling (MVS)

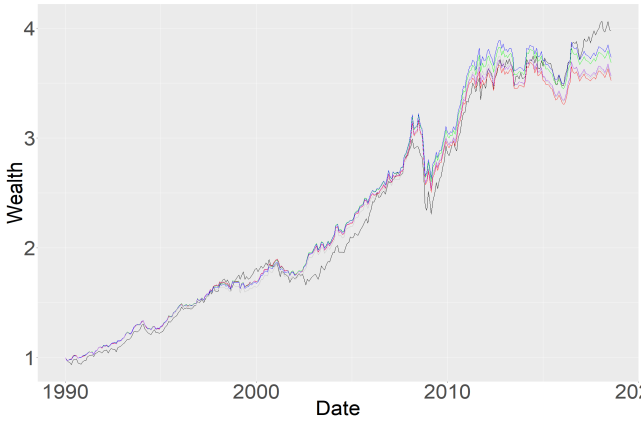


(b) Mean-Variance with short-selling (MVS)

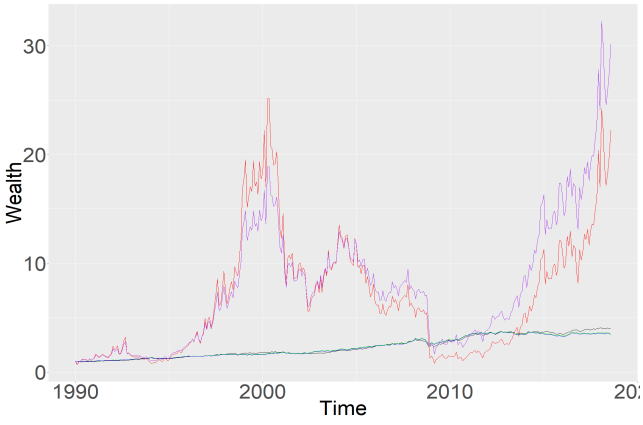


(c) Mean-Variance without short-selling (MVSC)

(d) Mean-Variance without short-selling (MVSC)

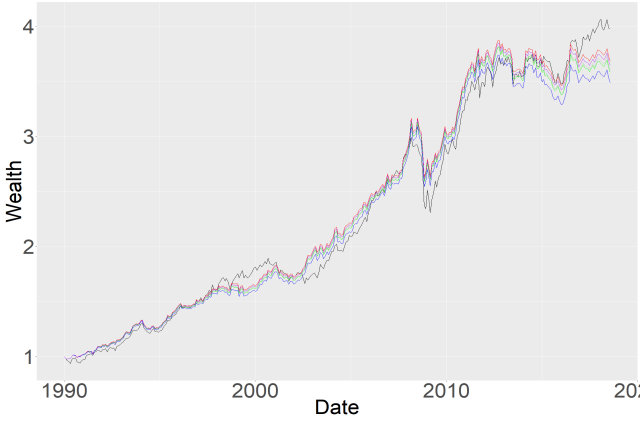


(e) Minimum-Variance with short-selling (MIN)



(f) Minimum-Variance with short-selling (MIN)

(g) Minimum-Variance without short-selling (MINC)



(h) Minimum-Variance without short-selling (MINC)

Sampling window size of 240 trading months from January 1980 to January 2000.

Table 8.9: Portfolio target weights across the optimization rules investigated ($M = 240$).

Portfolio strategy	FTSE-All-Share	S&P500 Comp.	BMUK10Y Index	BMUS10Y Index	TR-Equal-Weight-CCI	S&P-GSCI-Gold
Panel X1: *MVS portfolio target weights estimated with global correlation estimates*						
Markowitz	-0.00274	0.08937	0.27556	0.33902	0.32263	-0.02385
Panel A1: *MVS portfolio target weights estimated with local correlation and global s.d.*						
Bear	0.02948	0.06636	0.24698	0.36008	0.34512	-0.04802
Norm.	-0.01869	0.10581	0.28642	0.32464	0.32312	-0.02131
Bull	-0.03267	0.11574	0.30687	0.30461	0.31829	-0.01284
80-20%	0.01705	0.07624	0.25896	0.34899	0.33976	-0.04099
Panel B1: *MVS portfolio target weights estimated with local correlation and local s.d.*						
Bear	-0.64068	0.76499	0.46147	0.05189	0.40037	-0.03805
Norm.	0.00620	0.07737	0.27461	0.34459	0.31915	-0.02193
Bull	0.02721	0.06823	0.26572	0.33983	0.31895	-0.01995
80-20%	-0.50710	0.62564	0.42232	0.10948	0.38408	-0.03443
Panel X2: *MIN portfolio target weights estimated with global correlation estimates*						
Markowitz	0.00049	0.04693	0.19538	0.35574	0.41513	-0.01367
Panel A2: *MIN portfolio target weights estimated with local correlation and global s.d.*						
Bear	0.05836	-0.00120	0.16937	0.38490	0.42996	-0.04139
Norm.	-0.00657	0.05494	0.20523	0.34862	0.41067	-0.01288
Bull	-0.02260	0.06499	0.22044	0.33240	0.40529	-0.00053
80-20%	0.04217	0.01204	0.17958	0.37440	0.42503	-0.03322
Panel B2: *MIN portfolio target weights estimated with local correlation and local s.d.*						
Bear	-4.25489	4.92151	2.08459	-1.72816	0.09044	-0.11349
Norm.	0.01015	0.03201	0.19315	0.36199	0.41471	-0.01201
Bull	0.01762	0.02982	0.18459	0.35731	0.41589	-0.00523
80-20%	-3.40039	3.94317	1.70459	-1.31106	0.15553	-0.09184
Panel Y1: *MVSC portfolio target weights estimated with global correlation estimates*						
Markowitz	0.00	0.09021	0.27663	0.33603	0.29712	0.00
Panel C1: *MVSC portfolio target weights estimated with local correlation and global s.d.*						
Bear	0.04875	0.05160	0.24568	0.36205	0.29190	0.00
Norm.	0.00	0.09413	0.27859	0.32803	0.29926	0.00
Bull	0.00	0.09506	0.28704	0.31552	0.30238	0.00
80-20%	0.03901	0.06029	0.25396	0.35274	0.29400	0.00
Panel D1: *MVSC portfolio target weights estimated with local correlation and local s.d.*						
Bear	0.00	0.08705	0.2753	0.34216	0.29549	0.00
Norm.	0.00660	0.08023	0.27639	0.34088	0.29589	0.00
Bull	0.02668	0.07089	0.26770	0.33715	0.29758	0.00
80-20%	0.00534	0.08382	0.27379	0.34116	0.29590	0.00
Panel Y2: *MINC portfolio target weights estimated with global correlation estimates*						
Markowitz	0.00111	0.04728	0.19471	0.35407	0.40282	0.00
Panel C2: *MINC portfolio target weights estimated with local correlation and global s.d.*						
Bear	0.06310	0.00	0.16871	0.38120	0.38698	0.00
Norm.	0.00	0.05050	0.20103	0.34935	0.39911	0.00
Bull	0.00	0.04864	0.20418	0.34181	0.40537	0.00
80-20%	0.05048	0.00973	0.17581	0.37333	0.39066	0.00
Panel D2: *MINC portfolio target weights estimated with local correlation and local s.d.*						
Bear	0.00 0.04374	0.19764	0.35891	0.39971	00	
Normal	0.01043	0.03274	0.19264	0.36031	0.40388	0.00
Bull	0.01739	0.03010	0.18426	0.35681	0.41144	0.00
80-20%	0.00348	0.04102	0.19496	0.35849	0.40206	0.00

Table 8.10: Global correlation- and covariance matrix ($M = 240$).

	FTSE-All-Share	S&P500 Comp.	BMUK10Y Index	BMUS10Y Index	TR-Equal-Weight-CCI	S&P-GSCI-Gold
Panel A: *Correlation matrix*						
FTSE All Share	1.00					
S&P500 Comp.	0.69511	1.00				
BMUK10Y Index	0.33687	0.17212	1.00			
BMUS10Y Index	0.05428	0.21781	0.34339	1.00		
TR Equal Weight CCI	0.12878	0.10197	0.00520	-0.19422	1.00	
S&P GSCI Gold	0.02696	0.02915	0.05149	-0.00434	0.45683	1.00
Panel B: *Covariance matrix*						
FTSE All Share	2.4110e-03					
S&P500 Comp.	1.4208e-03	1.7327e-03				
BMUK10Y Index	4.6266e-04	2.0039e-04	7.8234e-04			
BMUS10Y Index	7.0757e-05	2.4068e-04	2.5497e-04	7.0471e-04		
TR Equal Weight CCI	1.7181e-04	1.1533e-04	3.9541e-06	-1.4009e-04	7.3824e-04	
S&P GSCI Gold	7.1259e-05	6.5311e-05	7.7526e-05	-6.2081e-06	6.6806e-04	2.8969e-03

Correlation- and covariances used to estimate benchmark mean-variance and minimum-variance portfolios.

Table 8.11: Local correlation- and covariance estimates under bear market conditions ($M = 240$).

	FTSE-All-Share	S&P500 Comp.	BMUK10Y Index	BMUS10Y Index	TR-Equal-Weight-CCI	S&P-GSCI-Gold
Panel A1: *Correlation matrix*						
FTSE All Share 1.00						
S&P500 Comp.	0.87171	1.00				
BMUK10Y Index	0.20149	0.09996	1.00			
BMUS10Y Index	-0.04761	0.14252	0.31573	1.00		
TR Equal Weight CCI	0.14727	0.16306	0.04212	-0.19011	1.00	
S&P GSCI Gold	-0.01646	0.06493	0.07526	0.006011	0.55338	1.00
Panel B1: *Covariance matrix w/ global s.d.*						
FTSE All Share	2.4110e-03					
S&P500 Comp.	1.7817e-03	1.7327e-03				
BMUK10Y Index	2.7673e-04	1.1639e-04	7.8234e-04			
BMUS10Y Index	-6.2067e-05	1.5750e-04	2.3444e-04	7.0471e-04		
TR Equal Weight CCI	1.9649e-04	1.6657e-04	3.2015e-05	-1.3712e-04	7.3824e-04	
S&P GSCI Gold	-4.3508e-05	1.4548e-04	1.1330e-04	8.5888e-06	8.0927e-04	2.8969e-03
Panel A2: *Correlation matrix used to estimate portfolio weights*						
FTSE All Share	1.00					
S&P500 Comp.	0.7571768	1.00				
BMUK10Y Index	0.2976254	0.147775	1.00			
BMUS10Y Index	0.0245473	0.195701	0.332297	1.00		
TR Equal Weight CCI	0.1363272	0.124365	0.018281	-0.192798	1.00	
S&P GSCI Gold	0.0099515	0.036893	0.057704	0.001265	0.487257	1.00
Panel B2: *Covariance matrix w/ local s.d*						
FTSE All Share	2.4110e-03					
S&P500 Comp.	1.9522e-03	1.7327e-03				
BMUK10Y Index	4.3753e-04	1.7841e-04	7.8234e-04			
BMUS10Y Index	3.3424e-05	2.2012e-04	2.4342e-04	7.0471e-04		
TR Equal Weight CCI	1.9720e-04	1.6354e-04	1.4106e-05	-1.3819e-04	7.3824e-04	
S&P GSCI Gold	2.7698e-05	8.5452e-05	8.6922e-05	1.7723e-06	7.2930e-04	2.8969e-03

Note: Different correlation matrices are used to estimate covariances with global- and local s.d. (see Panel A1 and A2), respectively. This is based on NA produced portfolio weights, see section 7.1.2 for details.

Table 8.12: Local correlation- and covariance estimates under normal market conditions ($M = 240$).

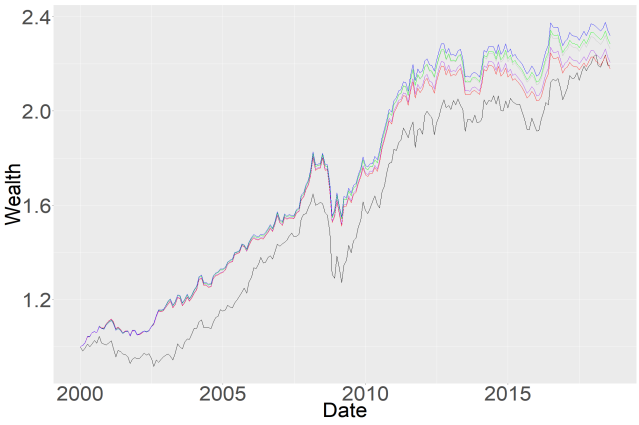
	FTSE-All-Share	S&P500 Comp.	BMUK10Y Index	BMUS10Y Index	TR-Equal-Weight-CCI	S&P-GSCI-Gold
Panel A: *Correlation matrix*						
FTSE All Share	1.00					
S&P500 Comp.	0.69102	1.00				
BMUK10Y Index	0.34133	0.17370	1.00			
BMUS10Y Index	0.05698	0.21983	0.34348	1.00		
TR Equal Weight CCI	0.12851	0.10135	0.00504	-0.19435	1.00	
S&P GSCI Gold	0.02652	0.02798	0.05116	-0.00343	0.45557	1.00
Panel B: *Covariance matrix w/ global s.d.*						
FTSE All Share	2.4110e-03					
S&P500 Comp.	1.4124e-03	1.7327e-03				
BMUK10Y Index	4.6879e-04	2.0224e-04	7.8234e-04			
BMUS10Y Index	7.4277e-05	2.4291e-04	2.5504e-04	7.0471e-04		
TR Equal Weight CCI	1.7146e-04	1.4535e-04	3.8341e-06	-1.4018e-04	7.3824e-04	
S&P GSCI Gold	7.0079e-05	6.2697e-05	7.7015e-05	-4.8946e-06	6.6623e-04	2.8969e-03
Panel C: *Covariance matrix w/ local s.d.*						
FTSE All Share	2.4110e-03					
S&P500 Comp.	1.3874e-03	1.7327e-03				
BMUK10Y Index	4.6429e-04	2.0086e-04	7.8234e-04			
BMUS10Y Index	7.3559e-05	2.4125e-04	2.5382e-04	7.0471e-04		
TR Equal Weight CCI	1.6985e-04	1.4435e-04	3.8161e-06	-1.3952e-04	7.3824e-04	
S&P GSCI Gold	6.9243e-05	6.2092e-05	7.6441e-05	-4.8575e-06	6.6047e-04	2.8969e-03

Table 8.13: Local correlation- and covariances estimated under bull market conditions ($M = 240$).

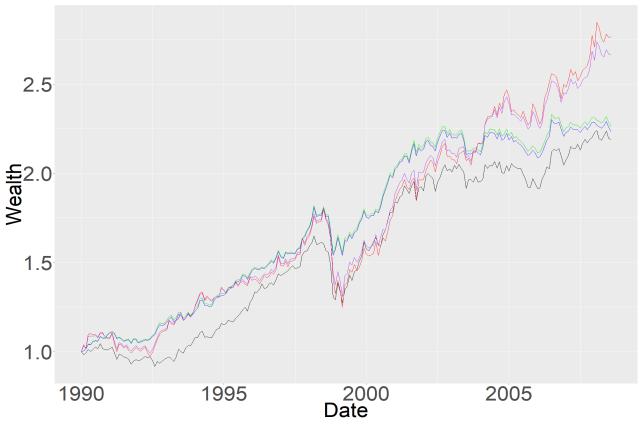
	FTSE-All-Share	S&P500 Comp.	BMUK10Y Index	BMUS10Y Index	TR-Equal-Weight-CCI	S&P-GSCI-Gold
Panel A: *Correlation matrix*						
FTSE All Share	1.00					
S&P500 Comp.	0.60948	1.00				
BMUK10Y Index	0.40467	0.22847	1.00			
BMUS10Y Index	0.10295	0.25165	0.37298	1.00		
TR Equal Weight CCI	0.10966	0.05461	-0.02173	-0.19694	1.00	
S&P GSCI Gold	0.08510	0.04334	0.05349	-0.02133	0.41413	1.00
Panel B: *Covariance matrix w/ global s.d.*						
FTSE All Share	2.41103e-03					
S&P500 Comp.	1.24575e-03	1.7327e-03				
BMUK10Y Index	5.5577e-04	2.6601e-04	7.8234e-04			
BMUS10Y Index	1.3420e-04	2.7809e-04	2.7694e-04	7.0471e-04		
TR Equal Weight CCI	1.4630e-04	1.2402e-04	-1.6518e-05	-1.4205e-04	7.3824e-04	
S&P GSCI Gold	2.2491e-04	9.7104e-05	8.0539e-05	-3.0472e-05	6.0563e-04	2.8969e-03
Panel C: *Covariance matrix w/ local s.d.*						
FTSE All Share	2.41103e-03					
S&P500 Comp.	9.5017e-04	1.7327e-03				
BMUK10Y Index	4.8239e-04	2.4656e-04	7.8234e-04			
BMUS10Y Index	1.2331e-04	2.6736e-04	2.8517e-04	7.0471e-04		
TR Equal Weight CCI	1.2587e-04	1.1105e-04	-1.5838e-05	-1.4336e-04	7.3824e-04	
S&P GSCI Gold	2.1819e-04	9.7178e-05	8.3900e-05	-3.2823e-05	6.2468e-04	2.8969e-03

The plots on left side of Figure 8.11 use global s.d.-, whereas those on right side use local s.d. to estimate covariances. Color code; - benchmark $1/N$-, - global-, - bear-, - normal-, - and bull-market conditions. The - 80-20% weighted portfolio are estimated from bear- and normal market conditions, respectively.

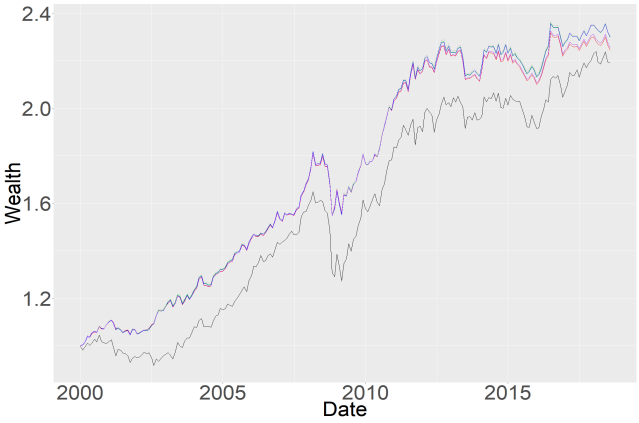
Figure 8.11: Buy-and-Hold portfolio performance using a sampling window of 240 trading months.



