COVARIANCE MATRIX CONSTRUCTION AND ESTIMATION: CRITICAL ANALYSES AND EMPIRICAL CASES FOR PORTFOLIO APPLICATIONS

MARCO NEFFELLI

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
PHD PROGRAMME IN ECONOMICS, XXXI CYCLE

DEPARTMENT OF ECONOMICS AND BUSINESS STUDIES
UNIVERSITY OF GENOVA

2019

Contents

List of Tables

List of Figures

List of Algorithms

Acknowledgments

Almost all the research contained in this PhD thesis has arisen from a close working relationship with Professor Marina Resta, to which I owe a huge debt of gratitude for her trust, support and encouragement during the past few years.

I am also deeply indebted with Professor Maria Elena De Giuli, for her enthusiasm and advice, constantly reminding me that greatness lies in the smallest detail.

Furthermore, this journey would not have been possible without the closeness, help and patience of all the beautiful people who surround me everyday. My deepest thank goes to my girlfriend Altea, my mum, dad and sister.

Lastly, a special mention goes to all my PhD fellows, the PhD coordinator Professor Anna Bottasso and the PhD officer Viviana Pastorino, for constant support along this path.

Life can only be understood backwards;
but it must be lived forwards.
Søren Kierkegaard

Declaration

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Marco Neffelli
January 2019

Abstract

The thesis contributes to the financial econometrics literature by improving the estimation of the covariance matrix among financial time series. To such aim, existing econometrics tools have been investigated and improved, while new ones have been introduced in the field. The main goal is to improve portfolio construction for financial hedging, asset allocation and interest rates risk management. The empirical applicability of the proposed innovations has been tested trough several case studies, involving real and simulated datasets.

The thesis is organised in three main chapters, each of those dealing with a specific financial challenge where the covariance matrix plays a central role. Chapter 2 tackles on the problem of hedging portfolios composed by energy commodities. Here, the underlying multivariate volatility among spot and futures securities is modelled with multivariate GARCH models. Under this specific framework, we propose two novel approaches to construct the covariance matrix among commodities, and hence the resulting long-short hedging portfolios. On the one hand, we propose to calculate the hedge ratio of each portfolio constituent to combine them later on in a unique hedged position. On the other hand, we propose to directly hedge the spot portfolio, incorporating in such way investor's risk and return preferences. Trough a comprehensive numerical case study, we assess the sensitivity of both approaches to volatility and correlation misspecification. Moreover, we empirically show how the two approaches should be implemented to hedge a crude oil portfolio.

Chapter 3 focuses on the covariance matrix estimation when the underlying data show non–Normality and High–Dimensionality. To this extent, we introduce a novel estimator for the covariance matrix and its inverse – the Minimum Regularised Covariance Determinant estimator (MRCD) – from chemistry and criminology into our field. The aim is twofold: first, we improve the estimation of the Global Minimum Variance Portfolio by exploiting the MRCD closed form solution for the covariance matrix inverse. Trough an extensive Monte Carlo simulation study we check the effectiveness of the proposed approach in comparison to the sample estimator. Furthermore, we take on an empirical case

study featuring five real investment universes characterised by different stylised facts and dimensions. Both simulation and empirical analysis clearly demonstrate the out–of–sample performance improvement while using the MRCD. Second, we turn our attention on modelling the relationships among interest rates, comparing five covariance matrix estimators. Here, we extract the principal components driving the yield curve volatility to give important insights on fixed income portfolio construction and risk management. An empirical application involving the US term structure illustrates the inferiority of the sample covariance matrix to deal with interest rates.

In chapter 4, we improve the shrinkage estimator for four risk-based portfolios. In particular, we focus on the target matrix, investigating six different estimators. By the mean of an extensive numerical example, we check the sensitivity of each risk-based portfolio to volatility and correlation misspecification in the target matrix. Furthermore, trough a comprehensive Monte Carlo experiment, we offer a comparative study of the target estimators, testing their ability in reproducing the *true* portfolio weights. Controlling for the dataset dimensionality and the shrinkage intensity, we find out that the Identity and Variance Identity target estimators are the best targets towards which to shrink, always holding good statistical properties.

Chapter 1

Introduction

The estimation of the covariance matrix plays a central role in portfolio construction and optimisation. [Markowitz, 1952] defined the covariance matrix among financial time series as a key input for his well–known portfolio building technique. In recent years, findings suggested that investment portfolios solely based on the covariance matrix yield better out–of–sample stability and performance [DeMiguel and Uppal, 2009]. Hence, the general tendency has been to put risk – and so the covariance matrix – at the very centre of the portfolio construction process. Seminal contributions as the Global Minimum Variance [Black and Litterman, 1992], the Maximum Diversification [Choueifaty and Coignard, 2008], the Equal–Risk–Contribution [Maillard et al., 2010] and more recently the Inverse Volatility portfolios [Leote et al., 2012] have clearly shown that on the one hand, this object conveys almost all the important information needed to allocate wealth. On the other hand, it carries a lower estimation error than the mean, leading to more robust asset allocation choices for investors.

From a statistical viewpoint, the *true* covariance matrix is an unfeasible object, given that we only observe a sample of the whole population data. Hence, the covariance matrix used in empirical applications is merely an estimate, suffering from different errors: estimation error from the inferential process; model error from employing a certain model with specific assumptions; curse of dimensionality because the number of parameters to estimate must be lower than the number of available observations. Ignoring these issues leads to severe inferential distortions, with pernicious consequences for investment portfolio allocations, especially when the covariance matrix is the only input. As largely documented in the literature (see e.g. [Ardia et al., 2017]), these distortions have striking implications in out–of–sample portfolio weights and out–of–sample portfolio performances.

In light of the above considerations, we focus our efforts on improving the estimation of the covariance matrix for portfolio applications. To do so, we build on existing econometric methodologies and propose the adoption of new ones. In particular, we contribute to the existing literature under different angles. Special emphasis have been given to issues in portfolio construction and optimisation. In this area, we investigate the estimation of both the covariance matrix and its inverse, considering several portfolio building rules for asset allocation purposes. Moreover, we also face issues related to portfolio hedging as well as interest rates modelling for risk management purposes. Henceforth, this work has been divided into three main chapters to accommodate the heterogeneous nature of the analysis.