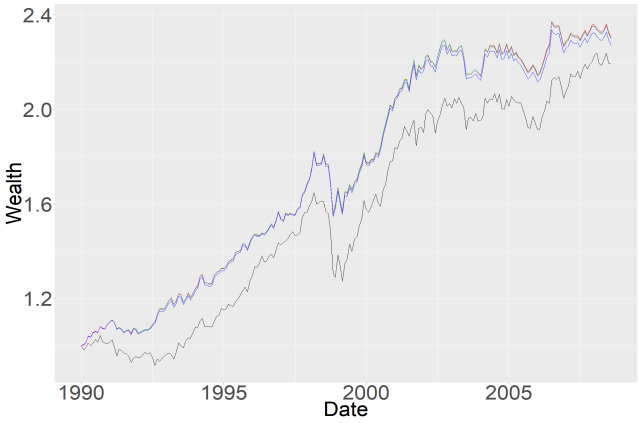
(a) Mean-Variance with short-selling (MVS)



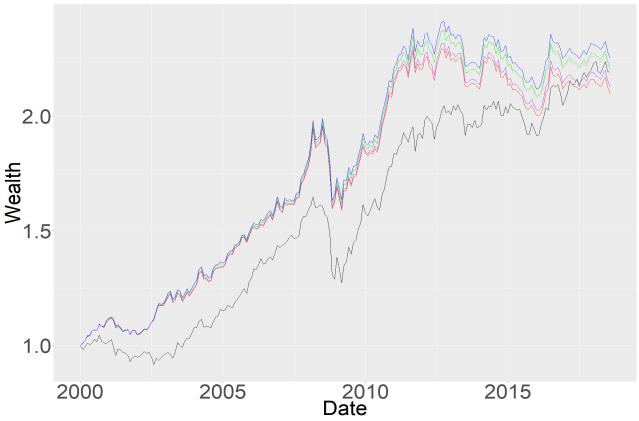
(b) Mean-Variance with short-selling (MVS)



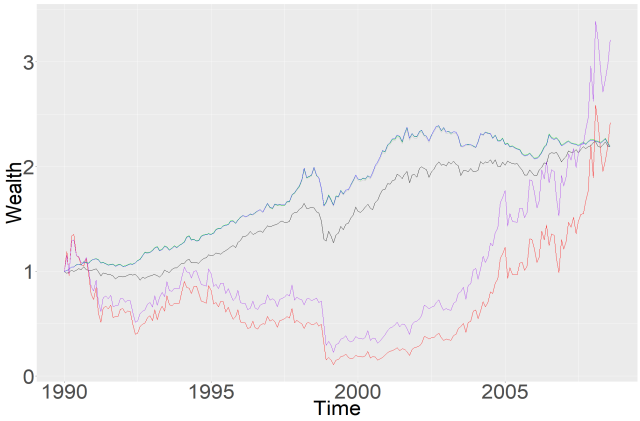
(c) Mean-Variance without short-selling (MVSC)



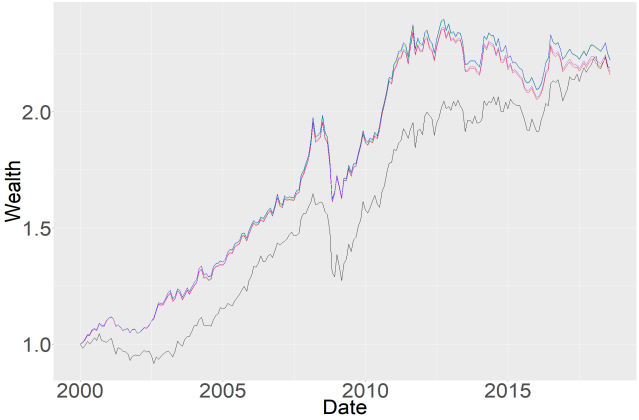
(d) Mean-Variance without short-selling (MVSC)



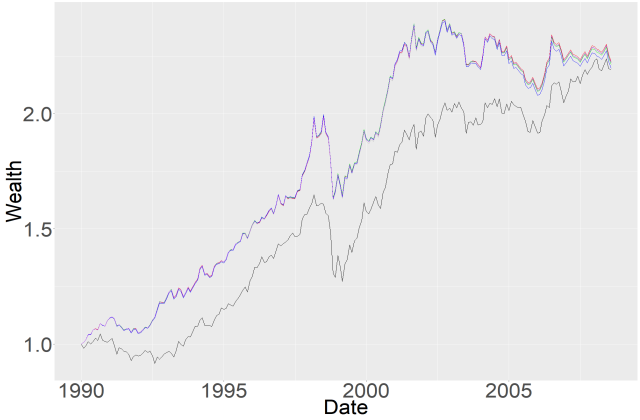
(e) Minimum-Variance without short-selling (MIN)



(f) Minimum-Variance without short-selling (MIN)



(g) Minimum-Variance without short-selling (MINC)



(h) Minimum-Variance without short-selling (MINC)

Furthermore, Table 8.14 and 8.15 reports the performance in the sense of Sharpe Ratio for the given portfolio from the plots in Figure 8.10, 8.11 and A.4, respectively. As can be seen in Panel B and D of Table 8.14, higher Sharpe Ratios are produced when we exclude short-selling constraints under bear market conditions. Notice that Sharpe ratios of portfolio returns increases as ρ decreases, especially when we use global standard deviation to calculate local covariances. These results are consistent when we apply for longer historical sampling window as well, see Table 8.15, however, lower Sharpe Ratios are produced when longer sampling windows are applied. This indicates that when ρ is small, the portfolios have a smaller risk given the expected return compared to when ρ is large. As discussed in section 5.1.4, the Sharpe Ratio is a measure of the incremental return for taking on added risk. Therefore, we want portfolios of low correlated assets, intuitively. Note that local MINC estimates generally performs quite bad in the sense of Sharpe Ratio.

Table 8.14: Sharpe Ratio of Buy-and-Hold strategy using a sampling window of 120 trading months.

Portfolio strategy	$(1/N)$	Markowitz	LGC_{bear}	LGC_{normal}	LGC_{bull}	LGC_{80-20}
Panel A: *Mean-Variance with short-selling (MVS)*						
global s.d.	0.21176	0.26155	0.24424	$0.26924^{\uparrow}$	$0.27586^{\uparrow}$	0.25082
local s.d.			$0.29644^{\uparrow}$	0.26078	0.25745	$0.29717^{\uparrow}$
5-par-mf			$0.29323^{\uparrow}$	0.26084	0.25771	$0.29133^{\uparrow}$
Panel B: *Mean-Variance without short-selling (MVSC)*						
global s.d.		0.26978	0.24802	$0.27821^{\uparrow}$	$0.28055^{\uparrow}$	0.25495
local s.d.			$0.28037^{\uparrow}$	0.26856	0.26377	$0.27721^{\uparrow}$
5-par-mf			0.25615	0.26864	0.26522	0.25810
Panel C: *Minimum-Variance with short-selling (MIN)*						
global s.d.		0.23003	0.22670	$0.23872^{\uparrow}$	$0.24152^{\uparrow}$	0.23001
local s.d.			$0.24202^{\uparrow}$	0.22976	0.22573	0.14449
5-par-mf			0.14885	0.22975	0.22562	0.15742
Panel D: *Minimum-Variance without short-selling (MINC)*						
global s.d.		0.23656	0.23565	$0.24422^{\uparrow}$	$0.23966^{\uparrow}$	$0.23687^{\uparrow}$
local s.d.			0.13864	0.23567	0.22775	$0.23920^{\uparrow}$
5-par-mf			0.24434	0.23571	0.22880	0.24127

Table 8.15: Sharpe Ratio of Buy-and-Hold strategy using a sampling window of 240 trading months.

Portfolio strategy	$(1/N)$	Markowitz	LGC_{bear}	LGC_{normal}	LGC_{bull}	LGC_{80-20}
Panel A: *Mean-Variance with short-selling (MVS)*						
global s.d.	0.17414	0.23471	0.22161	$0.23601^{\uparrow}$	$0.24006^{\uparrow}$	0.22555
local s.d.			0.18488	$0.23567^{\uparrow}$	0.23105	0.19772
5-par-mf			0.18249	$0.23574^{\uparrow}$	0.23201	0.19550
Panel B: *Mean-Variance without short-selling (MVSC)*						
global s.d.		0.24109	0.23470	0.23970	0.23891	0.23568
local s.d.			$0.24214^{\uparrow}$	$0.24158^{\uparrow}$	0.23678	$0.24110^{\uparrow}$
5-par-mf			$0.24137^{\uparrow}$	$0.24171^{\uparrow}$	0.23867	0.24085
Panel C: *Minimum-Variance with short-selling (MIN)*						
global s.d.		0.19945	0.18891	$0.20196^{\uparrow}$	$0.20543^{\uparrow}$	0.19248
local s.d.			0.096452	0.19930	0.19633	0.10327
5-par-mf			0.09356	0.19928	0.19681	0.10100
Panel D: *Minimum-Variance without short-selling (MINC)*						
global s.d.		0.20214	0.19886	$0.20366^{\uparrow}$	0.20186	0.19959
local s.d.			$0.20417^{\uparrow}$	0.20163	0.19731	$0.20279^{\uparrow}$
5-par-mf			$0.20464^{\uparrow}$	0.20166	0.19858	$0.20342^{\uparrow}$

Table 8.14 and 8.15 provides the Sharpe Ratio for various portfolio rules (see Table 7.1) applied to the portfolio data set when Markowitz or model-based estimates are applied for the Buy-and-Hold strategy. Panel A (B) reports the Mean-Variance portfolio performance with (without) long-only constraints. Panel C (D) reports the Minimum-Variance portfolio performance with (without) long-only constraints, all using sampling windows of 120 and 240 trading months.

8.2.2 Out-of-sample performance of a rolling window

Before we present the final results of our analysis, I will clarify that the following three sections are highly inspired by the papers; DeMiguel and Uppal [2009], Støve and Tjøstheim [2014] and Yew-Low et al. [2016].[2] Table 8.16, 8.17 and 8.18 reports the Sharpe Ratio and CEQ values for the dataset investigated across every optimization rule. Panel A in Table 8.17 reports the Sharpe Ratio and CEQ values for the benchmark 1/N model. Overall, the 1/N portfolio performs poorly in the sense of the Sharpe Ratio and CEQ. In Table 8.16, Panel B ($M = 120$) display results for alternative optimization strategies when we apply moving-grid under normal market conditions. Higher Sharpe Ratio and CEQ values are produced for 4 (4) optimization strategies when local estimates with global s.d. are applied as opposed to global covariance estimates, whereas 4 (4)

Table 8.16: Portfolio performance results across different optimization rules with moving-grid.

Portfolio Strategy		Sharpe Ratio (SR)			Certainty-equivalent (CEQ)	
Panel A: *benchmark model*						
$\text{EWR}_{M=120}$		0.16948			0.73655	
$\text{EWR}_{M=240}$		0.15823			0.47729	
Moving-grid	SR_M	$SR_{LGC}^{global-s.d.}$	$SR_{LGC}^{local-s.d.}$	CEQ_M	$CEQ_{LGC}^{global-s.d.}$	$CEQ_{LGC}^{local-s.d.}$
Panel B: *alternative portfolio rules, Window size M=120*						
Classic approach that ignores estimation error						
MVS	0.24257	$0.24913^{\uparrow}$	$0.24476^{\uparrow}$	0.94889	$0.97048^{\uparrow\#}$	$0.95746^{\uparrow\#}$
Moment restrictions						
MIN	0.25904	$0.26626^{\uparrow}$	$0.26190^{\uparrow}$	0.88889	$0.91311^{\uparrow\#}$	$0.90584^{\uparrow\#}$
Portfolio constraints						
MVSC	0.23705	$0.24110^{\uparrow}$	$0.23911^{\uparrow}$	0.90592	$0.93061^{\uparrow\#}$	$0.92274^{\uparrow\#}$
MINC	0.25535	$0.25893^{\uparrow}$	$0.25676^{\uparrow}$	0.85900	$0.86854^{\uparrow\#}$	$0.86669^{\uparrow\#}$
Panel C: *alternative portfolio rules, Window size M=240*						
Classic approach that ignores estimation error						
MVS	0.23145	$0.23607^{\uparrow}$	0.23109	0.67705	$0.69819^{\uparrow}$	$0.68355^{\uparrow}$
Moment restrictions						
MIN	0.25332	$0.26040^{\uparrow}$	$0.25468^{\uparrow}$	0.73848	$0.77272^{\uparrow}$	$0.75339^{\uparrow}$
Portfolio constraints						
MVSC	0.22810	$0.23242^{\uparrow*}$	$0.22957^{\uparrow}$	0.66111	$0.68074^{\uparrow}$	$0.67126^{\uparrow}$
MINC	0.25346	$0.26075^{\uparrow}$	$0.25494^{\uparrow}$	0.74296	$0.77464^{\uparrow}$	$0.75227^{\uparrow}$

Table 8.16 reports the SR and CEQ metric for various portfolio rules (see Table 7.1) applied to the portfolio data set when global (e.g., SR_M, CEQ_M) or local estimates (e.g., SR_{LGC}, CEQ_{LGC}) are applied. Panel A contains the benchmark models. Panel B and C report the out-of-sample portfolio performances for sampling windows of 120 and 240 months, respectively. The $^{\#}$ indicates higher CEQ when shorter historical sampling windows are applied.

of the optimization strategies perform better when local estimates with corresponding local s.d. are applied. In Panel C ($M = 240$), when longer sampling windows are applied, local portfolio estimates improves the Sharpe Ratio and CEQ values for 4 (4) when global s.d. are applied, and 4 (4) when local s.d. are applied, respectively. Further, we find that longer sampling window result in decreased portfolio performance, especially in the sense of CEQ, which we also did in the simulation study, respectively. Overall, we find that local portfolio estimates with moving-grid covariances as input generally performs quite well as opposed to global estimates that do not account for asymmetries. Thus, by accounting for asymmetric dependence within the portfolio data, and by using the moving-grid strategy, we have been able to advantage from the time-changing dependence structure within the market and manipulate the risk-and-return characteristics in our favor.

In Table 8.17 we are mainly interested in manipulating the risk-and-return characteristics by only relying on the covariances estimated at fixed grid-points on the left diagonal of the grid-map, that is, from bear- and normal market conditions. As mentioned earlier, our strategy is, by only relying on the higher correlation that is present in a falling market[3], then we as an investment manager, will be able to produce smoother and less volatile average portfolio excess returns, hopefully. Panel A in Table 8.17 display results for alternative optimization strategies when global s.d. are applied, and a sampling window length of 120 trading months, respectively. As reported, higher Sharpe Ratio and CEQ values are produced for 12 (12) of the optimization

[2]Other references of importance; Bessler and Wolff [2015] and Daskalaki and Skiadopoulos [2011].

[3]Higher correlation under bear market conditions are especially true for stock and commodity indices in time of financial crisis.

rules when local portfolio estimates are applied as opposed to global estimates. As for Panel B ($M = 120$), using the corresponding local s.d. to calculate covariances, higher Sharpe Ratio and CEQ values are produced for 1 (6) of the optimization rules when local portfolio estimates are applied as opposed to global estimates. Overall, local portofolio estimates using global s.d. generally produce higher Sharpe Ratio and CEQ values as opposed to when local s.d. are applied., indicating a slightly better performance. More specifically, local portfolio estimates with global s.d. produce higher Sharpe Ratio and CEQ values for 12 (9) as opposed to estimates with local s.d. In Panel C and D, where longer sampling windows of 240 trading months are used,

Table 8.17: Portfolio performance results across different optimization rules with fixed grid-points.

P.str.	SR_M	SR^{-4}_{LGC}	SR^{-2}_{LGC}	SR^{0}_{LGC}	CEQ_M	CEQ^{-4}_{LGC}	CEQ^{-2}_{LGC}	CEQ^{0}_{LGC}
Panel A: *alternative portfolio rules, Window size M=120, global s.d.*								
Classic approach that ignores estimation error								
MVS	0.24257	$0.25295^{\uparrow}$	$0.25097^{\uparrow}$	$0.24717^{\uparrow}$	0.94889	$1.00090^{\uparrow\#}$	$0.98623^{\uparrow\#}$	$0.95865^{\uparrow\#}$
Moment restrictions								
MIN	0.25904	$0.26519^{\uparrow}$	$0.26749^{\uparrow}$	$0.26325^{\uparrow}$	0.88889	$0.92334^{\uparrow\#}$	$0.92282^{\uparrow\#}$	$0.88723^{\#}$
Portfolio constraints								
MVSC	0.23705	$0.23908^{\uparrow}$	$0.23917^{\uparrow}$	$0.23897^{\uparrow}$	0.90592	$0.91480^{\uparrow\#}$	$0.91217^{\uparrow\#}$	$0.91268^{\uparrow\#}$
MINC	0.25535	$0.26280^{\uparrow}$	$0.26093^{\uparrow}$	$0.25765^{\uparrow}$	0.85900	$0.89324^{\uparrow\#}$	$0.88220^{\uparrow\#}$	$0.85943^{\uparrow\#}$
Panel B: *alternative portfolio rules, Window size M=120, local s.d.*								
Classic approach that ignores estimation error								
MVS	–	0.16804	$0.24271^{\uparrow}$	0.24162	–	$1.07850^{\uparrow\#}$	$0.95016^{\uparrow\#}$	$0.94162^{\#}$
Moment restrictions								
MIN	–	0.04824	0.25350	0.25138	–	-17.0830	$0.86221^{\#}$	$0.84506^{\#}$
Portfolio constraints								
MVSC	–	0.23460	0.23546	0.23562	–	$0.93731^{\uparrow\#}$	$0.91523^{\uparrow\#}$	$0.91216^{\uparrow\#}$
MINC	–	0.25022	0.25064	0.24874	–	$0.87302^{\uparrow\#}$	$0.84862^{\#}$	$0.83251^{\#}$
Panel C: *alternative portfolio rules, Window size M=240, global s.d.*								
Classic approach that ignores estimation error								
MVS	0.23145	$0.24417^{\uparrow}$	$0.24010^{\uparrow}$	$0.23690^{\uparrow}$	0.67705	$0.69641^{\uparrow}$	$0.69503^{\uparrow}$	$0.69254^{\uparrow}$
Moment restrictions								
MIN	0.25332	$0.26843^{\uparrow}$	$0.26359^{\uparrow}$	$0.25986^{\uparrow}$	0.73848	$0.76767^{\uparrow}$	$0.76414^{\uparrow}$	$0.76006^{\uparrow}$
Portfolio constraints								
MVSC	0.22810	$0.23200^{\uparrow}$	$0.23205^{\uparrow}$	$0.23137^{\uparrow}$	0.66111	0.65606	$0.66758^{\uparrow}$	$0.67220^{\uparrow}$
MINC	0.25346	$0.26371^{\uparrow}$	$0.26196^{\uparrow}$	$0.25960^{\uparrow}$	0.74296	$0.76148^{\uparrow}$	$0.76632^{\uparrow}$	$0.76580^{\uparrow}$
Panel D: *alternative portfolio rules, Window size M=240, local s.d.*								
Classic approach that ignores estimation error								
MVS	–	0.21953	$0.23866^{\uparrow}$	$0.23629^{\uparrow}$	–	$0.71216^{\uparrow}$	$0.70579^{\uparrow}$	$0.70288^{\uparrow}$
Moment restrictions								
MIN	–	0.10333	$0.26117^{\uparrow}$	$0.25948^{\uparrow}$	–	0.51592	$0.78648^{\uparrow}$	$0.78584^{\uparrow}$
Portfolio constraints								
MVSC	–	$0.23225^{\uparrow}$	$0.23698^{\uparrow}$	$0.23591^{\uparrow}$	–	$0.68643^{\uparrow}$	$0.70175^{\uparrow}$	$0.70390^{\uparrow}$
MINC	–	$0.25618^{\uparrow}$	$0.26101^{\uparrow}$	$0.25950^{\uparrow}$	–	0.78665	$0.78747^{\uparrow}$	$0.78672^{\uparrow}$

Table 8.17 reports the SR and CEQ metric for various portfolio rules (see Table 7.1) applied to the portfolio data set when global (e.g., SR_M, CEQ_M) or local estimates (e.g., SR_{LGC}, CEQ_{LGC}) are applied. Panel A and C report the out-of-sample portfolio performance for global s.d., while panel B and D for local s.d. using sampling windows of 120 and 240 months, respectively. The $^{\#}$ indicates higher CEQ when shorter historical sampling windows are applied.

applying local dependence estimates improves the Sharpe Ratio and CEQ values beyond the use of global portfolio estimates for 12 (12) strategies with global s.d., and for 9 (10) strategies with local s.d., respectively. Further analysis show that higher Sharpe Ratio and CEQ values are produced for 9 (1) of the portfolio optimization strategies when local estimates with global s.d. are applied as opposed to local estimates with local s.d., suggesting that the local covariances actually do quite well by high overall portfolio return but also high return volatility. Overall, model-based estimates using local correlation estimates and global s.d. do in general perform better in the sense of smoother and less risky estimates. Table 8.18 reports the weighted portfolio performance. More precisely, the target weights within our portfolio is a mixture of estimates from bear market-

(grid-point -4), and normal market condition (grid-point 0), that is, $w_k = q_1 \times w_i^{-4} + q_2 \times w_j^0$, where q_1 and q_2 are values of 20-, 50- and 80%. The proportions of q_1 and q_2 could of course be chosen in a more sophisticated way, but for now, they are fixed and chosen on the basis of risk-averse investment interests. Panel A ($M = 120$) display results for alternative optimization strategies with global s.d., respectively. A higher Sharpe Ratio and CEQ values are produced for 12 (12) of the optimization rules when local dependence estimates are applied as opposed to global estimates. As for Panel B ($M = 120$), higher Sharpe Ratio and CEQ values are produced for 1 (8) of the strategies when local covariance estimates are applied as opposed to global estimates. Overall,

Table 8.18: Weighted portfolio performance results across different optimization rules with fixed grid-points.

P.str.	SR_M	SR_{LGC}^{80-20}	SR_{LGC}^{50-50}	SR_{LGC}^{20-80}	CEQ_M	CEQ_{LGC}^{80-20}	CEQ_{LGC}^{50-50}	CEQ_{LGC}^{20-80}
Panel A: *alternative portfolio rules, Window size M=120, global s.d.*								
Classic approach that ignores estimation error								
MVS	0.24257	$0.25247^{\uparrow}$	$0.25111^{\uparrow}$	$0.24836^{\uparrow}$	0.94889	$0.99303^{\uparrow\#}$	$0.98064^{\uparrow\#}$	$0.97186^{\uparrow\#}$
Moment restrictions								
MIN	0.25904	$0.26606^{\uparrow}$	$0.26622^{\uparrow}$	$0.26493^{\uparrow}$	0.88889	$0.91692^{\uparrow\#}$	$0.90654^{\uparrow\#}$	$0.89524^{\uparrow\#}$
Portfolio constraints								
MVSC	0.23705	$0.23915^{\uparrow}$	$0.23917^{\uparrow}$	$0.23908^{\uparrow}$	0.90592	$0.91449^{\uparrow\#}$	$0.91392^{\uparrow\#}$	$0.91322^{\uparrow\#}$
MINC	0.25535	$0.26199^{\uparrow}$	$0.26058^{\uparrow}$	$0.25890^{\uparrow}$	0.85900	$0.88657^{\uparrow\#}$	$0.87647^{\uparrow\#}$	$0.86628^{\uparrow\#}$
Panel B: *alternative portfolio rules, Window size M=120, local s.d.*								
Classic approach that ignores estimation error								
MVS	–	0.18375	0.21563	$0.24784^{\uparrow}$	–	$1.10510^{\uparrow\#}$	$1.10180^{\uparrow\#}$	$1.03180^{\uparrow\#}$
Moment restrictions								
MIN	–	0.06448	0.099417	0.19041	–	-7.70390	-0.44202	$1.16860^{\uparrow\#}$
Portfolio constraints								
MVSC	–	0.23502	0.23545	0.23563	–	$0.93248^{\uparrow\#}$	$0.92504^{\uparrow\#}$	$0.91738^{\uparrow\#}$
MINC	–	0.25023	0.24996	0.24935	–	$0.86503^{\uparrow\#}$	0.85294	0.84072
Panel C: *alternative portfolio rules, Window size M=240, global s.d.*								
Classic approach that ignores estimation error								
MVS	0.23145	$0.24296^{\uparrow}$	$0.24091^{\uparrow}$	$0.23875^{\uparrow}$	0.67705	$0.69581^{\uparrow}$	$0.69475^{\uparrow}$	$0.69073^{\uparrow}$
Moment restrictions								
MIN	0.25332	$0.26700^{\uparrow}$	$0.26458^{\uparrow}$	$0.26185^{\uparrow}$	0.73848	$0.76633^{\uparrow}$	$0.76414^{\uparrow}$	$0.76176^{\uparrow}$
Portfolio constraints								
MVSC	0.22810	$0.23193^{\uparrow}$	$0.23177^{\uparrow}$	$0.23155^{\uparrow}$	0.66111	$0.66900^{\uparrow}$	$0.66417^{\uparrow}$	$0.65932^{\uparrow}$
MINC	0.25346	$0.26300^{\uparrow}$	$0.26182^{\uparrow}$	$0.26053^{\uparrow}$	0.74296	$0.76502^{\uparrow}$	$0.76376^{\uparrow}$	$0.76243^{\uparrow}$
Panel D: *alternative portfolio rules, Window size M=240, local s.d.*								
Classic approach that ignores estimation error								
MVS	–	0.22522	$0.23191^{\uparrow}$	$0.23569^{\uparrow}$	–	$0.71342^{\uparrow}$	$0.71239^{\uparrow}$	$0.70785^{\uparrow}$
Moment restrictions								
MIN	–	0.13125	0.18610	0.24603	–	0.67197	$0.80701^{\uparrow}$	$0.82754^{\uparrow}$
Portfolio constraints								
MVSC	–	$0.23312^{\uparrow}$	$0.23430^{\uparrow}$	$0.23532^{\uparrow}$	–	$0.68999^{\uparrow}$	$0.69527^{\uparrow}$	$0.70048^{\uparrow}$
MINC	–	$0.25707^{\uparrow}$	$0.25820^{\uparrow}$	$0.25907^{\uparrow}$	–	$0.78681^{\uparrow}$	$0.78691^{\uparrow}$	$0.78685^{\uparrow}$

Table 8.18 reports the SR and CEQ metric for various portfolio rules (see Table 7.1) applied to the portfolio data set when global (e.g., SR_M, CEQ_M) or local estimates (e.g., SR_{LGC}, CEQ_{LGC}) are applied. Panel A and C report the out-of-sample portfolio performance with global s.d., while panel B and D with local s.d. for sample windows of 120 and 240 months, respectively. The $^{\#}$ indicates higher CEQ when shorter historical sampling windows are applied.

optimization strategies using global s.d. generally produce higher Sharpe Ratio and CEQ values as opposed to when local s.d. are applied., indicating better performance. More specifically, local dependence estimates with global s.d. produces higher Sharpe Ratio and CEQ values for 12 (5) as opposed to local dependence estimates with local s.d. In Panel C and D ($M = 240$), the local dependence estimates improves the Sharpe Ratio and CEQ values beyond the use of global estimates for 12 (12) strategies using global s.d., whereas 8 (11) strategies using local s.d., respectively. Further analysis show that higher Sharpe Ratio and CEQ values are produced for 9 (1) strategies when local dependence estimates with global s.d. are applied as opposed to local local s.d., indicating that the local covariance estimates actually perform quite well, as concluded in Table 8.17.

Discussion

Across the portfolio strategies investigated, we find that by accounting for asymmetry within the return distribution generally results in higher Sharpe Ratio and CEQ values, although, there are some exceptions (for example, in Table 8.19: Panel A: MIN at grid point 0, Panel B: most except moving-grid estimates; Panel C: 80-20% weighted MVSC; Panel D: MVS and MIN strategies under bear market conditions. We also find that the portfolio performance do in general decrease when we apply longer historical sampling windows. Thus, our results contradicts to that of Tu and Zhou [2011] who report improved performance when longer historical sampling windows are applied[4]. Furthermore, we see from Panel A and C that the best performing strategies

Table 8.19: Economic value of asymmetry.

Portfolio Strategy	SR_{Δ}^{MVS}	CEQ_{Δ}^{MVS}	SR_{Δ}^{MVSC}	CEQ_{Δ}^{MVSC}	SR_{Δ}^{MIN}	CEQ_{Δ}^{MIN}	SR_{Δ}^{MINC}	CEQ_{Δ}^{MINC}
Panel A: *asymmetry, Window size M=120, global s.d.*								
Fixed grid-point								
LGC_{-4}	0.01038#	0.05201#	0.00203	0.00888	0.00615	0.03445#	0.00745#	0.03424#
LGC_{-2}	0.00840	0.03734	0.00212	0.00625	0.00845#	0.03393	0.00558	0.02320
LGC_{0}	0.00460	0.00976	0.00192	0.00676	0.00421	-0.00166	0.00230	0.00043
LGC_{80-20}	0.00990	0.04414	0.00210	0.00857	0.00702	0.02803	0.00664	0.02757
LGC_{50-50}	0.00854	0.03175	0.00212	0.00800	0.00718	0.01765	0.00523	0.01747
LGC_{20-80}	0.00579	0.02297	0.00203	0.00730	0.00589	0.00653	0.00355	0.00728
Moving-grid	0.00656	0.02159	0.00405#	0.02469#	0.00722	0.02422	0.00358	0.00954
Panel B: *asymmetry, Window size M=120, local s.d.*								
Fixed grid-point								
LGC_{-4}	-0.07453	0.12961	-0.00245	0.03139#	-0.21081	-17.9889	-0.00513	0.01402#
LGC_{-2}	0.00014	0.00127	-0.00159	0.00931	-0.00554	-0.02669	-0.00471	-0.01038
LGC_{0}	-0.00095	-0.00727	-0.00143	0.00624	-0.00766	-0.04383	-0.00661	-0.02649
LGC_{80-20}	-0.05882	0.15621#	-0.00203	0.02656	-0.19456	-8.59279	-0.00512	0.00603
LGC_{50-50}	-0.02694	0.15291	-0.00160	0.01912	-0.15962	-1.33091	-0.01990	-0.00606
LGC_{20-80}	0.00527	0.08291	-0.00142	-0.03151	-0.06863	0.27971#	-0.00600	-0.01828
Moving-grid	0.00219#	0.00857	0.00206#	0.01682	0.00286#	0.01695	0.00141#	0.00769
Panel C: *asymmetry, Window size M=240, global s.d.*								
Fixed grid-point								
LGC_{-4}	0.01272#	0.01936	0.00390	0.01852	0.01511#	0.02919	0.01025#	0.01852
LGC_{-2}	0.00865	0.01798	0.00395	0.00647	0.01027	0.02566	0.00850	0.02336
LGC_{0}	0.00545	0.01549	0.00327	0.01109	0.00654	0.02158	0.00614	0.02284
LGC_{80-20}	0.01151	0.01876	0.00383	-0.00179	0.01368	0.02785	0.00954	0.01947
LGC_{50-50}	0.00946	0.01770	0.00367	0.00306	0.01126	0.02566	0.00836	0.02080
LGC_{20-80}	0.00730	0.01368	0.00345	0.00789	0.00853	0.02328	0.00707	0.02206
Moving-grid	0.00462	0.02114#	0.00432#	0.01963#	0.00708	0.03424#	0.00729	0.03168#
Panel D: *asymmetry, Window size M=240, local s.d.*								
Fixed grid-point								
LGC_{-4}	-0.01192	0.03511	0.00415	0.02532	-0.14999	-0.22256	0.00272	0.04369
LGC_{-2}	0.00721#	0.02874	0.00888#	0.04064	0.00785#	0.04800	0.00755#	0.04451#
LGC_{0}	0.00484	0.02583	0.00781	0.04279#	0.00616	0.04736	0.00604	0.04376
LGC_{80-20}	-0.00623	0.03637#	0.00502	0.02888	-0.12207	-0.06651	0.00361	0.04385
LGC_{50-50}	0.00046	0.03534	0.00620	0.03416	-0.06722	0.06853	0.00474	0.04395
LGC_{20-80}	0.00424	0.03080	0.00722	0.03937	-0.00729	0.08906#	0.00561	0.04389
Moving-grid	-0.00036	0.00650	0.00147	0.01015	0.00136	0.01491	0.00148	0.00931

Table 8.19 reports the performance differences for each pair of confronted models in the sense of SR (SR_{Δ}) and CEQ (CEQ_{Δ}) between the competing strategies. The Δ columns indicates the performance differences between the portfolios; local dependence estimates minus global estimates - with a positive value indicating an improvement (i.e., increase) in the performance for the local dependence strategy. Similarly, negative values indicates that incorporating asymmetry in the correlation fails to enhance economic value. The # indicates the best performing model-based portfolio strategies.

[4]Tu and Zhou [2011] report increased reliability of the expected return and covariances as the sampling window size increases.

are variants of the moving-grid portfolio, and local estimates under bear market conditions, more precisely at grid point -4. The moving-grid portfolio is superior for the MVSC strategy using a sampling window of $M = 120$ and 240 trading months, and for the MVS, MIN and MINC strategy in the sense of CEQ, when we apply a rolling sampling window of 240 trading months, respectively. The local portfolio estimates from grid-point -4 using MVS and MINC are superior for a sampling window of 120 trading months, and the MIN strategy in the sense of CEQ. For a longer sampling window, the MVS, MIN and MINC strategy performs best in the sense of Sharpe Ratio. In Panel B ($M = 120$) however, we find that by accounting for asymmetric dependence and using local s.d., do in fact improve portfolio performance for some optimization strategies in the sense of CEQ, but in general produce negative Sharpe Ratio values. From the results in Panel D ($M = 240$), we can tell that the local estimates provides more stable portfolio estimates, and that it actually is outperforming the benchmarks for some strategies. Note that moving-grid estimates perform best in the sense of Sharpe Ratio, and variants of fixed-grid estimates outperforms in the sense of CEQ. Summarizing Table 8.19, the best performing strategies across all panels are variants of the moving-grid portfolios, and portfolio strategies (and weighted portfolios) relying on fixed grid-point estimates under bear market conditions, more specifically at grid-point -4.

8.2.3 Coefficients of the optimization rules

The analysis in Section 8.2.2 indicates that the best performing strategies are often variants of the MIN and MINC strategies. In Table 8.20, 8.21 and 8.22 we report the average portfolio return- and the standard deviations of the portfolio's excess returns, and as can be seen from eq.(5.27) and eq.(5.28), when the average portfolio

Table 8.20: Coefficients applied to the different optimization strategies investigated in Table 8.16

Portfolio Strategy		$\mu_{1/N}$			$\sigma_{1/N}$	
Panel A: *Benchmark model*						
$\text{EWR}_{M=120}$		0.86756			0.051189	
$\text{EWR}_{M=240}$		0.53429			0.033766	
Moving-grid	μ_M	$\mu_{LGC}^{global-s.d.}$	$\mu_{LGC}^{local-s.d.}$	σ_M	$\sigma_{LGC}^{global-s.d.}$	$\sigma_{LGC}^{local-s.d.}$
Panel B: *alternative portfolio rules, Window size M=120*						
Classic approach that ignores estimation error						
MVS	1.04100	$1.0612^{\uparrow\#}$	$1.04940^{\uparrow\#}$	0.042915	$0.042597^{\downarrow}$	$0.042874^{\downarrow}$
Moment restrictions						
MIN	0.95715	$0.98097^{\uparrow\#}$	$0.97516^{\uparrow\#}$	0.036951	$0.036842^{\downarrow}$	0.037235
Portfolio constraints						
MVSC	0.99380	$1.0201^{\uparrow\#}$	$1.01240^{\uparrow\#}$	0.041923	0.042311	0.042339
MINC	0.92454	$0.93353^{\uparrow\#}$	$0.93267^{\uparrow\#}$	0.036207	$0.036053^{\downarrow}$	0.036325
Panel C: *alternative portfolio rules, Window size M=240*						
Classic approach that ignores estimation error						
MVS	0.72629	$0.74845^{\uparrow\#}$	$0.73399^{\uparrow\#}$	0.031380	0.031705	0.031762
Moment restrictions						
MIN	0.78670	$0.82262^{\uparrow\#}$	$0.80311^{\uparrow\#}$	0.031056	0.031590	0.031534
Portfolio constraints						
MVSC	0.70948	$0.73007^{\uparrow\#}$	$0.72052^{\uparrow\#}$	0.031103	0.031412	0.031385
MINC	0.79175	$0.82465^{\uparrow\#}$	$0.80171^{\uparrow\#}$	0.031237	0.031625	0.031447

Table 8.20 reports the volatility- and average portfolio's excess return when global (μ_M, σ_M) or local (μ_{LGC}, σ_{LGC}) estimates are applied. The average returns are presented as percentage points. Panel A contains the benchmarks, Panel B and C reports the global and local estimates using rolling sampling windows of 120 and 240 months, respectively. The $^{\#}$ indicates a higher CEQ value when local dependence estimates are applied as opposed global estimates.

return increases, or, as the standard deviation of the portfolio's excess return decreases, the portfolio performance will increase. To assess whether the use of local dependence estimates are able to simultaneously improve portfolio performance and improve both coefficients, we identify cases where the local portfolio performance is superior to that based on global estimates by "#". Thus, in Table 8.20, Panel B ($M = 120$), moving-grid estimates produce higher average portfolio excess return for 4 (4) strategies, whereas only 3 (1) strategies report lower return volatility. Accordingly, this leads to increased performance across the portfolio strategies, as can be seen in the previous section. As for Panel C ($M = 240$), the strategies exhibit higher average portfolio excess return for 4 (4) strategies, whereas 0 (0) report lower volatility. However, these strategies are improved in terms of their CEQ values across all strategies. Therefore, application of local dependence estimates using

moving-grid around normal market conditions, do in fact show performance improvements that are supported by the Sharpe Ratio and CEQ values, for both sampling windows. Also note that the benchmark $1/N$ portfolio is outperformed by every optimization strategy, by far. Summarizing Table 8.20, the moving-grid strategy do in general produce portfolios with higher average excess return, but again, also higher return volatility.

In Table 8.21 across panel A and C, local portfolio estimates do increases the average portfolio excess return for 10 models (except MIN (MVSC) at grid-point -4 (0) for $M = 120$ (240)) and lower s.d. for 5 models. In Panel B ($M = 120$) however, when local covariances are applied, 7 strategies exhibit higher average portfolio

Table 8.21: Coefficients applied to the different optimization strategies investigated in Table 8.17

P.str.	μ_M	μ_{LGC}^{-4}	μ_{LGC}^{-2}	μ_{LGC}^{0}	σ_M	σ_{LGC}^{-4}	σ_{LGC}^{-2}	σ_{LGC}^{0}
Panel A: *alternative portfolio rules, Window size M=120, global s.d.*								
Classic approach that ignores estimation error								
MVS	1.04100	$1.04870^{\uparrow\#}$	$1.07860^{\uparrow\#}$	$1.09460^{\uparrow\#}$	0.042915	$0.042427^{\downarrow}$	0.042977	0.043272
Moment restrictions								
MIN	0.95715	0.95271	$0.99152^{\uparrow\#}$	$0.99352^{\uparrow\#}$	0.036951	$0.036190^{\downarrow}$	0.037067	0.037464
Portfolio constraints								
MVSC	0.99380	$1.00030^{\uparrow\#}$	$0.99948^{\uparrow\#}$	$1.00280^{\uparrow\#}$	0.041923	$0.041859^{\downarrow}$	$0.041789^{\downarrow}$	0.041942
MINC	0.92454	$0.92370^{\#}$	$0.94823^{\uparrow\#}$	$0.95995^{\uparrow\#}$	0.036207	$0.035851^{\downarrow}$	0.036341	0.036528
Panel B: *alternative portfolio rules, Window size M=120, local s.d.*								
Classic approach that ignores estimation error								
MVS	–	$1.45170^{\uparrow\#}$	$1.04240^{\uparrow\#}$	1.03300	–	0.086386	0.042948	$0.042753^{\downarrow}$
Moment restrictions								
MIN	–	$3.06170^{\uparrow}$	0.92942	0.91069	–	0.63474	$0.036664^{\downarrow}$	$0.036228^{\downarrow}$
Portfolio constraints								
MVSC	–	$1.03450^{\uparrow\#}$	$1.00660^{\uparrow\#}$	$1.00270^{\uparrow\#}$	–	0.044098	0.042750	0.042557
MINC	–	$0.94422^{\uparrow\#}$	0.91530	0.89762	–	0.037736	0.036519	$0.036086^{\downarrow}$
Panel C: *alternative portfolio rules, Window size M=240, global s.d.*								
Classic approach that ignores estimation error								
MVS	0.72629	$0.74153^{\uparrow\#}$	$0.74290^{\uparrow\#}$	$0.74267^{\uparrow\#}$	0.031380	$0.031302^{\downarrow}$	$0.030941^{\downarrow}$	$0.030416^{\downarrow}$
Moment restrictions								
MIN	0.78670	$0.80846^{\uparrow\#}$	$0.81153^{\uparrow\#}$	$0.81361^{\uparrow\#}$	0.031056	0.031112	$0.030788^{\downarrow}$	$0.030310^{\downarrow}$
Portfolio constraints								
MVSC	0.70948	$0.72072^{\uparrow\#}$	$0.71506^{\uparrow\#}$	0.70182	0.031103	0.031150	$0.030814^{\downarrow}$	$0.030252^{\downarrow}$
MINC	0.79175	$0.81509^{\uparrow\#}$	$0.81468^{\uparrow\#}$	$0.80848^{\uparrow\#}$	0.031237	0.031398	$0.031099^{\downarrow}$	$0.030658^{\downarrow}$
Panel D: *alternative portfolio rules, Window size M=240, local s.d.*								
Classic approach that ignores estimation error								
MVS	–	$0.77220^{\uparrow\#}$	$0.76710^{\uparrow\#}$	$0.75982^{\uparrow\#}$	–	0.034286	0.033078	0.032237
Moment restrictions								
MIN	–	$0.91494^{\uparrow}$	$0.93257^{\uparrow\#}$	$0.89348^{\uparrow\#}$	–	0.069709	0.050111	0.036316
Portfolio constraints								
MVSC	–	$0.74043^{\uparrow\#}$	$0.74596^{\uparrow\#}$	$0.75147^{\uparrow\#}$	–	0.031762	0.031838	0.031934
MINC	–	$0.84022^{\uparrow}$	$0.83981^{\uparrow\#}$	$0.83934^{\uparrow\#}$	–	0.032684	0.032526	0.032398

Table 8.21 reports the volatility- and return of the portfolio's when global (μ_M, σ_M) or local (μ_{LGC}, σ_{LGC}) estimates are applied. The average returns are presented as percentage points. Panel A (C) and Panel B (D) contains the global and local portfolios for sample windows of 120 (240) trading months, respectively. The $^{\#}$ indicates a higher CEQ value when local dependence estimates are applied as opposed global estimates.

excess return, and for 4 strategies, lower standard deviation. In Panel D ($M = 240$), higher average returns are reported for 12 strategies, whereas 0 strategies report lower volatility. Thus, the increase in average portfolio excess return across the panels are supported by improved CEQ values. Summarizing Table 8.21, we do get more reliable and informative estimates of the covariance matrix by accounting for the asymmetric dependence within the market, especially under bear market conditions, and thus improve the portfolio performances across the optimization rules investigated by either increasing the average return or decreasing the return volatility. This effect is thus less pronounced for longer sampling windows. Furthermore, Panel A and B in Table 8.22 reports improved average return for 12 (10) strategies, whereas only 9 (3) have improved return volatility. Panel

C and D however, reports 11 (9) increased average returns, and 12 (3) reduced return volatilities.

Cumulative Performance and Portfolio target weights of re-balanced portfolios

As discussed in Table 8.19, the best performing portfolios are often variants of the moving-grid strategy, and strategies relying on fixed-grid estimates under bear market conditions. Hence, we compare the best performing strategies by plotting the portfolio's cumulative excess returns for each strategy, see Figure 8.12 and 8.13.