Chapther 2 focuses on hedging energy commodity portfolios in the multivariate GARCH framework. To such aim, we take on the case when the underlying spot and futures process are driven by heteroskedastic volatility and correlation. We introduce and discuss two novel methodologies to construct the hedging long–short portfolio: the Block Approach, where each spot position is hedged in a stand–alone fashion to be aggregated to the hedged portfolio in a second step; and the Objective–Driven approach, where we account for investor's risk and return preferences by hedging the spot positions in a portfolio fashion. This chapter features a numerical illustration to assess the sensitivity of the hedge to both approaches, and an empirical case study where a crude oil portfolio is hedged.

Chapter 3 revolves around the introduction of the Minimum Regularised Covariance estimator (MRCD) into the financial econometrics literature. This is a robust estimator for the covariance matrix and its inverse, particularly suitable to work with non–Normal and high–dimensional datasets. The chapter is divided into two main sections: first, we try to exploit the convenient closed form solution of the MRCD covariance inverse to improve the out–of–sample performance of the Global Minimum Variance Portfolio. This is achieved trough a comprehensive comparison with the sample covariance matrix estimator, carried on with both simulated and real–world dataset. Second, we turn our attention to fixed income modelling for portfolio risk management purposes. In this case, we compare the MRCD against four alternative covariance matrix estimators with different statistical features. Through an empirical case study, we search for the principal components driving the volatility of the US term structure, giving a freshly new application of the MRCD within the fixed income framework.

Chapter 4 concentrates on improving the shrinkage estimator for four risk–based portfolios. The chapter compares several estimators for the shrinkage target matrix with the aim of reducing the impact of covariance misspecification on the resulting portfolio weights. Following the results in [Ardia et al., 2017], we first design a numerical illustration to assess the effect of misspecification in the shrinkage target matrix. In a second instance, the study features a comprehensive Monte Carlo exercise where we control for the dimensionality of the dataset as well as the shrinkage intensity parameter. This allows us to pinpoint which target matrix estimator leads to portfolio weights closer to the population ones.

Finally, chapter 5 concludes the thesis, highlighting possible futures research paths and extensions of this work.

1.1 Contribution of the thesis

The main contribution of this thesis revolves around improving the estimation of the covariance matrix and its inverse for asset allocation and risk management purposes under a portfolio framework. This is achieved by analysing and improving existing approaches as well as by introducing novel techniques from other research fields. Furthermore, a huge variety of asset classes and datasets has been employed along the thesis to verify the empirical soundness of the proposed solutions. The next subsections offer a detailed description of each contributions for the next three chapters.

1.1.1 Chapter 2

Chapter 2 deals with two alternative approaches to hedge a portfolio of crude oil commodities. The addressed research question is whether it is preferable combining the hedge ratio of each portfolio component to derive the portfolio hedge position – Block Approach – rather than directly hedging the portfolio position – Objective–Driven Approach – thus incorporating into the hedge the investor's objectives and preferences in terms of risk and return. We compare the approaches within a multivariate GARCH framework, hence considering the underlying multivariate volatility process driven by heteroskedasticity. Using several multivariate GARCH models for the covariance matrix (namely: Dynamic Conditional Correlation (DCC), Diagonal–BEKK, Diagonal Rotated BEKK and Diagonal Rotated DCC) and under two alternative hedging strategies (fixed and dynamically rebalanced), we devise an empirical case study to practically hedge a long spot position in both WTI and Brent with related futures contracts. Our research concludes that feeding the hedging procedure with the investor's risk-return preference lowers the hedge sensitivity to shifts in the portfolio composition. The Objective–Driven approach makes then possible to get better risk–return performances, especially out–of–sample, whose outcomes become more stable and less prone to negative downfalls. Conclusions are robust with respect to changes in both the variance estimator and in the hedging strategy.

Despite of the relevance of the issue, it has been scarcely examined in the existing literature: [Roberts et al., 2018] analysed the problem for currencies portfolios, without taking into account our multivariate volatility framework; while [Chang et al., 2011] discussed the effectiveness of several multivariate GARCH models for hedging one commodity per time. Therefore, there is enough room to claim that, to the best of the authors' knowledge, this work is a primer in introducing and discussing the issue of comparing the Block and Objective–Driven approaches in the MGARCH framework.

1.1.2 Chapter 3

Chapter 3 contributes to the existing literature by introducing a novel estimator for the covariance matrix and its inverse – the Minimum Regularised Covariance Determinant estimator (MRCD) – already applied in chemistry and criminology. In particular, we add value towards two main directions, since we implement the MRCD in two different financial econometrics' areas: asset allocation and interest rate risk management. First, we use the MRCD to limit portfolio weights misspecification within the Global Minimum Variance Portfolio (GMVP) framework. Estimating the inverse covariance matrix (precision matrix), in fact, is a key task that can generate estimation errors directly affecting the GMVP out–of–sample performance. The MRCD, on the other hand, is a direct estimator of the precision matrix, designed to deal with high–dimensional, non–Normal datasets likewise the financial ones. These features make of particular appeal using the MRCD instead of the sample estimator to build the covariance matrix in the portfolio optimisation problem. Our study includes an extensive Monte Carlo simulation analysis to check the effectiveness of the proposed approach in comparison to the sample estimator. Furthermore, the Monte Carlo analysis shows that applying the MRCD technique lowers the GMVP weights misspecification. Moreover, we take on an empirical case study featuring five real investment universes with different stylised facts and dimensions. Both the simulated and empirical applications clearly demonstrate that the out–of–sample performance of the GMVP benefits from the use of the MRCD estimator: results suggest a reduction in the portfolio turnover at no cost for the portfolio variance, and an increase in portfolio expected returns. Second, we compare various methodologies to estimate the covariance matrix in a fixed income portfolio. Adopting a statistical approach for the robust estimation of the covariance matrix, we compared the Shrinkage, the Nonlinear Shrinkage, the Minimum Covariance Determinant and the MRCD estimators against the sample covariance matrix, here employed as a benchmark. The comparison was run in an application aimed at individuating the principal components of the US term structure curve. The contribution of the work mainly resides in the fact that we give a freshly new application of the MRCD and the NS robust covariance estimators within the fixed income framework. Results confirm that, likewise financial portfolios, also fixed income portfolios can benefit of using robust statistical methodologies for the estimation of the covariance matrix.