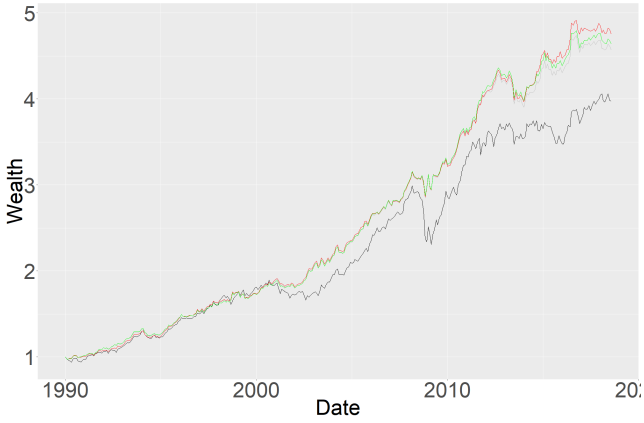
Table 8.22: Coefficients applied to the different optimization strategies investigated in Table 8.18

P.str.	μ_M	μ_{LGC}^{80-20}	μ_{LGC}^{50-50}	μ_{LGC}^{20-80}	σ_M	σ_{LGC}^{80-20}	σ_{LGC}^{50-50}	σ_{LGC}^{20-80}
Panel A: *alternative portfolio rules, Window size M=120, global s.d.*								
Classic approach that ignores estimation error								
MVS	1.04100	$1.08540^{\uparrow\#}$	$1.07170^{\uparrow\#}$	$1.06350^{\uparrow\#}$	0.042915	0.042993	$0.042679^{\downarrow}$	$0.042823^{\downarrow}$
Moment restrictions								
MIN	0.95715	$0.98553^{\uparrow\#}$	$0.97338^{\uparrow\#}$	$0.96104^{\uparrow\#}$	0.036951	0.037041	$0.036563^{\downarrow}$	$0.036276^{\downarrow}$
Portfolio constraints								
MVSC	0.99380	$1.00230^{\uparrow\#}$	$1.00160^{\uparrow\#}$	$1.00080^{\uparrow\#}$	0.041923	$0.041912^{\downarrow}$	$0.041879^{\downarrow}$	$0.041862^{\downarrow}$
MINC	0.92454	$0.95268^{\uparrow\#}$	$0.94179^{\uparrow\#}$	$0.93093^{\uparrow\#}$	0.036207	0.036363	$0.036143^{\downarrow}$	$0.035956^{\downarrow}$
Panel B: *alternative portfolio rules, Window size M=120, local s.d.*								
Classic approach that ignores estimation error								
MVS	–	$1.39220^{\uparrow\#}$	$1.27730^{\uparrow\#}$	$1.1371^{\uparrow\#}$	–	$0.042804^{\downarrow}$	0.043297	0.04588
Moment restrictions								
MIN	–	$2.98070^{\uparrow}$	$2.34880^{\uparrow}$	$1.46430^{\uparrow\#}$	–	$0.036530^{\downarrow}$	0.037390	0.038540
Portfolio constraints								
MVSC	–	$1.02820^{\uparrow\#}$	$1.0091^{\uparrow\#}$	$1.01860^{\uparrow\#}$	–	0.042146	0.042697	0.042824
MINC	–	$0.93481^{\uparrow\#}$	0.92079	0.90685	–	$0.036032^{\downarrow}$	0.036369	0.036763
Panel C: *alternative portfolio rules, Window size M=240, global s.d.*								
Classic approach that ignores estimation error								
MVS	0.72629	$0.73858^{\uparrow\#}$	$0.74221^{\uparrow\#}$	$0.74251^{\uparrow\#}$	0.031380	$0.030935^{\downarrow}$	$0.030809^{\downarrow}$	$0.030561^{\downarrow}$
Moment restrictions								
MIN	0.78670	$0.80956^{\uparrow\#}$	$0.81114^{\uparrow\#}$	$0.81264^{\uparrow\#}$	0.031056	$0.030917^{\downarrow}$	$0.030657^{\downarrow}$	$0.030436^{\downarrow}$
Portfolio constraints								
MVSC	0.70948	$0.71693^{\uparrow\#}$	$0.71126^{\uparrow\#}$	$0.70560^{\#}$	0.031103	$0.030962^{\downarrow}$	$0.030688^{\downarrow}$	$0.030423^{\downarrow}$
MINC	0.79175	$0.81380^{\uparrow\#}$	$0.81184^{\uparrow\#}$	$0.80983^{\uparrow\#}$	0.031237	$0.031236^{\downarrow}$	$0.031007^{\downarrow}$	$0.030793^{\downarrow}$
Panel D: *alternative portfolio rules, Window size M=240, local s.d.*								
Classic approach that ignores estimation error								
MVS	–	$0.71450^{\#}$	$0.70872^{\#}$	$0.70872^{\#}$	–	$0.030428^{\downarrow}$	$0.030815^{\downarrow}$	$0.030815^{\downarrow}$
Moment restrictions								
MIN	–	$0.81351^{\uparrow}$	$0.82171^{\uparrow\#}$	$0.82965^{\uparrow\#}$	–	0.031354	0.031852	0.032479
Portfolio constraints								
MVSC	–	$0.72268^{\uparrow\#}$	$0.72613^{\uparrow\#}$	$0.72613^{\uparrow\#}$	–	0.031168	0.031259	0.031259
MINC	–	$0.81896^{\uparrow\#}$	$0.82553^{\uparrow\#}$	$0.83196^{\uparrow\#}$	–	0.031549	0.031882	0.032308

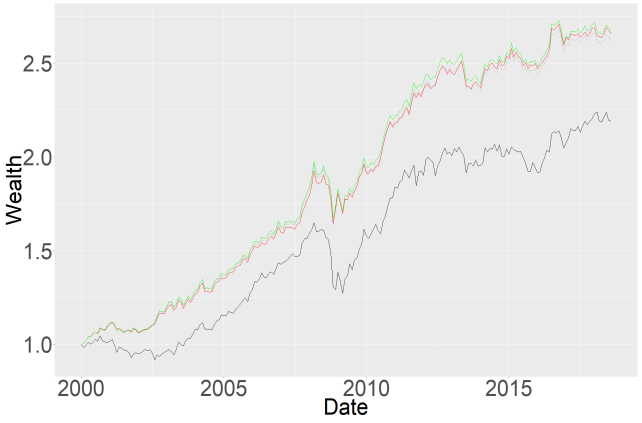
Table 8.22 reports the volatility- and return of the portfolio's when global (μ_M, σ_M) or local (μ_{LGC}, σ_{LGC}) estimates are applied. The average returns are presented as percentage points. Panel A (C) and Panel B (D) contains the global and local portfolios for sample windows of 120 (240) trading months, respectively. The $^{\#}$ indicates a higher CEQ value when local dependence estimates are applied as opposed global estimates.

Summarizing Figure 8.12 and 8.13, we do see that moving-grid estimates outperforms both the fixed-grid estimates and the benchmark models across all optimization rules, when we apply a rolling sampling window of 240 trading months. For a shorter sampling window of 120 trading months, the moving-grid estimates are still superior compared to the benchmark, but the fixed grid estimates under bear market conditions do perform better for some strategies. Further, superior performance for MVS and MIN strategy in Figure 8.13. Furthermore, we also give a visual representation of the corresponding rolling portfolio target weights by ploting the moving-grid- and fixed grid target estimates under bear market conditions, more specifically, at grid-point -4, both using global s.d. in the covariances. For more information, see Figure 8.14 and 8.15.

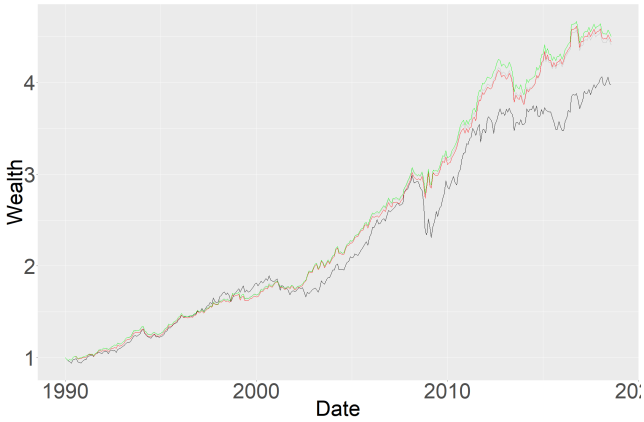
Figure 8.12: Cumulative performance using rolling sampling windows of 120 and 240 trading months. Color code; - benchmark $1/N$-, - global-, - bear market-, - moving-grid portfolio estimates.



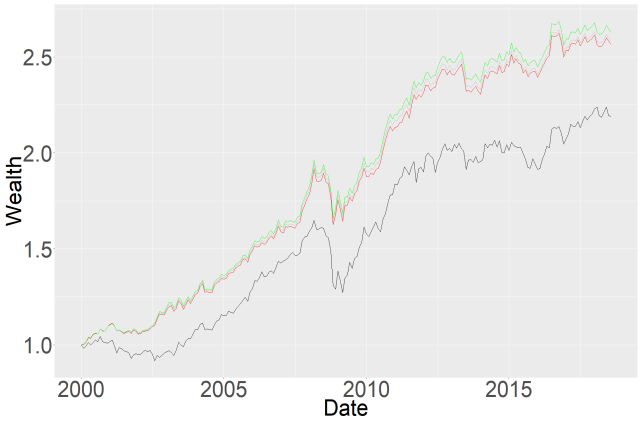
(a) Mean-Variance with short-selling (MVS)



(b) Mean-Variance with short-selling (MVS)

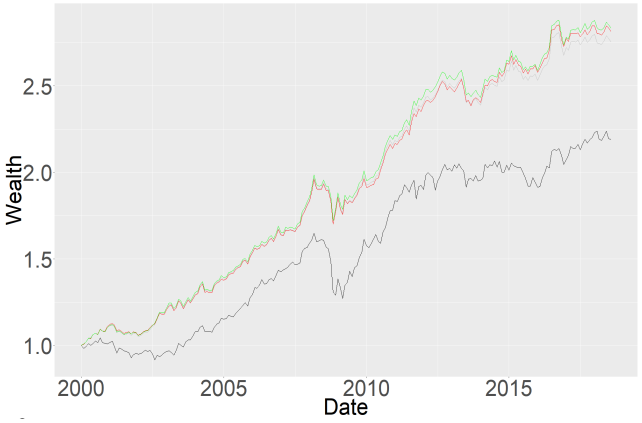


(c) Mean-Variance without short-selling (MVSC)

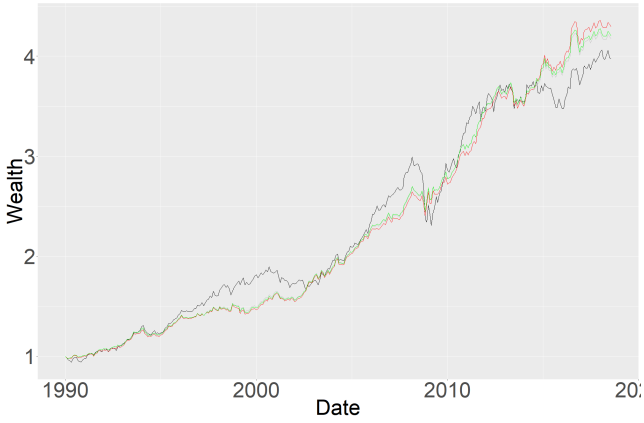


(d) Mean-Variance without short-selling (MVSC)

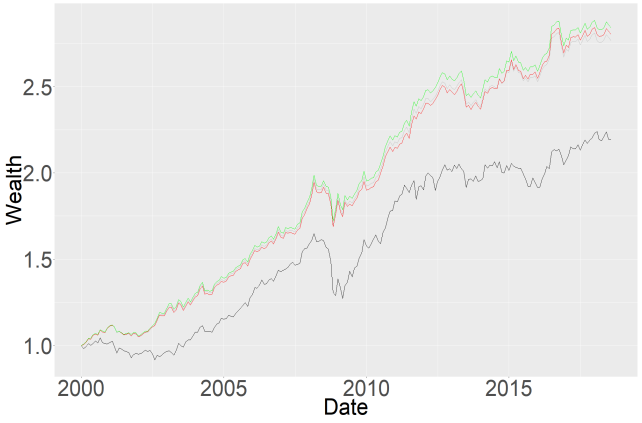
(e) Minimum-Variance with short-selling (MIN)



(f) Minimum-Variance with short-selling (MIN)

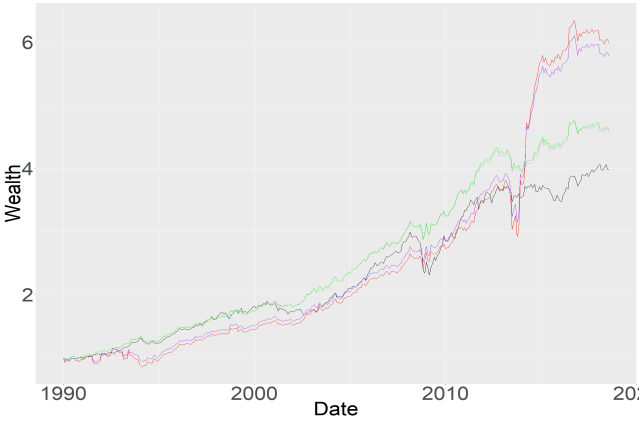


(g) Minimum-Variance without short-selling (MINC)

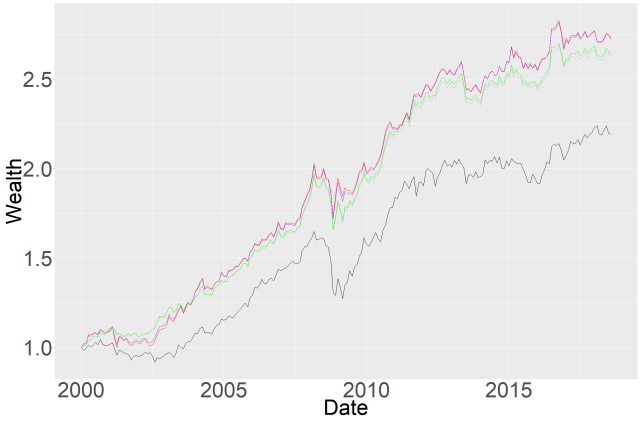


(h) Minimum-Variance without short-selling (MINC)

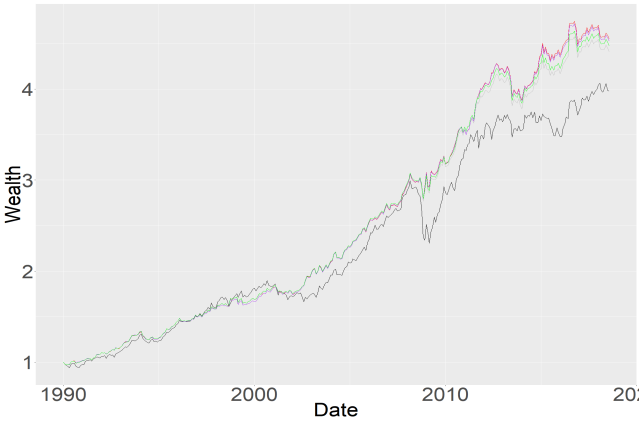
Figure 8.13: Cumulative performance using rolling sampling windows of 120 and 240 trading months. Color code; - benchmark $1/N$-, - global-, - bear market-, - 80-20% weighted-, and - moving-grid portfolio estimates.



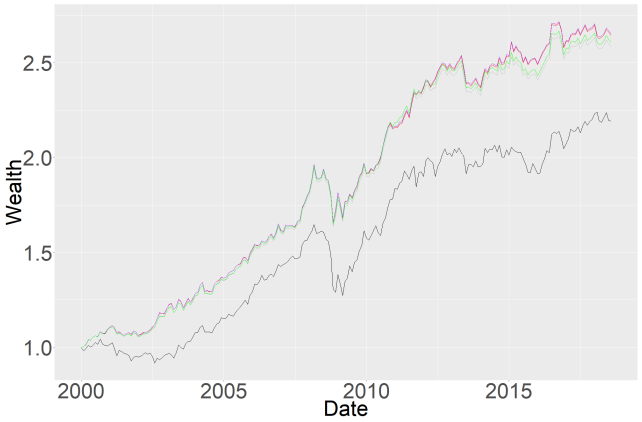
(a) Mean-Variance with short-selling (MVS)



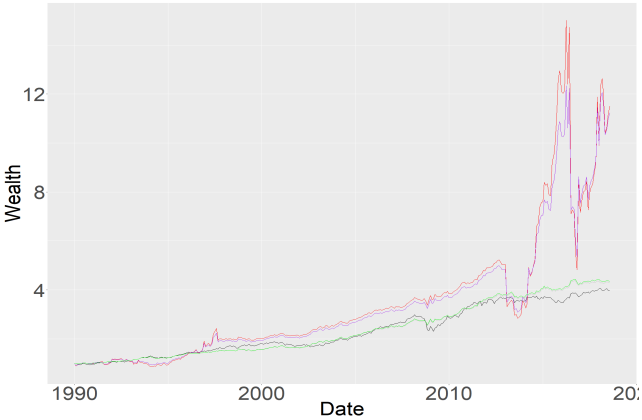
(b) Mean-Variance with short-selling (MVS)



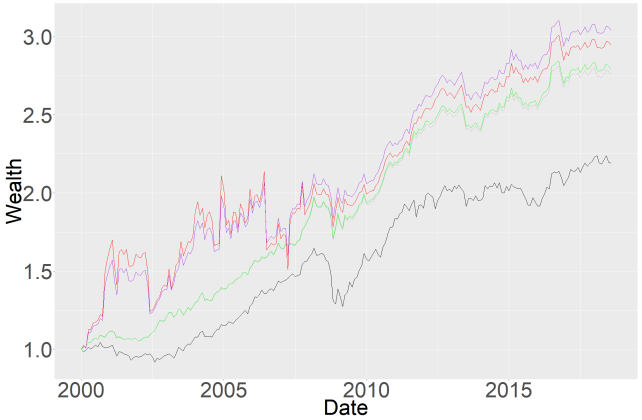
(c) Mean-Variance without short-selling (MVSC)



(d) Mean-Variance without short-selling (MVSC)

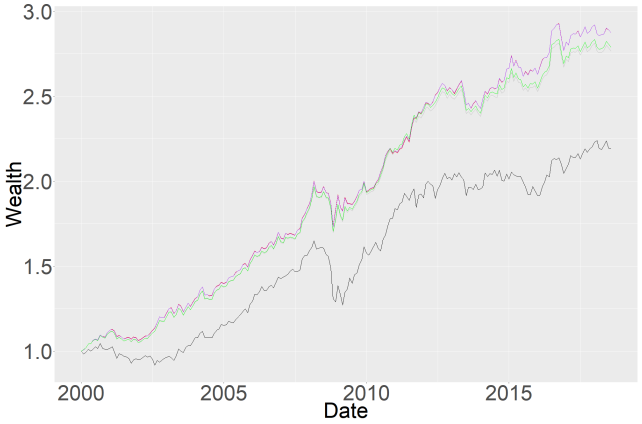


(e) Minimum-Variance with short-selling (MIN)



(f) Minimum-Variance with short-selling (MIN)

(g) Minimum-Variance without short-selling (MINC)



(h) Minimum-Variance without short-selling (MINC)

8.2.4 Portfolio re-balancing analysis

The primary goal of a re-balancing strategy is to minimize risk relative to a target asset allocation, rather than to maximize returns. Yet over time, the asset classes will produce different returns, and the portfolio's asset allocation changes. Then the portfolio will likely drift away from its primary target asset allocation, acquiring a risk-and-return characteristics that may be inconsistent with an investor's goals and preferences. Therefore, to re-capture the portfolio's original risk-and-return characteristics, the portfolio's should be re-balanced. As part of the portfolio-construction process, an optimal investment strategy involves two key issues: (1) frequency of re-balancing; and (2) re-balancing fully towards the target allocation or an intermediate allocation. The only clear advantage for any of these strategies, as far as maintaining a portfolio's risk-and-return characteristics, and without factoring in re-balancing costs, is that a re-balanced portfolio more closely aligns with the characteristics of the target asset allocation than a portfolio that is never re-balanced. If one is unable to re-balance fully towards the target portfolio weights as required by the portfolio strategy, this results in sub-optimal diversification. Therefore, all other things equal, a strategy that leads to greater stability in target portfolio weights is desirable as it is easier for a practitioner to implement (DeMiguel and Uppal [2009]). As such, in addition to risk-adjusted performance, assessing the average standard deviation in target weights is a criteria that a practitioner uses in the selection of a portfolio strategy.

Table 8.23: Average volatility of portfolio target weights across optimization strategies using moving-grid.

Portfolio Strategy	σ_M	$\sigma_{LGC}^{global-s.d.}$	Δ	$\sigma_{LGC}^{local-s.d.}$	Δ
Panel A: alternative portfolio rules, Window size M=120					
Classic approach that ignores estimation error					
MVS	0.85752	$0.85682^{\#}$	0.00070	$0.85587^{\#}$	0.00165
Moment restrictions					
MIN	0.85328	$0.85221^{\#}$	0.00107	$0.85156^{\#}$	0.00172
Portfolio constraints					
MVSC	0.85094	$0.85042^{\#}$	0.00052	$0.85004^{\#}$	0.00090
MINC	0.85176	$0.85136^{\#}$	0.00040	$0.85064^{\#}$	0.00112
Panel B: alternative portfolio rules, Window size M=240					
Classic approach that ignores estimation error					
MVS	0.84736	$0.84632^{\#}$	0.00104	$0.84639^{\#}$	0.00097
Moment restrictions					
MIN	0.84663	$0.84584^{\#}$	0.00079	$0.84565^{\#}$	0.00098
Portfolio constraints					
MVSC	0.84596	$0.84546^{\#}$	0.00050	$0.84536^{\#}$	0.00060
MINC	0.84614	$0.84570^{\#}$	0.00044	$0.84543^{\#}$	0.00071

Table 8.23 reports the average standard deviation in target weights when global (σ_M) or local (σ_{LGC}) estimates are used for the data set investigated. Panel A and B report the average volatility within the portfolio target weights across all strategies using different sampling windows of 120 and 240 months, respectively. $^{\#}$ indicates higher CEQ value when local estimates are applied as opposed to global estimates.

Accordingly, Table 8.23, 8.24 and 8.25 shows an analysis of the variability within the target portfolio weights across each of our previous settings. The average standard deviation within target portfolio weights across the entire out-of-sample time period is calculated as follows:

$$\bar{\sigma}(\hat{\mathbf{w}}_{k,c,M}) = \frac{\sum_{t=1}^{T-M} \sigma(\hat{w}_{k,t,c,M})}{T-M}, \tag{8.1}$$

where

$$\sigma(\hat{\mathbf{w}}_{k,c,M}) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{w}_{k,t,c,M,i} - \bar{\hat{w}}_{k,t,c,M}\right)^2} \tag{8.2}$$

measures a sort of Euclidean distance to compare the differences in allocation between the assets within the portfolio. Furthermore, the $\hat{w}_{k,t,c,M}$ is the N vector of target portfolio weights at time t under strategy k using estimates of expected returns from the dataset based upon a sampling window size of M trading months. Similarly, $\hat{w}_{k,t,c,M,i}$ is the target portfolio weight for asset i in a portfolio of N assets, and $\bar{\hat{w}}_{k,t,c,M}$ is the average target portfolio weight across the portfolio of N assets. The Δ column reports the differences between

the average standard deviation of the portfolio weights of the different strategies; global dependence estimates minus local estimates - with a positive value indicating an improvement (i.e., reduction) when local dependence estimates are applied, as opposed to global estimates. To assess whether the use of local dependence estimates are able to simultaneously improve portfolio performance and reduce the average standard deviation of portfolio weights, we identify cases where the portfolio strategy produces superior portfolio performance compared to the global estimates by $^{\#}$. From Panel A (Table 8.23), we find that all portfolio strategies investigated reports a reduction in the standard deviation of target portfolio weights, but as can be seen, when applying longer histo-

Table 8.24: Average volatility of portfolio target weights across all optimization rules with fixed grid-points.

P.str.	σ_M	σ_{LGC}^{-4}	Δ	σ_{LGC}^{-2}	Δ	σ_{LGC}^{0}	Δ
Panel A: *alternative portfolio rules, Window size M=120, global s.d.*							
Classic approach that ignores estimation error							
MVS	0.85752	0.86402$^{\#}$	-0.00650	0.85944$^{\#}$	-0.00192	0.85754$^{\#}$	-0.00002
Moment restrictions							
MIN	0.85328	0.86072$^{\#}$	-0.00744	0.85555$^{\#}$	-0.00227	0.85357	-0.00029
Portfolio constraints							
MVSC	0.85094	0.85192$^{\#}$	-0.00098	0.85147$^{\#}$	-0.00053	0.85121$^{\#}$	-0.00027
MINC	0.85176	0.85400$^{\#}$	-0.00224	0.85296$^{\#}$	-0.00120	0.85216$^{\#}$	-0.00040
Panel B: *alternative portfolio rules, Window size M=120, local s.d.*							
Classic approach that ignores estimation error							
MVS	–	1.00200$^{\#}$	-0.14448	0.86116$^{\#}$	-0.00364	0.86064	-0.00312
Moment restrictions							
MIN	–	1.48730	-0.63402	0.85436	-0.00108	0.85385	-0.00057
Portfolio constraints							
MVSC	–	0.85487$^{\#}$	-0.00393	0.85388$^{\#}$	-0.00294	0.85376$^{\#}$	-0.00282
MINC	–	0.85484$^{\#}$	-0.00308	0.85361	-0.00185	0.85328	-0.00152
Panel C: *alternative portfolio rules, Window size M=240, global s.d.*							
Classic approach that ignores estimation error							
MVS	0.84736	0.84852$^{\#}$	-0.00116	0.84755$^{\#}$	-0.00019	0.84716$^{\#}$	-0.00020
Moment restrictions							
MIN	0.84663	0.84938$^{\#}$	-0.00275	0.84746$^{\#}$	-0.00083	0.84678$^{\#}$	-0.00015
Portfolio constraints							
MVSC	0.84596	0.84642	-0.00046	0.84618$^{\#}$	-0.00022	0.84603$^{\#}$	-0.00007
MINC	0.84614	0.84740$^{\#}$	-0.00126	0.84674$^{\#}$	-0.00060	0.84637$^{\#}$	-0.00023
Panel D: *alternative portfolio rules, Window size M=240, local s.d.*							
Classic approach that ignores estimation error							
MVS	–	0.86682$^{\#}$	-0.01946	0.84851$^{\#}$	-0.00115	0.84827$^{\#}$	0.00091
Moment restrictions							
MIN	–	1.20500	-0.35837	0.84691$^{\#}$	-0.00028	0.84666$^{\#}$	-0.00003
Portfolio constraints							
MVSC	–	0.84961$^{\#}$	-0.00365	0.84806$^{\#}$	-0.00210	0.84791$^{\#}$	-0.00195
MINC	–	0.84856	-0.00242	0.84683$^{\#}$	-0.00069	0.84661$^{\#}$	-0.00047

Table 8.24 shows the average standard deviation in target weights when Markowitz (σ_M) or model-based (σ_{LGC}) estimates are used for the data set investigated. Panel A (B) and Panel C (D) show the average volatility within the portfolio across all strategies using different sample windows length of 120 and 240 months, respectively. $^{\#}$ indicates higher CEQ value when model-based estimates are applied as opposed to Markowitz estimates.

rical sampling windows of 240 trading months (Panel C), the reduction in average volatility are much smaller. This makes sense, since applying for longer historical sampling windows do in fact give more robust parameter estimates and less room for improvements. This is similar effects as can be found in previous sections, where estimates from shorter sampling windows often enhances a larger number of portfolio strategies that are improved, but the degree of improvement increases. Intuitively, with longer sampling windows, the local dependence estimates have less error to reduce. Thus, shorter windows that exhibit greater error result in greater improvement when the local dependence estimates that account for asymmetries are accounted for. Also, note that portfolio strategies with short-sale constraints exhibit less improvement in average volatility.

This is logical as the portfolio weights are already constrained to vary within a narrow band of values in the covariance matrix. Hence, we can see that the application of local dependence estimates in Table 8.23 "fine-tunes" the optimization procedure as the standard errors are made smaller. In Table 8.24, when portfolio target weights are estimated based on fixed grid-points, we see that the portfolio strategies across all panels exhibit higher average volatility within the target portfolio weights, when local dependence estimates are applied. We also find that as the sampling windows are lengthened, the reduction in average volatility also decreases when long-only constraints are present. Furthermore, there is in general higher volatility within the portfolio target

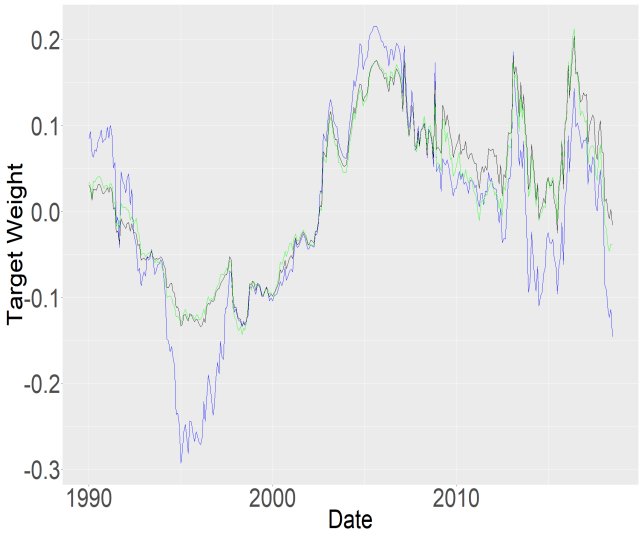
Table 8.25: Average volatility of weighted portfolio targets across all optimization rules with fixed grid-points.

P.str.	σ_M	σ_{LGC}^{80-20}	Δ	σ_{LGC}^{50-50}	Δ	σ_{LGC}^{20-80}	Δ
Panel A: *alternative portfolio rules, Window size M=120, global s.d.*							
Classic approach that ignores estimation error							
MVS	0.85752	$0.86247^{\#}$	-0.00495	$0.86038^{\#}$	-0.00286	$0.85857^{\#}$	-0.00105
Moment restrictions							
MIN	0.85328	$0.85894^{\#}$	-0.00566	$0.85659^{\#}$	-0.00331	$0.85464^{\#}$	-0.00136
Portfolio constraints							
MVSC	0.85094	$0.85174^{\#}$	-0.00080	$0.85150^{\#}$	-0.00056	$0.85131^{\#}$	-0.00037
MINC	0.85176	$0.85354^{\#}$	-0.00178	$0.85293^{\#}$	-0.00117	$0.85243^{\#}$	-0.00067
Panel B: *alternative portfolio rules, Window size M=120, local s.d.*							
Classic approach that ignores estimation error							
MVS	–	$0.96023^{\#}$	-0.10271	$0.90590^{\#}$	-0.04838	$0.86846^{\#}$	-0.01094
Moment restrictions							
MIN	–	1.32660	-0.47332	1.10000	-0.24672	$0.91313^{\#}$	-0.05985
Portfolio constraints							
MVSC	–	$0.85460^{\#}$	-0.00366	$0.85424^{\#}$	-0.00476	$0.85394^{\#}$	-0.00300
MINC	–	$0.85443^{\#}$	-0.00267	0.85390	-0.00214	0.85349	-0.00173
Panel C: *alternative portfolio rules, Window size M=240, global s.d.*							
Classic approach that ignores estimation error							
MVS	0.84736	$0.84818^{\#}$	-0.00082	$0.84773^{\#}$	-0.00037	$0.84736^{\#}$	0.00
Moment restrictions							
MIN	0.84663	$0.84874^{\#}$	-0.00211	$0.84789^{\#}$	-0.00126	$0.84718^{\#}$	-0.00055
Portfolio constraints							
MVSC	0.84596	$0.84631^{\#}$	-0.00035	$0.84618^{\#}$	-0.00022	$0.84608^{\#}$	-0.00012
MINC	0.84614	$0.84713^{\#}$	-0.00099	$0.84679^{\#}$	-0.00065	$0.84651^{\#}$	-0.00037
Panel D: *alternative portfolio rules, Window size M=240, local s.d.*							
Classic approach that ignores estimation error							
MVS	–	$0.86009^{\#}$	-0.01273	$0.85272^{\#}$	-0.00536	$0.84884^{\#}$	-0.00148
Moment restrictions							
MIN	–	1.10560	-0.25897	$0.97111^{\#}$	-0.12448	$0.87134^{\#}$	-0.02471
Portfolio constraints							
MVSC	–	$0.84924^{\#}$	-0.00328	$0.84871^{\#}$	-0.00275	$0.84822^{\#}$	-0.00226
MINC	–	$0.84811^{\#}$	-0.00197	$0.84749^{\#}$	-0.00135	$0.84694^{\#}$	-0.00080

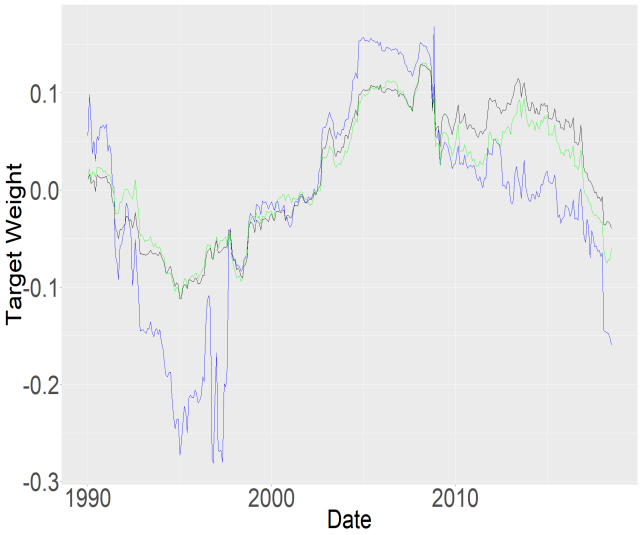
Table 8.25 shows the average standard deviation in target weights when Markowitz (σ_M) or model-based (σ^{LGC}) estimates are used for the data set investigated. Panel A (B) and Panel C (D) show the average volatility within the portfolio across all strategies using different sample windows length of 120 and 240 months, respectively. $^{\#}$ indicates a higher CEQ value when model-based estimates are applied as opposed to Markowitz estimates.

weights estimated at grid-point -4, and a decrease in volatility as we move towards normal market conditions. Furthermore, in Table 8.25, the same pattern is present as in Table 8.24, that is, less volatility within portfolio target weights near normal-market conditions as opposed to bear market conditions. Also, less volatility is present when longer sampling windows are applied. Summarizing Table 8.23, 8.24 and 8.25, we do find evidence of enhanced performance in the sense of smoother portfolio target weights when we apply moving-grid estimates, and decreased performance for fixed grid estimates. These findings, however, would result in reduced re-balancing cost when accounting for transaction costs, but again, we leave this for further investigation. For additional information on portfolio re-balancing analysis, see Vanguard [2010] and Tu and Zhou [2011].

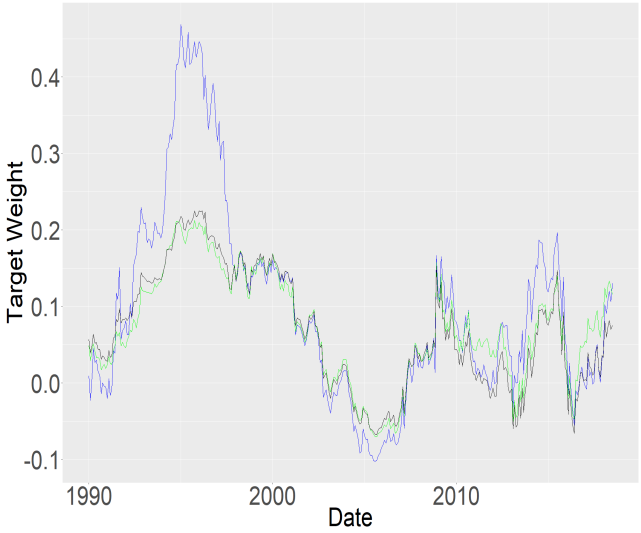
Figure 8.14: Portfolio target weights of the MVS- and MIN portfolio optimization strategies using a rolling sampling window of $M = 120$ trading months. Color code; - global estimates, - local estimates under bear market conditions- and - moving-grid estimates both using global standard deviation.



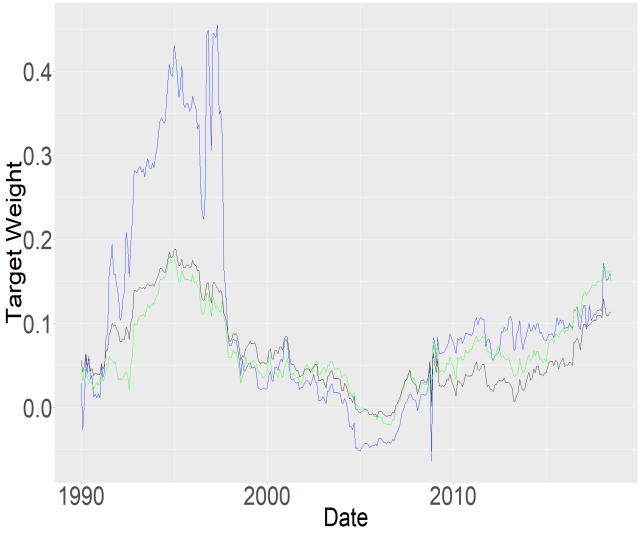
(a) FTSE All Share Index (MVS)



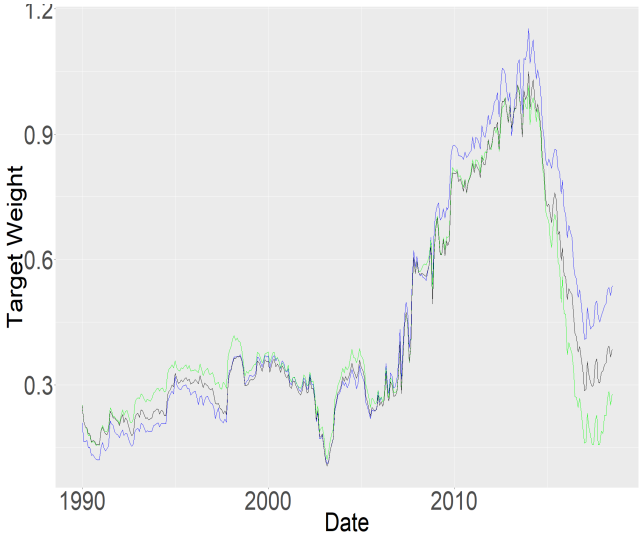
(b) FTSE All Share Index (MIN)



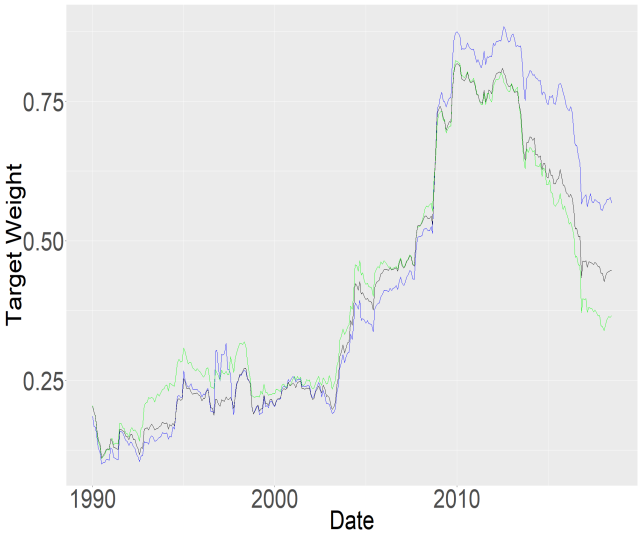
(c) S&P500 Comp. Index (MVS)



(d) S&P500 Comp. Index (MIN)

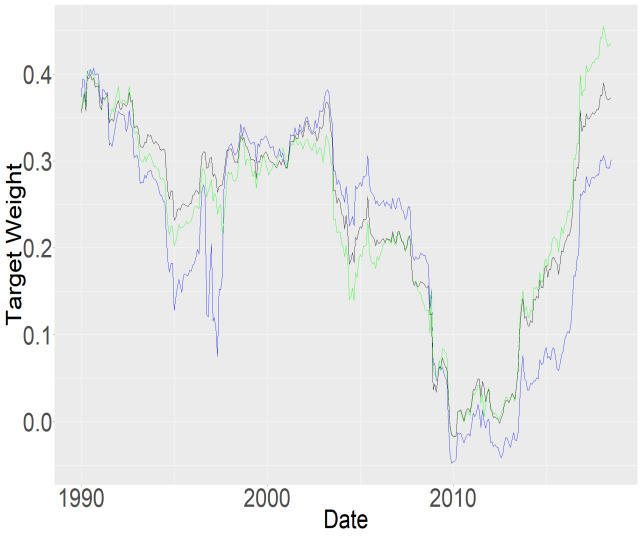


(e) UK Benchmark 10 Year DS GOVT. Index (MVS)

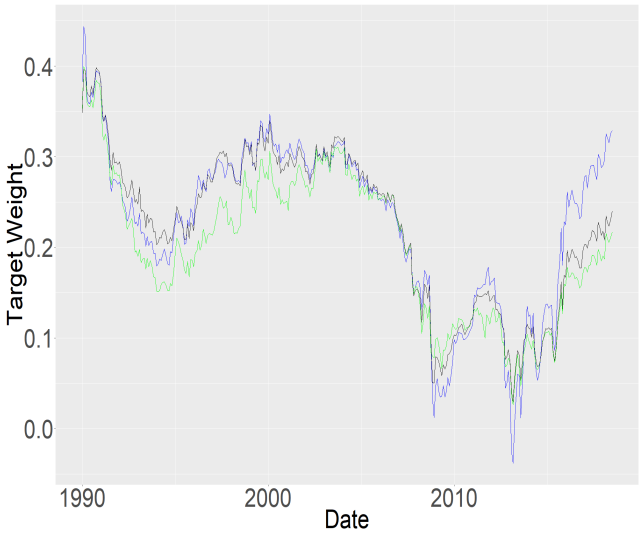


(f) UK Benchmark 10 Year DS GOVT. Index (MIN)

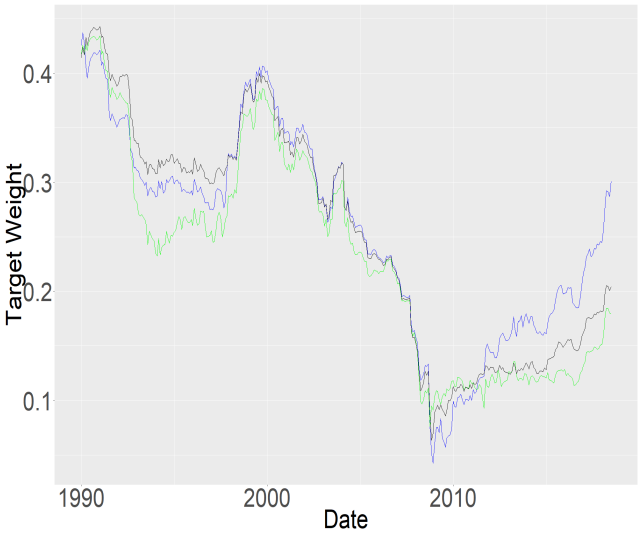
(g) US Benchmark 10 Year DS GOVT. Index (MVS)