1.1.3 Chapter 4

In Chapter 4, we tackle on the issue of minimising the misspecification from volatility and correlation in the estimation of risk–based portfolios. Even thought portfolio weights solely based on risk avoid estimation errors from the sample mean, they are still affected from the misspecification in the sample covariance matrix. To solve this problem, we shrink the covariance matrix towards the Identity, the Variance Identity, the Single-index model, the Common Covariance, the Constant Correlation, and the Exponential Weighted Moving Average target matrices. Therefore, our contribution resides in improving the shrinkage estimator specifically for the Minimum Variance, Inverse Volatility, Equal-Risk-Contribution, and Maximum Diversification portfolios. Using an extensive Monte Carlo simulation, we offer a comparative study of shrinkage target estimators, testing their ability in reproducing the *true* portfolio weights. We control for the dataset dimensionality and the shrinkage intensity in aforementioned portfolios. We find out that the Identity and Variance Identity have very good statistical properties, also being well conditioned in high-dimensional datasets. In addition, these two models are the best target towards which to shrink: they minimise the misspecification in risk–based portfolio weights, generating estimates very close to the population values. Overall, shrinking the sample covariance matrix helps to reduce weight misspecification, especially in the Minimum Variance and the Maximum Diversification portfolios. The Inverse Volatility and the Equal-Risk-Contribution portfolios are less sensitive to covariance misspecification and so benefit less from shrinkage.

1.2 Datasets, Programming Codes and Published Research

In order to ensure the full reproducibility of the obtained results, this section provides a summary, chapter by chapter, of the employed datasets as well as of the programming codes developed for this thesis. Both these elements are available in the GitHub page of the author at the following web address: `https://github.com/marconeffelli/PhDThesis`. In addition, this section offers a complete description of the published research linked with this manuscript.

In **Chapter 2**, we analysed WTI and Brent energy commodities. In particular, we used WTI and Brent spot prices freely available from the US Energy Information Administration (EIA[1]), while we used WTI and Brent futures prices obtained from Bloomberg database. Each series has a total of 1505 daily observations in the period 03 January 2012 – 02 January 2018.

A programming code written in MATLAB with related instructions to replicate the Block and the Objective–Driven approach is made available for download. In order to estimate multivariate GARCH models, the code is based upon Kevin Sheppard *MFE toolbox* [2]. The research papers *Commodities hedging with Multivariate GARCH models: an application to the energy market* and *Hedging crude oil portfolios: better using a block or an objective-driven approach?*, both co–authored with my PhD main supervisor, Marina Resta, are based on the results in this chapter and have been submitted for publication. The papers

[1] `www.eia.gov`.

[2] Available at `https://www.kevinsheppard.com/MFE_Toolbox`.

have been respectively presented at the 2018 International Conference on Energy Finance (ICEF – Beijing, April 03–04 2018) and at the 42nd annual meeting of the AMASES (Naples, September 13–15 2018).

In **Chapter 3**, with respect to the first contribution on the Global Minimum Variance portfolio, we analysed five investment universes. The first four are taken from Kenneth French's website [3]. They have a monthly frequency in the period 01 July 1926 – 01 August 2018. The fifth investment universe has been taken from Bloomberg database. It is composed by the prices of 300 stocks belonging to the S&P500 in the period 01 January 1996 – 01 June 2018.

Programming code written in R to simulate financial time series, calculate the Global Minimum Variance portfolio and all the investment strategies described along this chapter is made available at the aforementioned GitHub page of the author. The code for the MRCD estimator has been taken from the University of KU Leuven website [4].

The research paper *Portfolio Selection with Minimum Regularised Covariance Determinant*, co–authored with my PhD supervisors, Marina Resta and Maria Elena De Giuli, is based on the results in the first part of this chapter and has been submitted for publication. The paper has been presented at the XIX Quantitative Finance Workshop (University of Roma Tre, January 24–26 2018), at the 29th European Conference on Operational Research (EURO – Valencia, July 08–11 2018) and at the 42nd annual meeting of the AMASES (Naples, September 13–15 2018).

With respect to the second contribution on fixed income modelling, we analysed the US term structure of interest rates. Daily observations in the period 02 January 2014 – 08 September 2017 have been taken from Bloomberg database.

Programming code written in MATLAB for running the PCA analysis is made available. The MCD has been calculated using the MATLAB toolbox LIBRA [5]; the MRCD code shares the previously mentioned source, while the Nonlinear Shrinkage code has been taken from the CRAN R package *nlshrink* [6].

The Research paper *A comparison of Estimation Techniques for the Covariance Matrix in a Fixed–Income Framework*, co–authored with Marina Resta, is based on the results in the second part of this chapter and it has been published in *New Methods in Fixed Income Modelling*, pp. 99–115, Springer, Cham.

In **Chapter 4**, we used a simulated dataset. Programming code written in MATLAB to

[3] `http://mba.tuck.darthmouth.edu/pages/faculty/ken.french/data_library.html`

[4] `https://wis.kuleuven.be/stat/robust/Programs/MRCD`.

[5] `https://wis.kuleuven.be/stat/robust/LIBRA/LIBRA-home`.

[6] `https://cran.r-project.org/web/packages/nlshrink/nlshrink.pdf`.

calculate the four risk–based portfolios is made available at the GitHub page of the author.

The research paper *Target matrix estimators in Risk–based portfolios* is based on the results in this chapter and it has been published in *Risks*, 6(4), 1–20, Special Issue Computational Methods for Risk Management in Economics and Finance.