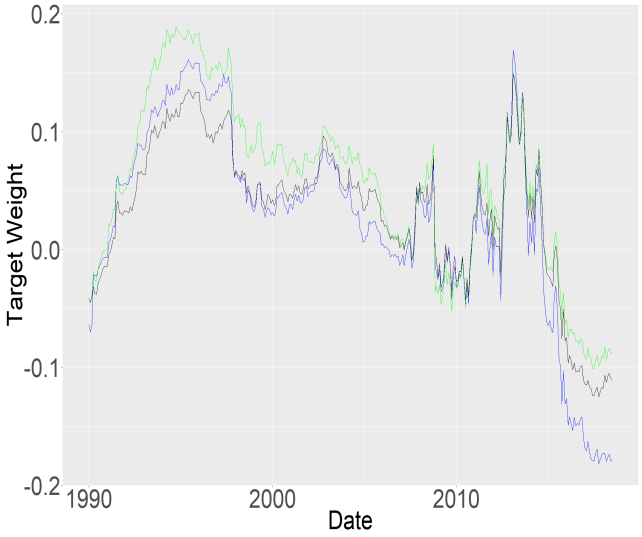
(h) US Benchmark 10 Year DS GOVT. Index (MIN)



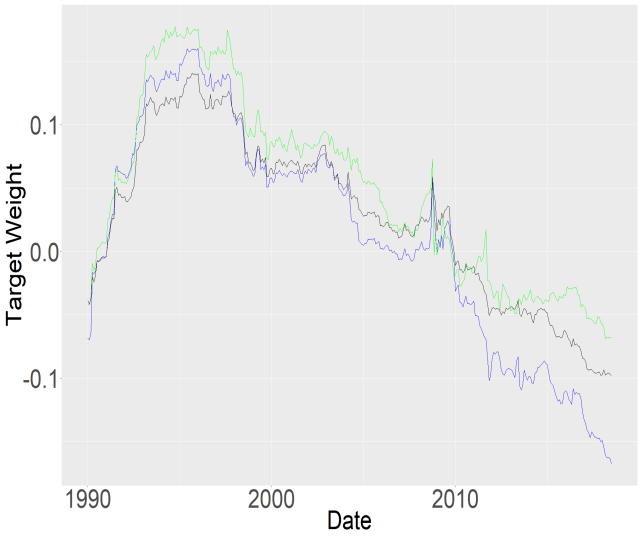
(i) TR Equal Weight CCI Index (MVS)



(j) TR Equal Weight CCI Index (MIN)



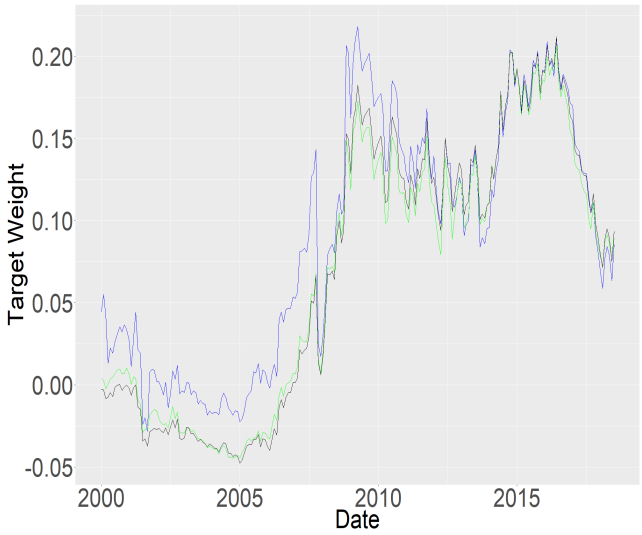
(k) S&P GSCI Gold Index (MVS)



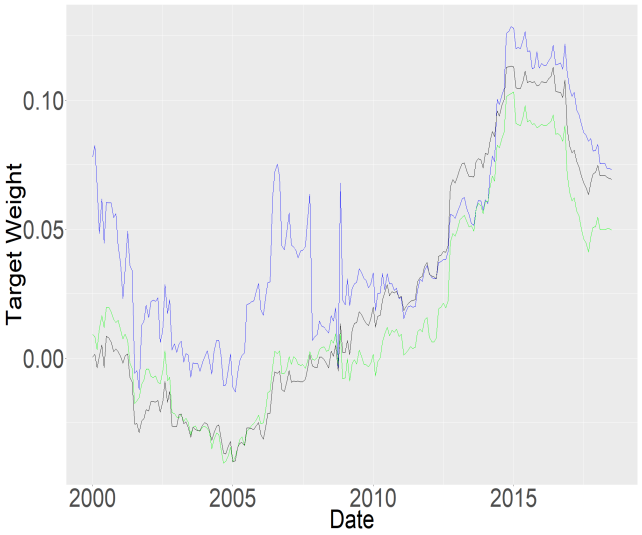
(l) S&P GSCI Gold Index (MIN)

In Figure 8.14, we do no report the rolling estimated target portfolio weights when local correlations and the corresponding local standard deviations are applied.

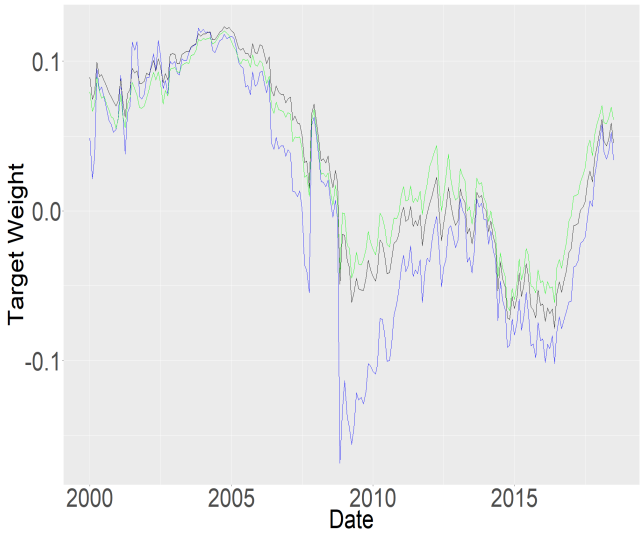
Figure 8.15: Portfolio target weights of the MVS- and MIN portfolio optimization strategies using a rolling sampling window of $M = 240$ trading months. Color code; - global estimates, - local estimates under bear market conditions- and - moving-grid estimates both using global standard deviation.



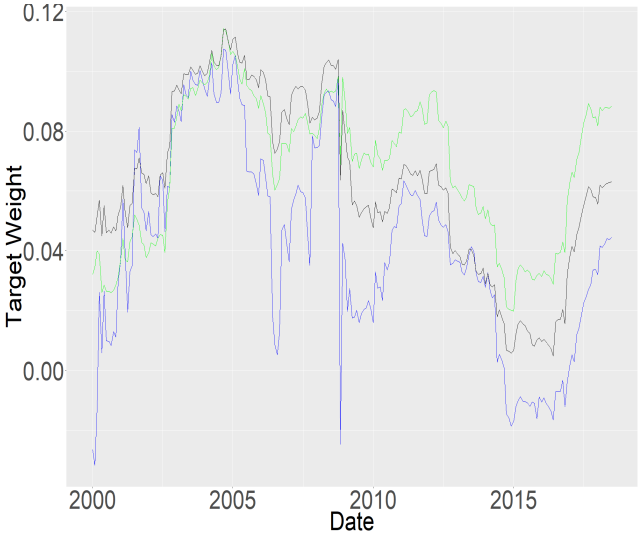
(a) FTSE All Share Index (MVS)



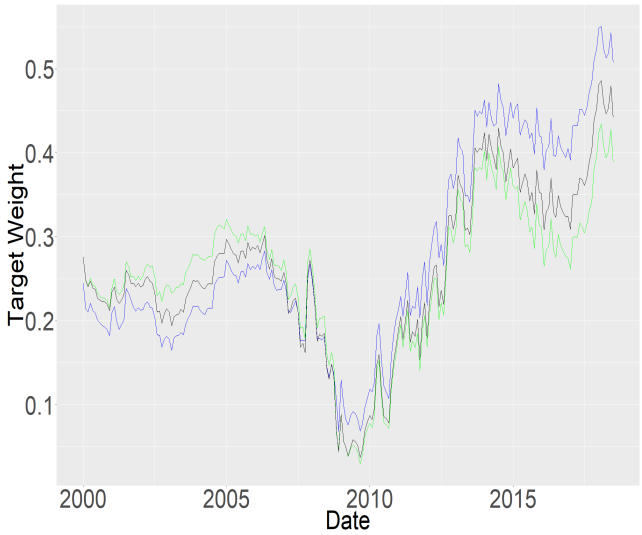
(b) FTSE All Share Index (MIN)



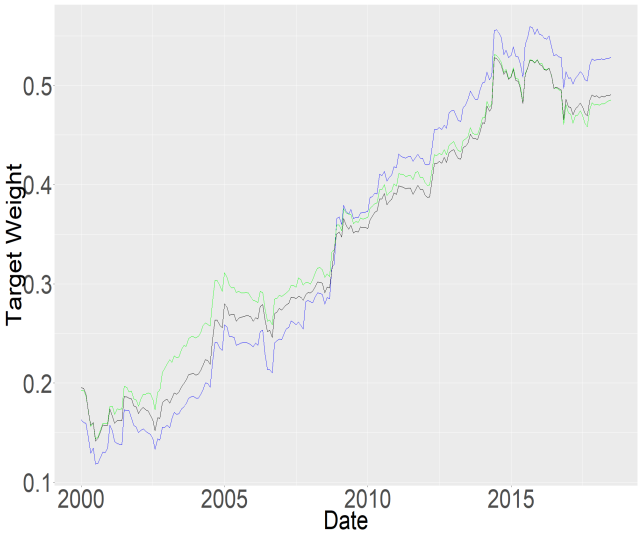
(c) S&P500 Comp. Index (MVS)



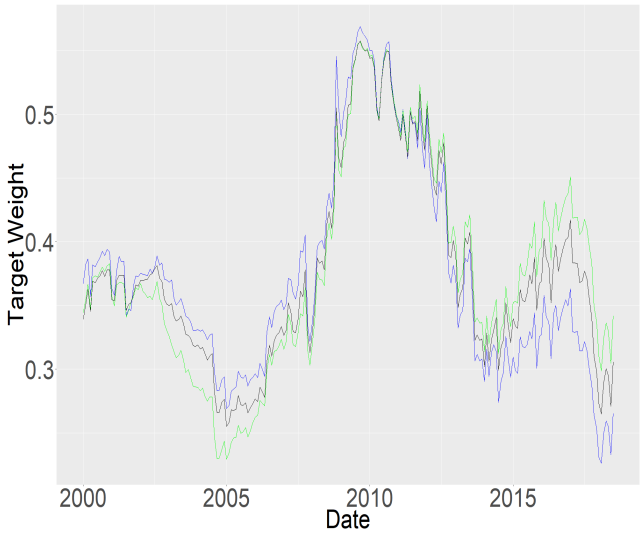
(d) S&P500 Comp. Index (MIN)



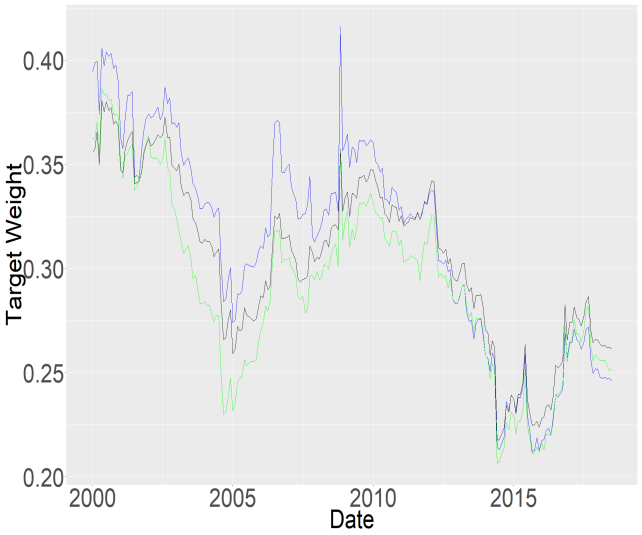
(e) UK Benchmark 10 Year DS GOVT. Index (MVS)



(f) UK Benchmark 10 Year DS GOVT. Index (MIN)

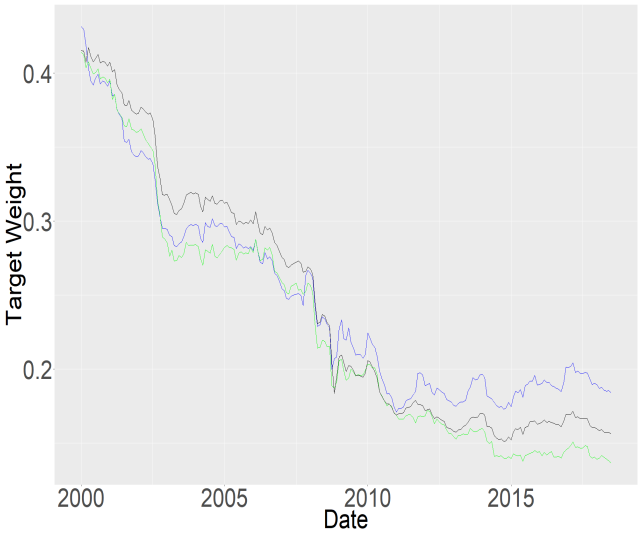


(g) US Benchmark 10 Year DS GOVT. Index (MVS)

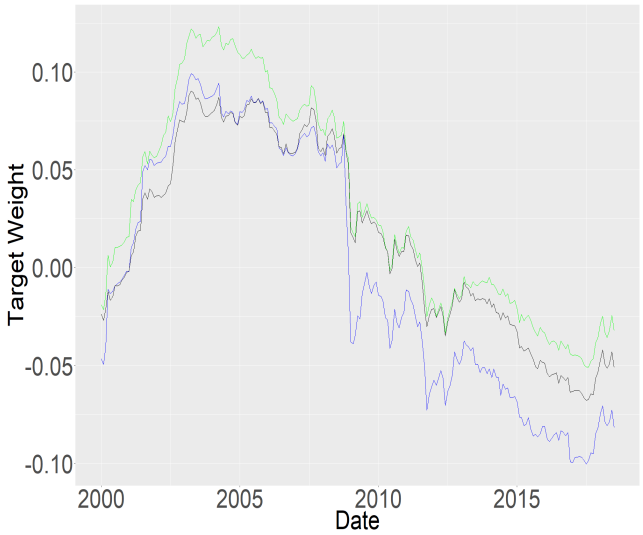


(h) US Benchmark 10 Year DS GOVT. Index (MIN)

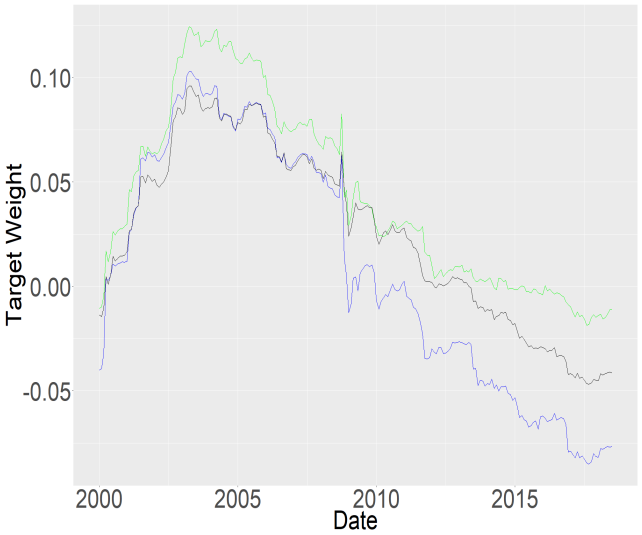
(i) TR Equal Weight CCI Index (MVS)



(j) TR Equal Weight CCI Index (MIN)



(k) S&P GSCI Gold Index (MVS)



(l) S&P GSCI Gold Index (MIN)

In Figure 8.15, we do no report the rolling estimated target portfolio weights when local correlations and the corresponding local standard deviations are applied. As mentioned earlier, these scripts can be sent upon request.

Chapter 9

Conclusion & future work

For decades, MV optimization by Markowitz [1952] has been taught across business schools globally and is widely used in the industry. But, despite the dual popularity, DeMiguel and Uppal [2009] who investigated the performance of several advances in MV optimization, find that none of these strategies consistently outperform the $1/N$ portfolio in terms of the Sharpe Ratio and CEQ, across a range of data sets. For academics and practitioners alike, these findings are troublesome and have intriguing implications regarding the empirical applications of portfolio optimization and modern portfolio theory. DeMiguel and Uppal [2009] use rolling-sampling windows (M=60 and 120) of historical returns which might not be very informative or reliable, and do not account for asymmetries in the return distributions. Notably, Markowitz [1952] explicitly recommends the use of a probability model to generate the inputs required by the MV model. Furthermore, Fantazzini [2009] finds that when the financial returns data exhibit asymmetries such as skewness, and symmetric marginals are applied, the estimated elliptical correlations can be negatively biased by as much as 70% of true values, thereby causing a large degree of estimation error. Accordingly, our study focuses on improving the covariance matrix used as input to a range of MV optimization rules by using a model-based approach. More specifically, in this thesis we have applied a new measure of dependence, called the local Gaussian correlation, to study the dependence structure between international stock market returns and thereby used the estimated local covariance matrix as input into various portfolio optimization models. Summarizing, the LGC provides us with a way of describing the changes in dependence and the departure from global normality for bivariate non-normal data.

By using the local measurements of LGC, we do find evidence of asymmetric dependence structures between international markets, and in particular between the US and European stock markets, where a quite strong bear market effect is present. Thus our findings support earlier studies (see e.g. Støve and Tjøstheim [2014]). We also demonstrate that in general, local correlation between pairwise stock and commodity indices from the US and UK markets do in fact increase with time, whereas the pairwise correlation against the government bond indices decreases with time. We conclude that the LGC gives a much more detailed picture than ordinary global correlation measures, and interestingly, the results provide evidence that accounting for correlation asymmetries within international markets enhance portfolio performances. As DeMiguel and Uppal [2009], we use rolling sampling windows of different sizes, but this might not be as informative and reliable as if we would have used the whole sampling window investigated, because we may do not account for the "true" asymmetry within the data, and hence exclude potential benefits of diversification when relying on the asymmetric dependence structure within the market. Therefore, we follow the same procedure as in Tu and Zhou [2011], and apply longer sampling windows of 120 and 240 trading months, respectively. Based on our analysis, we concur that poor performance of several MV optimization strategies as can be found in DeMiguel and Uppal [2009], are in fact a result of only relying on raw historical return data and the global covariance matrix. We have compared the performance of 4 different portfolio allocation models of optimizing portfolio target weights, comparing both global and local estimates of covariances to one another, and relative to the benchmark $1/N$ model. Our results show that across the optimization rules investigated, improvements in terms of the out-of-sample Sharpe Ratio and CEQ return values are obtained when we account for asymmetries within the return distributions, but without accounting for realistic transaction costs. We find evidence of portfolio strategies that do in general outperforms are often variants of the MIN and MINC strategies, which are also often the most robust to estimation errors since they only relies on the covariances and ignores the target returns. Although, strategies allowing for short-selling do improve the most across all strategies. These comparisons are undertaken using six different asset return series, as well as simulated data. Further, application of the local approach to the various portfolio optimization rules with the goal of outperforming the 1/N portfolio strategy is an elusive challenge, since the benefits to be gained may be outweighed by the impact of transaction costs. But, this has to be further investigated in future works. Thus, aside from very low turnover requirements, the equally weighted $1/N$ portfolio produces superior performance when evaluated over long sample periods of several decades as it avoids concentrated positions.

Hence, scrutinizing the range of models that are enhanced, we find that accounting for return asymmetries does improve the estimates of the covariance matrix, and thus we find evidence that several MV optimization strategies are improved and thus continues to be a viable empirical framework for investors.

Summarizing the empirical results; In section 8.2.1 we applied the Buy-and-Hold investment strategy, comparing local- against global estimates, and relative to the benchmark $1/N$ strategy. In the analysis, we did in fact find evidence of outstanding performance for the MVS and MIN strategies when local covariance estimates was fit into the various models (see Figure 8.10 and 8.11), which are supported by both estimation methods of "5par" and "5par-marginals-fixed", and sampling windows of 120 and 240 trading months, respectively. Although, as can be seen in Table 8.4 and 8.9, the MIN strategy produce extreme portfolio target weights. In section 8.2.2, we delved into the out-of-sample analysis of our portfolio optimization strategies using rolling sampling windows of 120 and 240 trading months. In Table 8.23 however, we do see that the LGC model provides enhanced estimates of the covariance matrix that lead to improved performance with smaller adjustments to portfolio target weights through the $T - M$ out-of-sample period. Results in Table 8.24 and 8.25, before accounting for transaction costs, show that application of local portfolio estimates continues to produce superior CEQ values. Although, we find that most of the portfolio strategies that demonstrate superior performances with the application of local covariance estimates, also exhibit the desirable attribute of reduced average standard deviation and thus give smoother target portfolio weights when using a moving-grid, whereas we do in fact produce more volatile target portfolio weights when we use the strategies relying on local estimates from fixed grid-points (see Figure 8.14 and 8.15). Even though the portfolios produce more volatile target weights, they do outperform the classical global portfolio estimates in the sense of Sharpe Ratio and CEQ. Concluding the empirical results in this thesis, by accounting for asymmetric dependence within the return distribution, we are able to manipulate the risk-and-return characteristics in our favor without sacrificing portfolio return, and thus by only relying on adjustments in the covariance matrix.[1]

The results in this thesis are supported by several robustness checks; (1) the sampling window lengths are held fixed and are rolling with time, rather than increasing with time, more specifically, equal to $M = 120$ and 240 trading months; (2) the holding period is one trading month rather than one trading day or year; (3) the portfolio evaluated is consisting of only risky assets; (4) the out-of-sample portfolio performances are measured relative to the benchmark $1/N$ strategy with re-balancing, rather than the $1/N$ strategy without re-balancing; (5) we use a risk aversion coefficient of $\gamma = 1$, rather than some other value, such as $\gamma = \{2, 3, 4, 5\}$. We have generated tables for the Sharpe Ratio and CEQ returns in order to check whether our results are sensitive to these assumptions. In addition, based on each of these two measures, we also report the volatility- and average portfolio's excess return of the various strategies investigated and the average volatility within the portfolio target weights, as done in section 8.2.3 and 8.2.4. Furthermore, we plot the cumulative return- and volatility within the portfolio target weights in order to give an insight in the portfolio re-balancing analysis.

In this thesis we have studied portfolio selection for stocks, bonds and commodity indices with certain characteristics, and the local dependence across these international markets. Using simulated data, we show that the extensions to the various mean-variance models that have been designed to deal with estimation error reduce only moderately the estimation error in a few models, but as mentioned, the dependence structure can be easily adjusted in order to demonstrate in even more fine details the possible diversification benefits by applying the LGC estimates. Furthermore, with the analysis and tools derived in this thesis, it is possible to delve even further into the study of correlation effects within financial markets. In principle, every financial or econometric analysis that depends on a covariance matrix can be subject to an analysis using the local Gaussian correlation model. Also note that models such as the Student t-GARCH(1,1) process can be fitted to the data before analyzing them. One particular paper cited describes the results from such models (see e.g. Støve and Tjøstheim [2014]). Clearly, this should be more closely investigated. A more sophisticated way of estimating the local covariances using moving-grid restriction under bear-market conditions only would be to remove the "diagonal" restriction and therefore let it work on the entire grid-map under bear market conditions. But, they do in fact produce superior results in form of portfolio's excess return, but again, they also produce highly volatile portfolio target weights, especially in periods of high volatility within the return series. Further, the moving-grid allowing for returns across the entire grid-map (e.g. 4×4) should also be more closely investigated, as the strategy is exposed to a large degree of estimation error (e.g., NAs in the portfolio target weights). A possible solution would be to apply weighted correlation values within the grid-map. Furthermore, applying predicted returns from other sophisticated models would be a possible extension to the various models investigated in this thesis. Another possible extension, as mentioned earlier in this thesis, would be introducing a sophisticated combination rule to be a sort of trade-off between bias and variance. It would also be of interest to see whether the local dependence patterns continue to hold for other comparable markets, such as the great East Asian economies, and how they are changing over time.

[1] Total Words: 20500; Headers: 60; Math Inline 538; Math Display: 81. The entire file in raw format can be sent upon request.

Appendices

Appendix A

Additional Theory, Figures & Tables

Definition A.0.1. *In statistics, the Pearson correlation coefficient is a measure of linear correlation between two variables X and Y. According to the Cauchy–Schwarz inequality it has a value in [-1,1] and defined as*

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y}. \tag{A.1}$$

Definition A.0.2. *One popular way of analyzing asymmetry is by conditional correlation, and for two return series, $\{X_t\}$ and $\{Y_t\}$, where we assume stationarity, the conditional correlation when restricted to a set A is*

$$\rho_A = corr(X_t, Y_t | (X_t, Y_t) \in A). \tag{A.2}$$

We can compute ρ_A for any set A, and as the number of observation increases, we can let the size of the restricted area A be made smaller. In this way we can obtain a conditional correlation in a small neighborhood A of (x,y). Such a localization is a basic feature of non-parametric analysis, where the size of A would be determined by a pair of bandwidths (b_1, b_2) in the x and y direction, respectively. Asymptotically, as the number of observations increases, (b_1, b_2) can be made smaller, making it possible to describe ever more fine details.

Extreme value theory (EVT), is a branch of statistics dealing with the extreme deviations from the median of probability distributions. It seeks to assess, from a given ordered sample of a given random variable, the probability of events that are more extreme than any previously observed. To illustrate EVT, imagine the normal distribution. The elegant thing about the normal distribution, is that it can be fully described by the first two moments. The first moment is the mean or the average of our data, and the second moment is a measure of dispersion, or the volatility for that matter. Further, the third moment is a measure of the tilt or skew, and intuitively, normal is always symmetrical, and therefore the skew is equal to 0. The fourth moment is a measure of the tail weight or density, and that is where the problem occurs.

Everything exceeding the given threshold is an extreme loss, and can further be described by its own distribution. This illustrates the problem with any EVT, because we in most cases are in an area where we don't have much data. The distribution specifically described for the "extreme tail loss" is an important solution to the problem of VaR for the purposes of risk measurement, and can further be described by eq. A.3 and A.4.

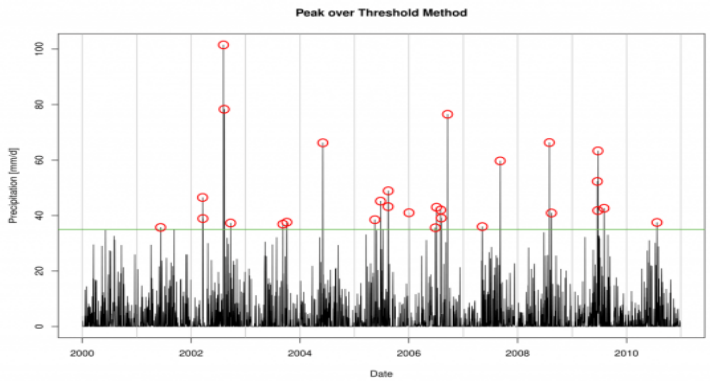


Figure A.1: Illustration of Peak-Over-Threshold. Source; Matthias [2015].

Financial risks often relate to negative developments in financial markets. As mentioned in chapter 4, financial risks are classified into the broad categories of market risks, credit risks, liquidity risks and operational

risks. The focus here will naturally be on the market risk, which arises from changes in the prices of financial assets. The global standard risk measure today used by regulators and investment bankers, is the Value-at-Risk (VaR) measurement. As a reminder, VaR represents the worst expected loss over a given time interval under normal market conditions at a given confidence level.

Estimation of VaR based on the GARCH(1,1) model

Now that we have implemented the routines, we would like to compare the different approaches that we discussed in section 4.1.2. We compare the methods by counting the number of points exceeding the VaR to assess the efficiency of the estimator, in this regard, using the dataset given in Table 8.1. We apply different VaR estimators based on a range of 464 historical points, make the one-step forecast, and count the number of events that exceed the forecast VaR. Such calculations can be executed in the form of parallel programming, using the recommended **R** package **parallel**.

Table A.1: VaR estimator based on the GARCH(1,1)-EVT, GARCH(1,1), EVT and empirical approaches.

	$\text{VaR}_{0.95}$	$\text{VaR}_{0.99}$	$\text{VaR}_{0.995}$
Expected violations	75	15	8
GARCH(1,1)-EVT	NA	NA	NA
GARCH(1,1)	44	14	9
EVT	72	20	12
Empirical	72	23	12

Table A.1 provides the comparison of a number of VaR results based on different approaches. Table A.1 report the expected number of points exceeding VaR at three levels of confidence, that is, 95%, 99%, 99.5% across five different approaches.

Definition A.0.3. *The probability density function of the GPD is defined as*

$$g_{\zeta,\beta}(x) = \begin{cases} \frac{1}{\beta}(1+\frac{\zeta x}{\beta})^{-\frac{1}{\zeta}-1} & \text{if } \zeta \neq 0, \\ \frac{1}{\beta} exp(-\frac{x}{\beta}) & \text{if } \zeta = 0, \end{cases} \tag{A.3}$$

where β is a scale parameter and ζ a shape parameter. The support of the distribution is, when $\zeta \leq 0, x \leq 0$, and when $\zeta < 0$, it is $0 \geq x \geq -\beta/\zeta$.

Definition A.0.4. *The VaR_α can be expressed in terms of the distribution of the exceeding points as*

$$VaR_\alpha = u + \frac{\hat{\beta}}{\hat{\zeta}}\Big[\Big(\frac{n}{N_u}(1-\alpha)\Big) - 1\Big], \tag{A.4}$$

where $\hat{\beta}$ and $\hat{\zeta}$ are the parameters of the GPD, u is the threshold, N_u is the number of points exceeding the threshold and, n is the number of total values of the dataset.

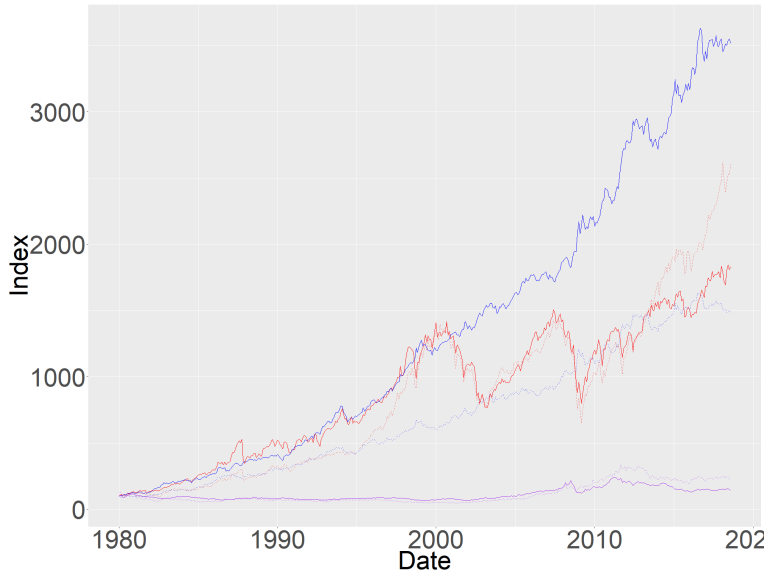
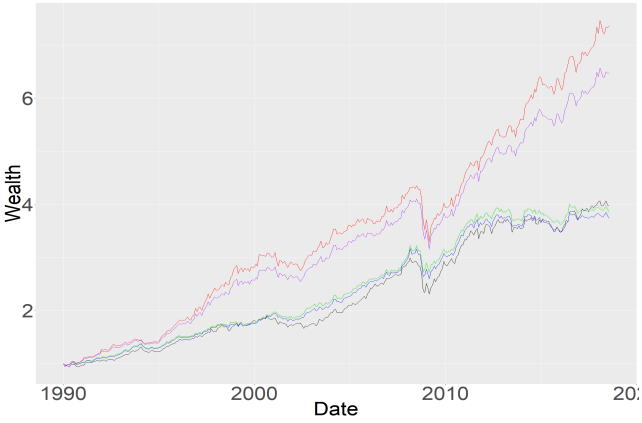
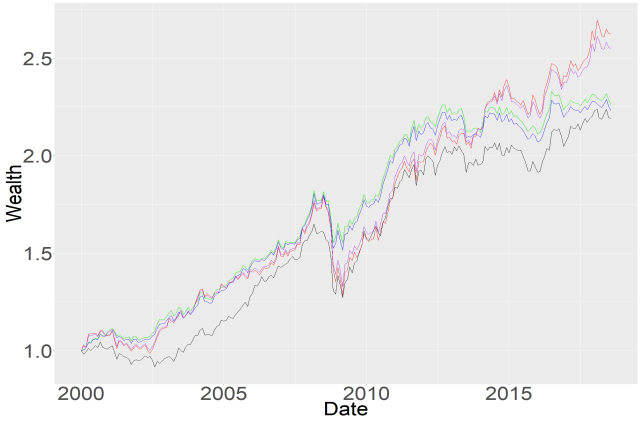


Figure A.2: Change in asset prices for stocks (-), gov. bonds (-) and commodity indices (-), normalised to 100 at the start date, denoted in %. Dotted lines; the S&P500 Comp., BMUS10Y and S&P GSCI Gold Indices.

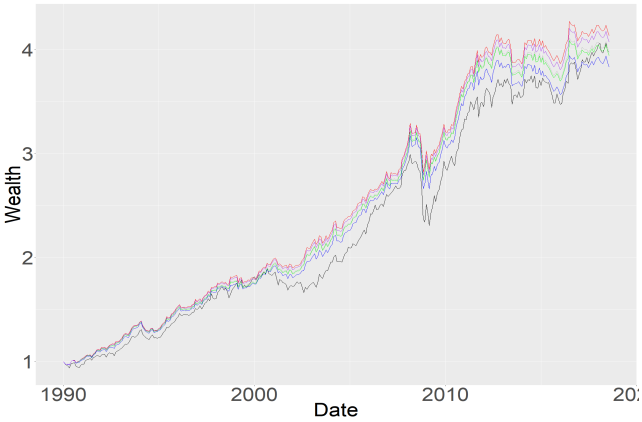
Figure A.3: Buy-and-Hold portfolio performance using local covariances (local correlation- and corresponding s.d.) and estimation method of "5par" and sampling windows of 120 and 240 trading months.



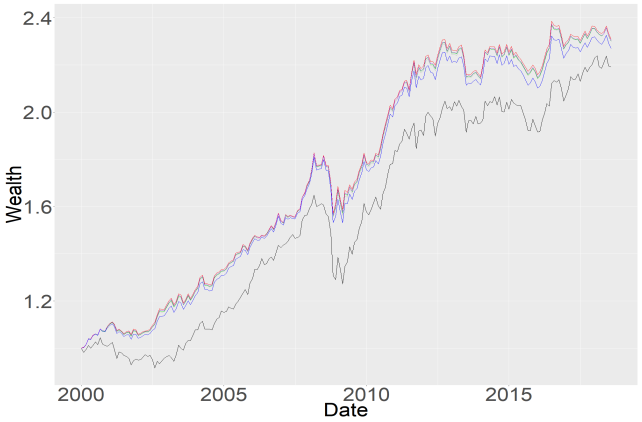
(a) Mean-Variance with short-selling (MVS)



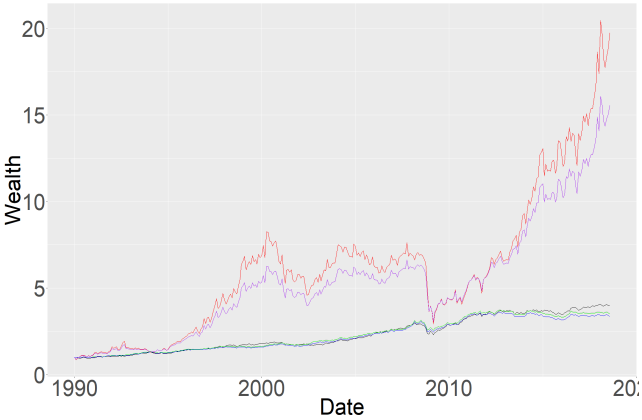
(b) Mean-Variance with short-selling (MVS)



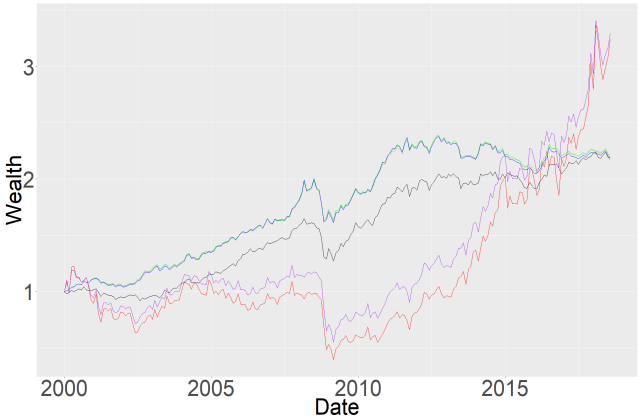
(c) Mean-Variance without short-selling (MVSC)



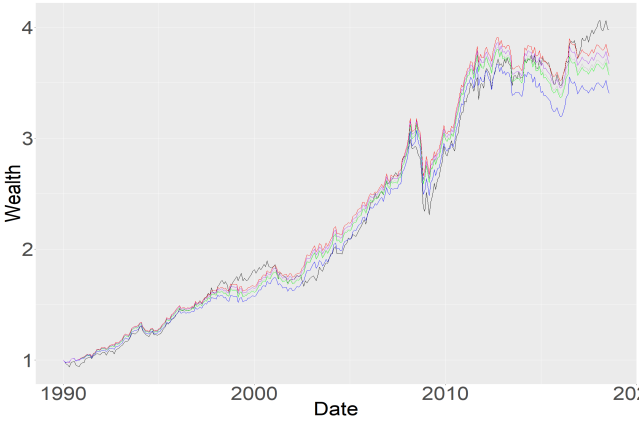
(d) Mean-Variance without short-selling (MVSC)



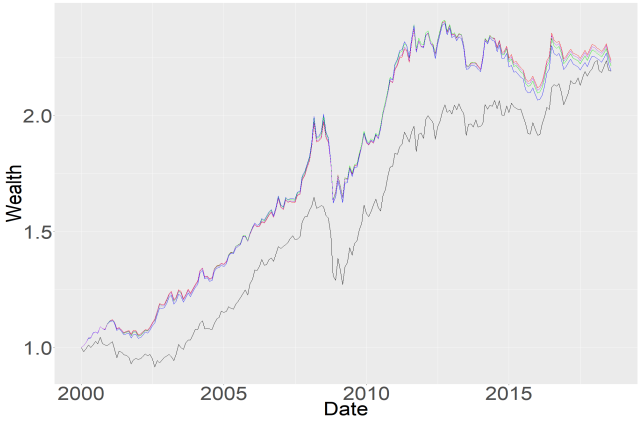
(e) Minimum-Variance with short-selling (MIN)



(f) Minimum-Variance with short-selling (MIN)

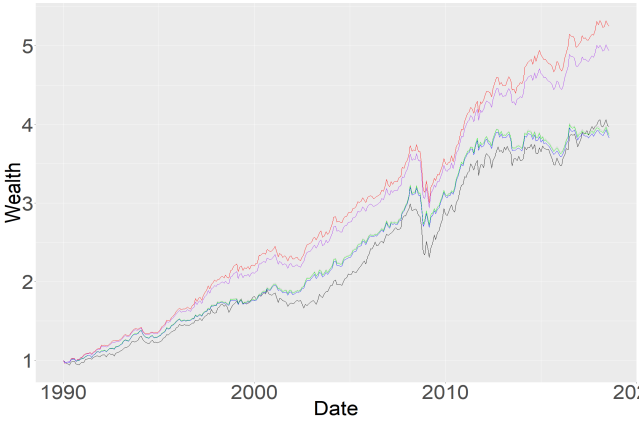


(g) Minimum-Variance without short-selling (MINC)

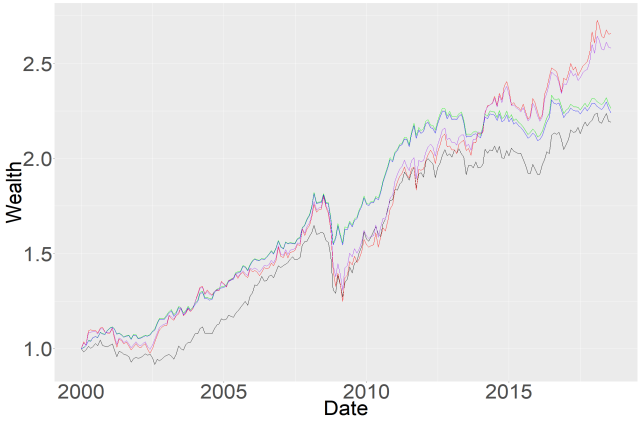


(h) Minimum-Variance without short-selling (MINC)

Figure A.4: Buy-and-Hold portfolio performance using local covariances (local correlation- and corresponding s.d.) and estimation method of "5par-marginals-fixed" and sampling windows of 120 and 240 trading months.