Chapter 2

Energy Portfolio Hedging in the MGARCH framework [1]

2.1 Introduction

This chapter provides an in–deep comparison and discussion of alternative hedging approaches for an energy commodities portfolio. The research question we are trying to address concerns the hedging procedure to restrain or reduce the risk of an investing position in energy commodities. Consider for instance an investor willing to allocate her wealth into two crude oil spot commodities, WTI and Brent, and to hedge the exposure by short–selling respective futures contracts. Under the Block Approach (BA), the investor calculates the Hedge Ratio (HR therein after) for each couple of spot and futures contracts (say: HR_{WTI} and HR_{BRENT}), at first, and in a second step she builds the hedged portfolio covering her position by a counterpart investment of $-HR_{WTI}$ and $-HR_{BRENT}$ in futures. This means that the overall hedging position of the portfolio (say: HR_{π}) is merely the aggregation of the HR of any component. Under the Objective–Driven Approach (ODA), on the other hand, the process described in previous rows is somewhat reverted: the HR is directly calculated on the spot portfolio so that the quote of WTI and BRENT to hedge depends on the weights given to the commodities in the portfolio optimisation procedure. The difference is not merely a matter of how aggregating the HRs, the ODA adds something further, because it makes possible to incorporate the investor risk–return preferences into the computation of the HR. This might potentially exercise a strong impact on the overall hedging decision.

Despite of the relevance of the issue, so far it has been examined only in a few cases: [Roberts et al., 2018] analyses the problem for currencies portfolios, while [Chang et al., 2011] discusses the effectiveness of the BA comparing several multivariate models for the portfolio variance;

[1] The research papers *Commodities hedging with Multivariate GARCH models: an application to the energy market* and *Hedging crude oil portfolios: better using a block or an object–driven approach?*, both co–authored with my PhD main supervisor, Marina Resta, are based on the results in this chapter and have been submitted for publication. The papers have been respectively presented at the 2018 International Conference on Energy Finance (ICEF - Beijing, April 03-04 2018) and at the 42$^{\text{nd}}$ annual meeting of the AMASES (Naples, September 13-15 2018)

however, to the best of the authors' knowledge the comparison between the BA and ODA has not been examined, at least for energy commodities. This clearly opens space enough for introducing and discussing the issue.

The experimental design carried on throughout the chapter is based on the following assumptions: (*i*) the investor hedges by shorting with futures contracts; (*ii*) she calculates the HR to reduce the variance of the hedged portfolio; (*iii*) to hedge the portfolio by risk, the investor can run either the BA or the ODA. With respect to the assumption labelled by (*iii*), in both cases it is necessary using a proper estimator for the portfolio variance. Stylized facts affecting crude oil time series, widely described in the specialised literature [Kang and Yoon, 2013] and later addressed in Section 5 of this chapter, too, suggest that the HR should be conditional and time–varying; this supports the use of an underlying conditional heteroskedastic process for the portfolio variance. We therefore selected four multivariate GARCH models, namely: the Dynamic Conditional Correlation (DCC) [Engle and Sheppard, 2001], the diagonal Baba, Engle, Kraft and Kroner (DBEKK) [Engle and Kroner, 1995], the diagonal rotated BEKK (RBEKK) and the diagonal rotated DCC (DRDCC) [Noureldin et al., 2014]. Finally, we tested the hedges derived from both the BA and ODA under two strategies: Fixed Hedging (FH) and Dynamically Rebalanced Hedging (DRH). The approaches were compared using WTI and BRENT daily data in the period: June 2012–June 2018, using earlier 1000 observations (June 2012–December 2015) as in–sample dataset, and the remaining 501 observations as out–of–sample, to test the hedging effectiveness.

Our results clearly highlight the superior performance of the ODA on the BA, supported by a number of evidences. First, the returns of the hedged portfolio are more stable and less prone to negative downfalls. This assertion is also supported by the behaviour of the Overall Returns Index (ORI), which computing the difference between the returns of the unhedged and hedged position provides by construction the same information as in the case of holding a long/short portfolio. Second, applying the ODA shrinks the variance of the hedged position, as resulting from the performance of two indicators which are generally employed to assess the goodness of the hedging: the Hedging Efficiency (HE) and the Expected Shortfall (ES). While the former assesses the goodness of the hedge in terms of variance reduction, the latter evaluates how much the risk is lowered; clearly, a desirable feature for the hedging is to get the highest HE and the closest to zero ES as possible. Throughout the chapter we are going to show that within our simulation framework the ODA makes possible to obtain higher HE and closer to zero ES values than in the BA, underpinning the lower risk associated to the former. Moreover, we highlight that results are stable and robust with respect to changes in the multivariate estimators of the portfolio variance and in the hedging strategies included in the chapter. Nevertheless, empirical results suggest that the DRBEKK is the best hedging model.

The chapter is organised as follows. Section 2.2 examines the literature on hedging, and offers a complete overview upon the room left to improve existing research. Section 2.3

introduces the notations used across this chapter, the derivation of the Minimum Variance Hedge Ratio and a detailed discussion on Multivariate GARCH models. Section 2.4 analyses the HR specifications for both the Block and the Objective–Driven approaches. Section 2.5 offers a numerical illustration to gauge the impact of volatility and covariance upon the two hedging approaches. Section 2.6 illustrates an empirical application to hedge Crude Oil portfolios. After a comprehensive data analysis, it establishes the general settings for our study, describing the implementation of the hedging approaches and related hedging strategies, to move then at how evaluating the hedge results. Section 2.7 concludes.

2.2 Literature Review

As described in [Chen et al., 2013], there is a large body of literature dealing with the Hedge Ratio (HR), variously focusing on its building components: the objective function, the assumption of static or dynamic basis risk, and the model for estimating the returns and the volatility of the financial exposure.

For what is concerning the objective functions, they can be mainly divided into two categories: risk–minimising and utility–maximising [Chen et al., 2013]. Within the former, the most common HR, as well as the first being described in the academic literature, is the Minimum Variance Hedge Ratio (MVHR) [Johnson, 1960]. The MVHR works by reducing the volatility of the hedged portfolio: the lower, the better. Risk-minimising HRs can be also estimated with risk proxies other than the variance: [Cheung et al., 1990] uses the Mean Extended–Gini coefficient (MEG), while [Chen et al., 2001] and [Turvey and Nayak, 2003] employ the Generalised Semi–Variance (GSV). On the other hand, the utility–maximising HRs maximise the investor's Risk to Reward Position (RtRP): an example is provided by the so–called Optimal Hedge Ratios (OHR) [Chen et al., 2013], a family of HRs that assumes various proxies for the RtRP: the Sharpe Ratio in [Howard and D'Antonio, 1984]; the expected utility in [Kroner and Sultan, 1993], the mean–MEG and the mean–GSV in [Shalit, 1995] and [Chen et al., 2001], respectively. Nevertheless, the MVHR is often preferred to the OHR: [Chen et al., 2001], for instance, points out that when futures prices behave as a pure–martingale process, the MVHR is consistent with the mean–variance approach; moreover, if spot and futures returns are jointly normally distributed, the MVHR would be also consistent with the expected utility maximisation principle.

Understanding the basis risk, i.e. the risk associated with imperfect hedging, is important, too, given its practical implications for investors. Assuming constant volatility through time (static basis risk), the naïve method [Johnson, 1960] short futures with the (1:1) ratio between spot and futures; moreover, by regressing spot on futures price changes [Benninga et al., 1984] demonstrated the empirical relevance of the OLS regression for hedging, and described the procedure as general and easy to apply. Indeed, [Yang and Allen, 2005] used an Error

Correction Model (ECM) for hedging the Australian futures market thus considering the eventuality of order one integration between spot and futures prices time series. Nevertheless, when the basis risk is time–varying (dynamic basis risk), the HR is easily estimated within the ARCH/GARCH framework. In this case, the HR itself becomes a time–varying object [Lence et al., 1993]: this implies that spot and futures prices are assumed being generated by different stochastic processes, as their behaviour differs through time, so that the unrealistic assumption that spot and futures prices move together is relaxed. Earlier applications of ARCH and GARCH models for building the HR have been discussed in [Cecchetti et al., 1988] and [Kroner and Sultan, 1993]. In particular, [Cecchetti et al., 1988] built an ARCH process to replicate the dynamics of the variance, trying to hedge a position in a 20–year Treasury Bond, while [Kroner and Sultan, 1993] hedged several currencies (namely: the British Pound, the Canadian Dollar and the German Mark) using a GARCH model and demonstrated the hedge ability of the so–obtained dynamic HR. The dynamic HR was tested on commodity markets by [Baillie and Myers, 1991] and [Moschini and Myers, 2002]: the former hedged six different non–energy commodities with a bivariate GARCH model, and found that the static HR is unsuitable, since the movements of the basis cannot be neglected over time; the latter derived similar conclusions with a BEKK (Baba-Engle-Kraft-Kroner) multivariate GARCH model [Engle and Kroner, 1995] to hedge a long position in corn. The appropriateness of univariate and multivariate GARCH models in hedging commodities was also shown in [Chan and Young, 2006], who successfully hedged an exposure in copper with a regime–switching GARCH model well working in the case of both bear and bull markets.

The superior capability of the GARCH and its extensions with respect to other hedging alternatives has been deeply investigated, too: [Laws and Thompson, 2005] found that the Exponentially Weighted Moving Average (EWMA) model [RiskMetrics, 1996] performs better than OLS and standard GARCH methodologies in a hedging application on stock indexes; an opposite conclusion, however, is drawn by [Choudhry, 2009], who compared the hedging effectiveness of three models (naïve, OLS and GARCH) for six stock indexes. Besides, [Floros and Vougas, 2004], hedging the Greek stock market, argued that GARCH performs better than OLS and Vector Error Correction Models (VECM). Actually, Multivariate GARCH (MGARCH) models represent the last frontier for hedging strategies involving the GARCH; [Bauwens et al., 2006] identified three approaches for constructing MGARCHs: (*i*) as direct generalisation of univariate GARCH models, likewise in VEC, BEKK and EWMA models; (*ii*) as linear combination of univariate GARCH, like in the orthogonal models; (*iii*) as nonlinear combination of univariate GARCH models, like in the case of Dynamic Conditional Correlation models (DCC). These models, in fact, are characterised by a low–parameterised structure, due to a two–steps approach for the estimation of univariate volatility components separately from the covariances [Bollerslev, 1990], [Engle and Sheppard, 2001].

MGARCH models have been widely used to replicate commodities and their relations [Silvennoinen and Teräsvirta, 2009]. There is a growing amount of literature that deals with their practical applications, at present oriented towards two main streams. In detail, [Park and Switzer, 1995] used a bivariate GARCH model to hedge stock indexes; the same model was employed by [Baillie and Myers, 1991] for hedging in six commodities markets, finding out evidence of non–stationarity in the HR; [Chang et al., 2010] computed the in–sample MVHR with several estimation techniques, and evaluated the hedge of a portfolio composed by oil and gasoline spot commodities together with the related futures contracts; [Lien et al., 2002] compared the hedging ability of OLS, Vector Auto–Regression (VAR), VECM and diagonal-VECM (DVECM) for the Australian equity market, and validates the assumption of a time–varying HR. A fruitful research vein concerns also the comparison among various MGARCH models and the OLS regression: [Yang and Allen, 2005] focused on the constant–correlation Vector GARCH and examined the out–of–sample forecast for ten financial instruments covering different asset classes such as currencies, commodities and equities; [Moon et al., 2009] analysed the in–sample and out–of–sample hedging performance in the Korean equity index and the respective futures for several MGARCH models, including BEKK, principal components GARCH, DVECM and CCC: as major finding, they highlight the greater suitability of those model for dynamic hedging. Besides, [Chang et al., 2011] tested the effectiveness of BEKK, CCC, DCC and Vector ARMA-GARCH models to hedge long positions in crude oil, pinpointing the notable efficiency of the DCC. Similarly, [Chang et al., 2013] applied MGARCHs to currencies, to define hedging strategies for the exchange rate of the US dollar against Euro, British Pound and Japanese Yen: four MVHRs calculated with diagonal BEKK, CCC, DCC and VARMA-Asymmetric GARCH (VARMA–AGARCH) models are compared for a two–assets portfolio including spot and correspondent futures contracts. Findings suggest the highest hedging efficiency of CCC and VARMA–AGARCH.

Within the above described framework, our work is supported by various instances. First, so far it has not still presented in the energy commodities market a systematic discussion about the issue under examination: whether it is better to hedge separately the components of the portfolio or the portfolio as whole at first, separating the contribution of each asset later, based on the weights of the optimisation procedure. Second, we merge the discussion within a multivariate portfolio variance framework. Third, we offer an extensive comparison based on various performance and risk indicators as well as on various hedging strategies.