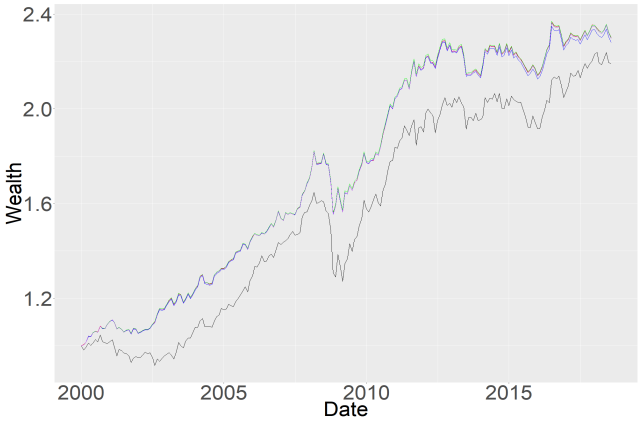


(a) Mean-Variance without short-selling (MVS)

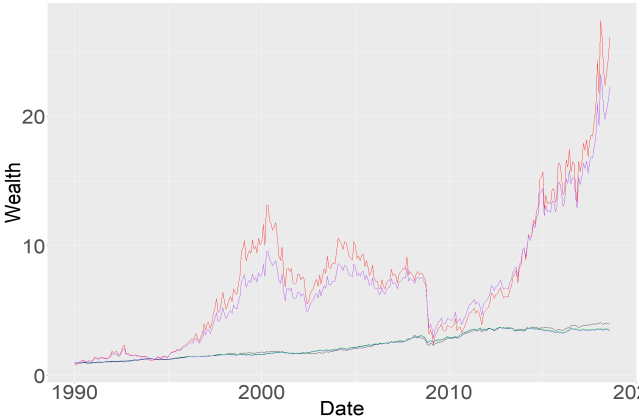


(b) Mean-Variance without short-selling (MVS)

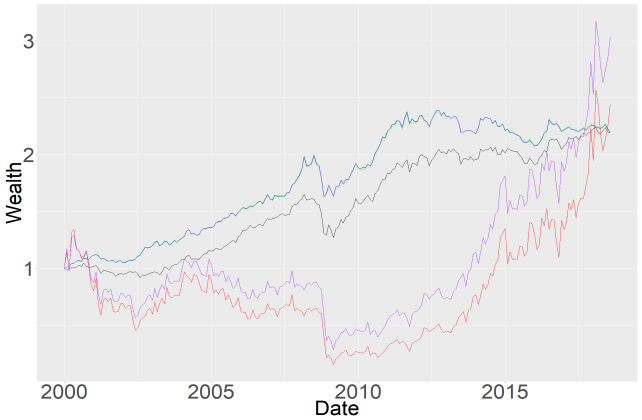
(c) Mean-Variance with short-selling (MVSC)



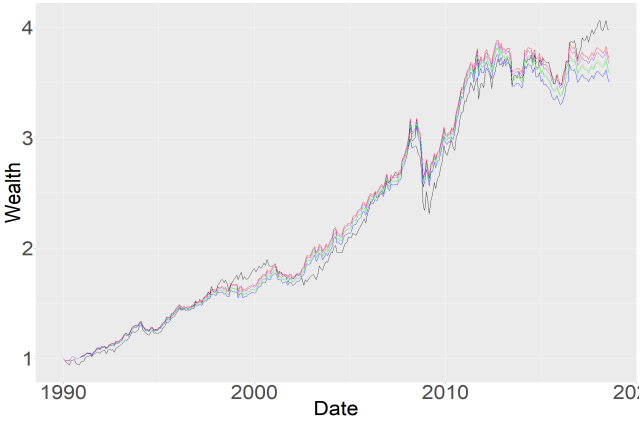
(d) Mean-Variance with short-selling (MVSC)



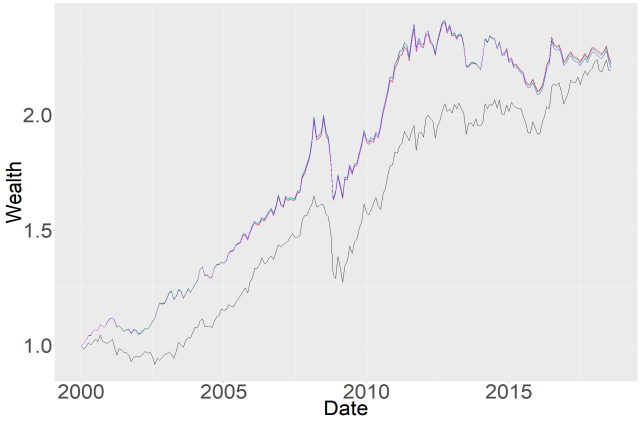
(e) Minimum-Variance without short-selling (MIN)



(f) Minimum-Variance without short-selling (MIN)

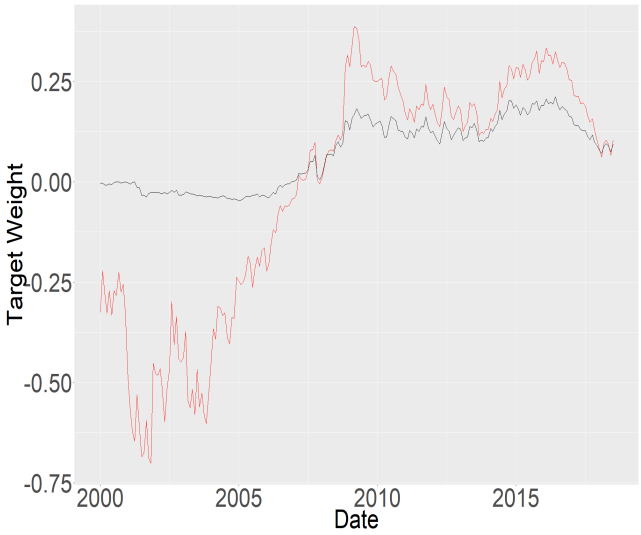


(g) Minimum-Variance with short-selling (MINC)



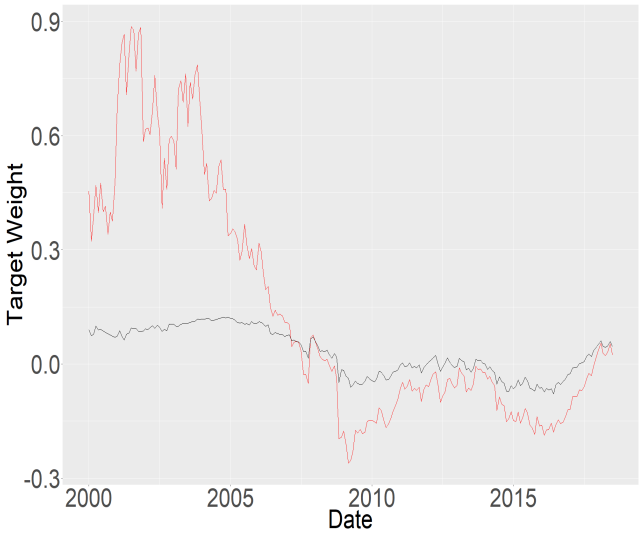
(h) Minimum-Variance with short-selling (MINC)

Figure A.5: Portfolio target weights of the MVS- and MIN portfolio optimization strategies using a rolling sampling window of $M = 240$ trading months. Color code; - global estimates, - local correlation- and the corresponding local s.d. estimates under bear market conditions.

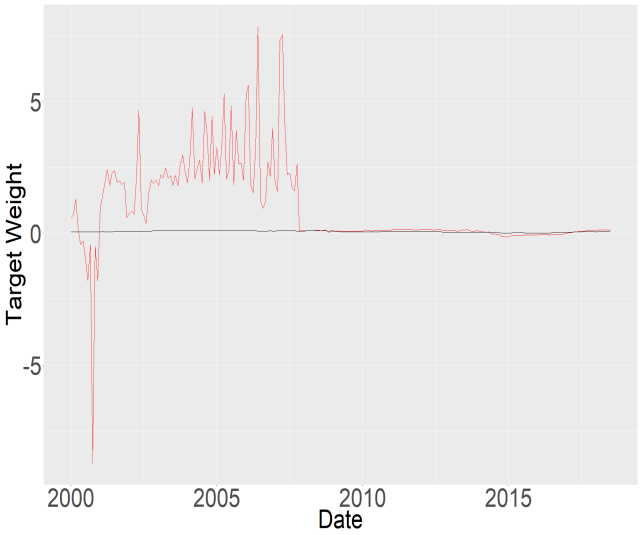


(a) FTSE All Share Index (MVS)

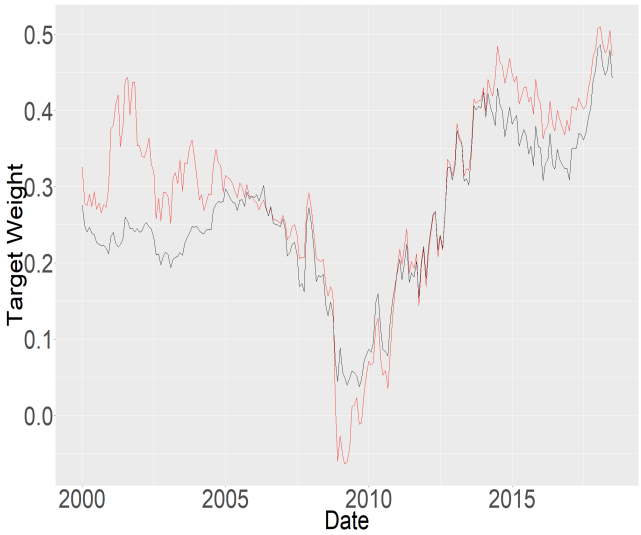
(b) FTSE All Share Index (MIN)



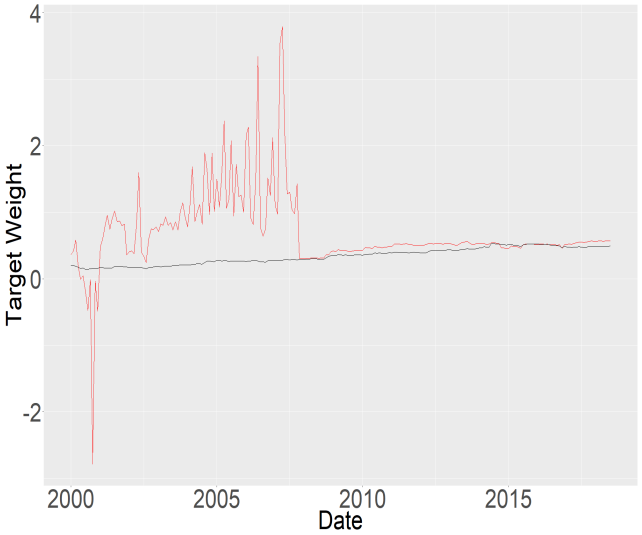
(c) S&P500 Comp. Index (MVS)



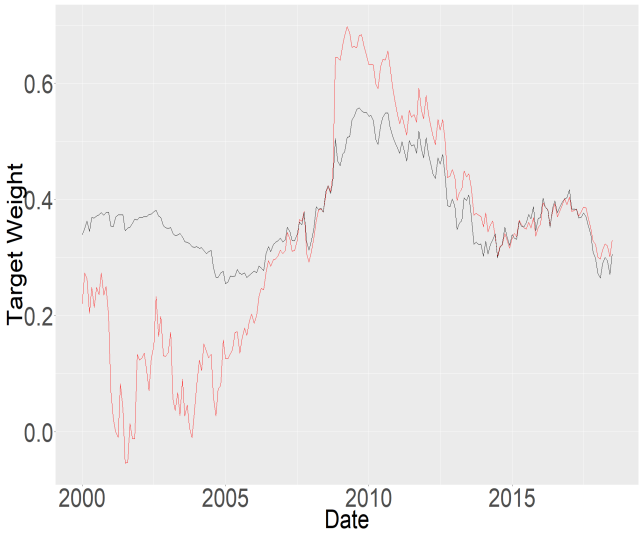
(d) S&P500 Comp. Index (MIN)



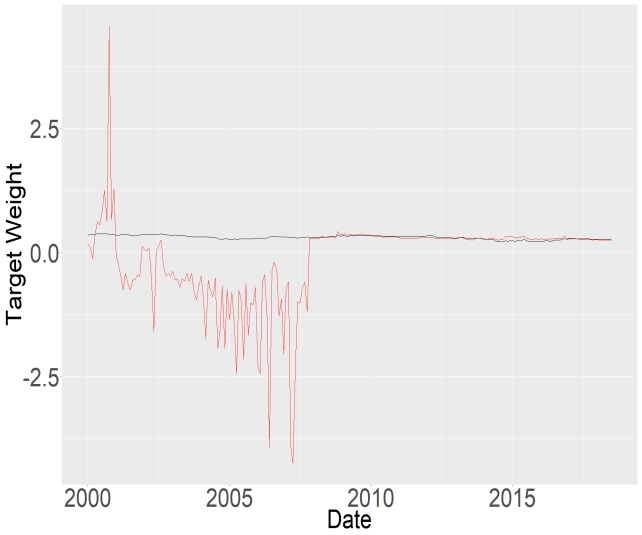
(e) UK Benchmark 10 Year DS GOVT. Index (MVS)



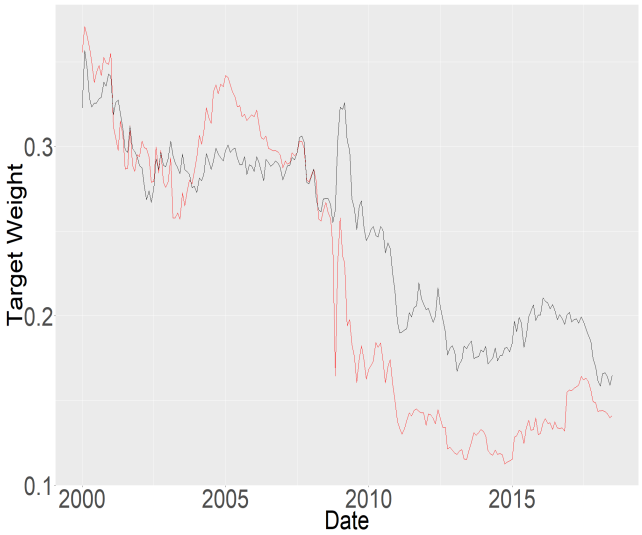
(f) UK Benchmark 10 Year DS GOVT. Index (MIN)



(g) US Benchmark 10 Year DS GOVT. Index (MVS)

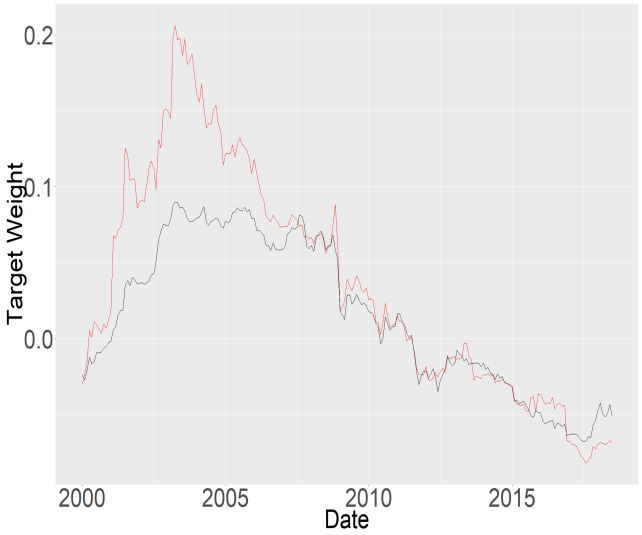


(h) US Benchmark 10 Year DS GOVT. Index (MIN)

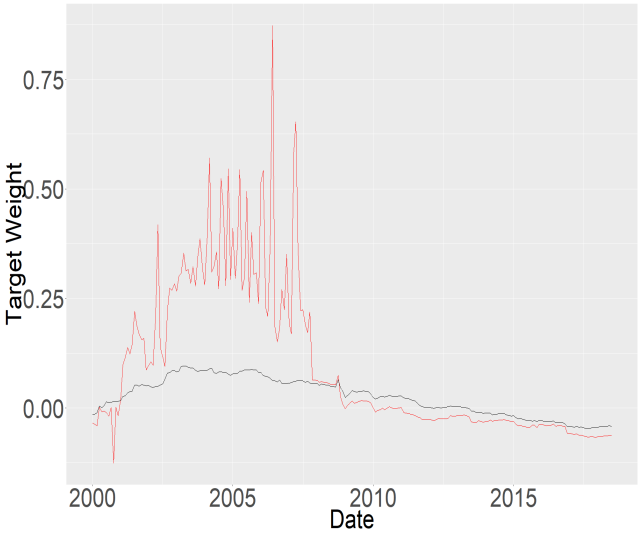


(i) TR Equal Weight CCI Index (MVS)

(j) TR Equal Weight CCI Index (MIN)



(k) S&P GSCI Gold Index (MVS)



(l) S&P GSCI Gold Index (MIN)

Note: The MVS- and MIN model deviates a lot from the global estimates, although, the MIN model produce extreme portfolio target weights in the period of 2000-2010. We did also investigate the portfolio target weights estimated with both local correlation- and corresponding local s.d., using a sampling window of $M = 120$ trading months (see Table 8.24 and 8.25 for more details), although, we do not report them in this thesis.

Bibliography

Aas, K. and X. K. Dimakos (2004). "Statistical Modelling of Financial Time Series: An Introduction". *Norwegian Computing Center 28*(March), page 37.

Anderson-Cook, C. M. *"Quantitative Risk Management"*. Number 476.

Artzner, Pasteur, D. H. E. G. H. (1996). "Coherent Measures of Risk". (June 1996), 1–24.

Aslanidis, N. and I. Casas (2013). "Nonparametric Correlation Models for Portfolio Allocation". *Journal of Banking & Finance 37*(7), 2268–2283.

Berentsen, Kleppe, T. (2014). "Introducing localgauss, an R Package for Estimating and Visualizing Local Gaussian Correlation". *Journal of Statistical Software 59*(10).

Bessler, W. and D. Wolff (2015). "Do commodities add value in multi-asset portfolios? An out-of-sample analysis for different investment strategies". *Journal of Banking and Finance 60*, 1–20.

Bollerslev, T. (1986). "Generalized Autoregressive Conditional Heteroskedasticity while conventional time series and econometric models operate under an assumption of constant variance, the ARCH process introduced in Engle (1982)". *31*, 307–327.

Bollerslev, T., R. Y. Chou, and F. Kroner (1992). "ARCH Modelling". *52*, 5–59.

Charpentier, A. (2015). "Computational Actuarial Science with R". *1*, 407–472.

Daskalaki and Skiadopoulos (2011). "Should investors include commodities in their portfolios after all? New evidence". *Journal of Banking and Finance 35*(10), 2606–2626.

DeMiguel, Garlappi, N. and Uppal (2009). "Optimal vs Naive Diversification: How inefficient is the 1/N Portfolio Strategy?".

Embrechts, P., F. Lindskog, and A. Mcneil (2001). "Modelling Dependence with Copulas and Applications to Risk Management".

Embrechts, P., A. Mcneil, and D. Strauman (2002). "Correlation and Dependency in Risk Management". (March).

Engle (1982). "Autoregressive Conditional Heteroscedacity with Estimates of variance of United Kingdom Inflation".

Engle, R. and R. Colacito (2006). "Testing and Valuing Dynamic Correlations for Asset Allocation". *Journal of Business and Economic Statistics 24*(2), 238–253.

Fantazzini, D. (2009). "the effects of misspecified marginals and copulas on computing the value at risk: A monte carlo study". *Computational Statistics Data Analysis 53*(6), 2168–2188.

Hjort and Jones (1996). "Locally Parametric/Non-parametric Density Estimation". *24*(4), 1619–1647.

Jagannathan and Ma (2003). "Risk Reduction In Large Portfolios".

Ling and Li (1997). "On fractionally integrated autoregressive moving-average time series models with conditional heteroscedasticity". *Journal of the American Statistical Association 92*(439), 1184–1194.

Ling and McAleer (2001). "Necessary and Sufficient Moment Conditions for the GARCH(r,s) and Asymmetric Power GARCH(r,s) Models". (534).

Markowitz, H. (1952). "Portfolio Selection - Harry Markowitz". *7*(1), 77–91.

Matthias (2015). "Peak over Threshold". (https://www.gis-blog.com/eva-intro-3/).

McNeil (2005). *"Quantitative Risk Management"*, Volume 101.

McNeil, A. J. and R. Frey (2000). "estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach". *Journal of Empirical Finance 7*(3-4), 271–300.

Munasca (2015). "Minimum-Variance frontier of MPT". (https://commons.wikimedia.org/wiki/File:Minimum_variance_flontier_of_MPT.svg).

Otneim (2017). "Package 'lg' R".

Otneim and Tjøstheim (2017). "Conditional Density Estimation using the Local Gaussian Correlation". *Statistics and Computing 28*(2), 303–321.

Posedel, P. (2005). "Properties and Estimation of GARCH (1,1) Model". *2*(2), 243–257.

Reider, R. (2009). "Volatility forecasting I: GARCH models". *New York*, 1–16.

Ruppert, D. and D. S. Matteson (2015). *"Statistics and Data Analysis for Financial Engineering"*.

Støve (2017). "Portfolio Allocation under Asymmetric Dependence (Working Paper)". (December), 1–13.

Støve and Tjøstheim (2014). "Measuring asymmetries in financial returns: an empirical investigation using local Gaussian correlation". (204321), 307–329.

Tjøstheim and Hufthammer (2012). "Local Gaussian Correlation: a new measure of dependence". (September).

Tu, J. and G. Zhou (2011). "Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies". *Journal of Financial Economics 99*(1), 204–215.

Vanguard (2010). "Best Practices for Portfolio Rebalancing". *Vanguard* (July), 1–17.

Yew-Low et al. (2016). "Enhancing Mean-Variance Portfolio Selection by Modeling Distributional Asymmetries". *Journal of Economics and Business 85*, 49–72.