2.3 Methodology

2.3.1 Nomenclature

In order to ensure a streamlined and easier understanding of this section, we preliminary summarise the nomenclatures in Table 2.1, recalling them when necessary. We then derive the analytic formulas for the MVHR to be applied both in the BA and ODA. Finally, we present and discuss the multivariate GARCH models selected for estimating the portfolio

Table 2.1: Variables in use throughout this chapter and their description

Row nr.	Variable	Description
1	k	number of spot assets
2	$n = 2k$	total number of assets, including spot and hedging instruments
3	σ^2_{hedged} $\left(\sigma^2_{t,hedged}\right)$	unconditional (conditional) hedged portfolio variance
4	$\sigma^2_{unhedged}$ $\left(\sigma^2_{t,unhedged}\right)$	unconditional (conditional) unhedged portfolio variance
5	$\mathbf{q}_t$	$k \times 1$ vector of hedging instruments cond. weights
6	$\mathbf{c}_t$	$k \times 1$ vector of cond.cov. between hedging instruments and the unhedged portfolio
7	G_t	$k \times k$ conditional Variance–Covariance matrix among hedging instruments
8	I_{t-1}	sigma field generated by the past information until time $t-1$
9	$\sigma^2_{hedging}$	unconditional hedging portfolio variance
10	$\sigma_{unhedged,hedging}$	unconditional covariance between unhedged and hedging instruments
11	H_t	$n \times n$ conditional Variance–Covariance matrix for the hedged portfolio
12	D_t	$n \times n$ diagonal matrix in DCC and RDCC models for the standard deviation
13	Y_t	conditional correlation matrix among the univariate returns in DCC and RDCC model
14	Q_t	matrix driving the correlation dynamics in DCC and RDCC models
15	C	upper triangular matrix of intercept equations in DBEKK and DRBEKK models
16	A	diagonal matrix of coefficients in in DBEKK and DRBEKK models
17	B	diagonal matrix of coefficients in in DBEKK and DRBEKK models
18	M	unconditional correlation eigenvectors matrix for the D-RDCC model
19	Λ	unconditional correlation eigenvalues matrix for the RCC model

variance.

2.3.2 Deriving the Minimum Variance Hedge Ratio

We derive the MVHR for portfolio whose Variance–Covariance (VC) matrix is conditional and time–varying. The derived formula can be used for hedging both in the BA and ODA cases.

Following [Johnson, 1960], consider an unhedged portfolio of k spot assets; we want to hedge its exposure by shorting an equal number k of related futures contracts. The resulting *hedged portfolio* is then formed by $n = 2k$ assets and the vector collecting the realisations of the returns for both the spot and futures contracts is:

$$\pi_{t,hedged}|I_{t-1} \sim N(\mu_{\mathbf{t}}|I_{t-1}, H_t|I_{t-1}), \tag{2.1}$$

where $H_t|I_{t-1}$ is the portfolio VC conditional matrix of dimension n, as in row 11 of Table 2.1, and $\mu_{\mathbf{t}}|I_{t-1}$ is the $n \times 1$ conditional mean vector. For sake of conciseness, since now on, all the quantities with the subscript t are meant being conditional to the sigma filter I_{t-1} that thereinafter will be omitted by the notations.

We assume to build the *hedged portfolio* by adding the hedging instruments (in our case: future contracts) to the *unhedged portfolio*, so that the *hedged portfolio variance* can be described as follows:

$$\sigma^2_{t,hedged} = \sigma^2_{t,unhedged} + \mathbf{q}'_t G_t \mathbf{q}_t + 2\mathbf{q}'_t \mathbf{c}_t, \tag{2.2}$$

where $\mathbf{q}_t$, $\mathbf{c}_t$ and G_t are as resulting from rows 5, 6 and 7 of Table 2.1. In this setting, $\mathbf{q}'_t G_t \mathbf{q}_t$ is the VC matrix of the hedging instruments with weights determined by the HR and $\mathbf{q}'_t \mathbf{c}_t$ is the covariance between the portfolio and the hedging instruments. By construction G_t is a block of H_t, as H_t is the VC matrix among all the assets, including futures contracts.

In order to get the MVHR, the variance of the hedged portfolio must be minimised with respect to $\mathbf{q}_t$; the First Order Conditions (FOC) from (2.2) is:

$$2G_t\mathbf{q}_t - 2\mathbf{c}_t = \mathbf{0}_k,$$

so that, as G_t must be positive definite, the Hedge Ratio is given by the optimal vector $\mathbf{q}^*_t$:

$$\mathbf{q}^*_t = -G_t^{-1}\mathbf{c}_t. \tag{2.3}$$

and the i–th component of $\mathbf{q}^*_t$ is:

$$\mathbf{q}^*_t(i) = \frac{\sigma_{uh}(i)}{\sigma^2_{hedging}(i)}, \tag{2.4}$$

where $\sigma_{uh}(i)$ is the covariance between the unhedged portfolio and the i–th hedging instrument and $\sigma^2_{hedging}(i)$ is the corresponding variance.

2.3.3 Multivariate GARCH: Technical Specifications

Within the above depicted framework, the matrix H_t plays an important role, as once it has been specified, since G_t is a block of it, the HR is fully determined too. For this reason, the conditional VC matrix H_t is our object of interest, and following [Engle and Sheppard, 2001], we modelled it by way of four multivariate GARCH models: the Dynamic Conditional Correlation (DCC), the diagonal Baba, Engle, Kraft and Kroner (DBEKK), the diagonal rotated BEKK (DRBEKK) and the diagonal rotated DCC (DRDCC).

The rationale in choosing the DCC and the Diagonal–BEKK models is that they have been already applied in [Chang et al., 2011], but limiting to the BA. Moreover, we consider the DRBEKK and the DRDCC because they are relatively new and promising tool and they can be easily implemented, as directly drawn from DCC and BEKK models with few additional computational efforts. [Noureldin et al., 2014].

In the next rows we provide a short description of above citepd MGARCH specifications of the matrix H_t: the rationale is providing the reader with a basic understanding of the efforts behind its estimation; references serve to address towards more in–depth analysis of any MGARCH model.

The BEKK [Engle and Kroner, 1995] is an extension of the VEC model [Bollerslev, 1988]. Formally, the BEKK equation for H_t is:

$$H_t = C'C + A'\epsilon_{t-1}\epsilon'_{t-1}A + B'H_{t-1}B, \tag{2.5}$$

where C, A and B are the matrices as specified at rows 15–17 in Table 2.1.

An alternative parametrisation of H_t is provided by the Dynamic Conditional Correlation (DCC) model [Engle and Sheppard, 2001]:

$$H_t = D_t Y_t D_t, \tag{2.6}$$

where D_t and Y_t are as in rows 12–13 of Table 2.1. The formulation of H_t given in (2.6) is particularly suitable for modelling purposes, as it allows to represent the univariate conditional variances in D_t in various ways, thus generating very different representations of the portfolio variance. In our case, for instance, the $h_{ii,t}$ elements of D_t are drawn from an univariate GJR(1,1) (Glosten, Jagannathan, and Runkle) process [Glosten et al., 1993] to include the evidence of greater volatility increases after negative shocks (leverage effect):

$$h_{ii,t} = \omega_i + \chi_i \varepsilon^2_{i,t-1} + \gamma_i \mathbf{1}_{\varepsilon^2_{i,t-1}<0} + \delta_i h_{ii,t-1} \tag{2.7}$$

where ω_i is the constant term associated to each element in the main diagonal of D_t, $\varepsilon_{i,t-1}$ is the error term at time $t-1$, $\mathbf{1}_{\varepsilon^2_{i,t-1}<0}$ is the indicator function which takes value 1 when $\epsilon_{i,t-1} < 0$ and 0 otherwise, χ_i weights the error terms at time $t-1$, γ_i is the leverage effect parameter and δ_i weights the variance at time $t-1$. This in turn makes possible to express Y_t in (2.6) through the following factorisation:

$$Y_t = \Psi_h Q_t^{DCC} \Psi_h, \tag{2.8}$$

where Ψ_h is the diagonal matrix with elements $1/\sqrt{ii,t}$, $i = 1, \ldots, n$ and the matrix Q_t given by:

$$Q_t = (1 - \chi - \gamma - \delta)\Omega + \theta_1 \hat{\nu}'_{t-1}\hat{\nu}_{t-1} + \theta_2 Q_{t-1}. \tag{2.9}$$

Here χ, δ and γ are as defined in the previous rows, $\hat{\nu}_{t-1}$ is the vector of standardised residuals at time $t-1$, Ω is the constant correlation matrix of the residuals and a and b are the scalars driving the correlation estimation process. In order to estimate it, we adopted the standard Quasi Maximum Likelihood technique with a three–stages procedure as in [Engle, 2002] that estimates the univariate variance components, the correlation intercepts and the dynamic parameters at three different steps.

The Diagonal–RBEKK and the Diagonal–RDCC models belong to the class of rotated ARCH (RARCH) models; as the names suggest, they are extensions of the BEKK and DCC models, respectively. The underlying idea is to apply a transformation (rotation) on raw returns that facilitates the adoption of covariance targeting technique [Noureldin et al., 2014], ending up in a less–parametrised estimation process.

Starting from the DRBEKK, the rotated returns are defined as:

$$e_t = M\Lambda^{-1/2}M'r_t = \bar{H}^{-1/2}r_t, \tag{2.10}$$

where M, Λ are as in rows 18–19 of Table 2.1. Moreover:

$$H_t = \bar{H}^{1/2}R_t\bar{H}^{1/2}. \tag{2.11}$$

with:

$$R_t = (I_n - AA' - BB') + Ae_{t-1}e'_{t-1}A' + BR_{t-1}B'. \tag{2.12}$$

Here $R_0 = I_n$, where I_n is the identity matrix of dimension n, A, B are diagonal matrices as at rows 16–17 in Table 2.1 and e_t are as previously specified. Clearly, estimating R_t is less demanding than estimating H_t, so that it is possible to build a simpler model, as this diagonal factorisation requires the estimation of only n elements.

As highlighted in [Noureldin et al., 2014], the RBEKK can be improved by applying the rotation to the returns after standardising them by the fitted conditional variances, thus obtaining the s.c. rotated DCC (RDCC). In this case, the general equation for H_t is the same as in (2.15) but: Y_t is now given by:

$$Y_t = (Q_t \odot I_n)^{-\frac{1}{2}}Q_t(Q_t \odot I_n)^{-\frac{1}{2}}.$$

Here the symbol $\odot$ indicates the Hadamard product and Q_t is obtained though the returns rotation procedure [Noureldin et al., 2014]:

$$Q_t = M\Lambda^{1/2}M'Q_t^*M'\Lambda^{1/2}M, \tag{2.13}$$

where M and Λ are as described in Table 2.1, rows 18 and 19, and Q_{t-1}^* is the variance of the standardised errors at time $t-1$. The expression in (2.13) is similar to (2.12) for the DRBEKK model, with the only difference that here we are modelling Q_t instead than R_t. Likewise in the case of R_t, this specification allows estimating a smaller number of parameters compared to the standard case because A and B are diagonal matrices by construction.

The DRDCC exploits the DCC parametrisation, proposing a robust alternative to the DCC itself, since it does not imply any bias in the estimation of the unconditional covariance, as the DCC does [Noureldin et al., 2014].

2.4 Better using a Block or an Objective–Driven Approach?

The difference between the Block and the Objective–Driven approaches originates when a portfolio manager faces the decision to hedge a portfolio of spot securities. In this chapter, we assume to have an initial investment universe composed by two spot securities, say S_1 and S_2. Each of those disposes of a related futures contract (F_1 and F_2) of whose it represents the underlying. The spot unhedged portfolio is given by:

$$\pi_{unhedged} = w_1 S_1 + w_2 S_2, \tag{2.14}$$

with w_1 and w_1 being the spot-1 and spot-2 weights. Spot weights always add up to 1. With hedging, a long-short portfolio featuring $k = 2$ spot securities and corresponding $k = 2$ futures contracts is built. Our question is: should the portfolio manager calculate the hedge ratios for both spot securities separately and then create a hedged portfolio (Block approach) merging the two hedged positions, or the first step should consist in hedging the two spot positions as a unique unhedged portfolio (ODA approach)? In the next rows we demonstrate how this decision leads to different hedge ratios. The main difference in choosing one of the proposed approach is the characterisation of the variance-covariance matrix among spot and futures securities, which is the main object of interest when hedging. In fact, in the case of the BA one just needs the covariances between spot and futures to work, while the ODA requires a richer set of information since it models the portfolio against the futures. We will highlight this along the section.

2.4.1 Block Approach

When hedging with the Block Approach, each asset is considered as a stand-alone block. This allows to calculate the hedge ratio separately for the i–th asset using (2.4):

$$HR_i^{BA} = \frac{cov(S_i, F_i)}{var(F_i)};$$

here $i = 1, 2$. To hedge the i–th spot position one must short a quantity of hedging instrument as described in HR_i^{BA}. With S_i describing a long position in the spot security i, we denote the outcome of the hedge for the i–th spot security as:

$$\pi_{hedged,i} = w_i S_i + (-w_i) HR_i^{BA} F_i,$$

where the futures position is weighted by the corresponding spot weight. Finally, the two hedged positions are merged to create the final hedged portfolio.

$$\pi_{hedged} = \pi_{unhedged} + (-w_1) HR_1^{BA} F_1 + (-w_2) HR_2^{BA} F_2. \tag{2.15}$$

More precisely, in an algorithmic fashion, the Block approach can be summarised in a few steps provided in the following rows.

Algorithm 2.1: Block Approach.

1 Allocate the initial wealth for spot assets as prescribed by a certain portfolio building rule.

2 Considering each spot/futures couple as stand-alone blocks, compute the hedge ratios.

3 Short futures contracts as prescribed by the HRs obtained in step 2.

4 Create the final hedged portfolio by merging the two hedged positions.

2.4.2 Objective–Driven Approach

Under the Objective–Driven approach, the portfolio manager builds the hedge directly on the spot portfolio. The hedge ratios are calculated using (2.3):

$$HR^{ODA} = [cov(\pi_{unhedged}, F_1), cov(\pi^{unhedged}, F_2)]G^{-1}.$$

HR^{ODA} is a vector composed by 2 elements. The hedged portfolio is easily built with the following:

$$\pi_{hedged} = \pi_{unhedged} + (-) HR_1^{ODA} F_1 + (-) HR_2^{ODA} F_2.$$

More precisely, in an algorithmic fashion, the Objective–Driven approach can be summarised in a few steps provided in the following rows.

Algorithm 2.2: Objective–Driven Approach.

1 Allocate the initial wealth for spot assets as prescribed by a certain portfolio building rule.

2 Compute the hedge ratios between the unhedged portfolio and the two futures instruments.

3 Short futures contracts as prescribed by the ODA HR obtained in the previous step to obtain the final hedged portfolio.

Remarks

The main takeaways of this section is to highlight the differences in the two hedging approaches. The point is that, under the above–described framework, the way in which the covariance among spot and futures is modelled changes dramatically. For example, in the case of the BA, the portfolio manager just needs the correlation among each spot/futures couple ($\rho(S_i, F_i)$) and the futures variances $var(F_i)$. On the contrary, hedging with the ODA approach means to calculate a richer set of information, since besides the correlations between the unhedged portfolio and the futures ($\rho(\pi_{unhedged}, F_1)$), the portfolio manager needs G, the variance–covariance matrix of futures. Assuming that the portfolio manager models the variances and covariances as multivariate GARCH processes, she needs to estimate two bivariate GARCH models under the BA, and one tri–variate GARCH model under the ODA. In conclusion, if the BA seems more practical to be implemented involving less parameterised models, on the other hand the ODA allows for a richer characterisation of variances and covariances, producing hedge ratios that contain more information.

2.5 A Numerical Illustration

To provide the reader with additional insights about the relevance of the addressed issue, we discuss a numerical example. Consider for instance a portfolio made by two assets, say S_1 and S_2, hedged by shorting with futures positions, F_1 and F_2. The covariance matrix of the hedged portfolio is therefore given by:

$$VC = \begin{bmatrix} \sigma_{S_1} & 0 & 0 & 0 \\ 0 & \sigma_{S_2} & 0 & 0 \\ 0 & 0 & \sigma_{F_1} & 0 \\ 0 & 0 & 0 & \sigma_{F_2} \end{bmatrix} \begin{bmatrix} 1 & \rho_{(S_1,S_2)} & \rho_{(S_1,F_1)} & \rho_{(S_1,F_2)} \\ \rho_{(S_1,S_2)} & 1 & \rho_{(S_2,F_1)} & \rho_{(S_2,F_2)} \\ \rho_{(S_1,F_1)} & \rho_{(S_2,F_1)} & 1 & \rho_{(F_1,F_2)} \\ \rho_{(S_1,F_2)} & \rho_{(S_2,F_2)} & \rho_{(F_1,F_2)} & 1 \end{bmatrix} \begin{bmatrix} \sigma_{S_1} & 0 & 0 & 0 \\ 0 & \sigma_{S_2} & 0 & 0 \\ 0 & 0 & \sigma_{F_1} & 0 \\ 0 & 0 & 0 & \sigma_{F_2} \end{bmatrix}$$

with $VC' = VC$, i.e, the matrix has simmetric entries with respect the main diagonal. This covariance matrix has entries that resemble the behaviour of WTI and BRENT spot and futures contracts during the period 2012–2018; therefore, $(\sigma_{S_1}, \sigma_{S_2}, \sigma_{F_1}, \sigma_{F_2}) = (0.02, 0.02, 0.02, 0.02)$ and $(\rho_{(S_1,S_2)}, \rho_{(S_1,F_1)}, \rho_{(S_1,F_2)}, \rho_{(S_2,F_1)}, \rho_{(S_2,F_2)}, \rho_{(F_1,F_2)}) = (0.6, 0.9, 0.8, 0.6, 0.7, 0.9)$.

We assume to monitor the behaviour of the HR when, varying the weights in the range [0,1], all is maintained unchanged except for the volatility (shifted in the range [0.01,0.05]) or the correlation (shifted in the range [0.5,1]), which are shifted one per time as shown in Table 2.2.

[Table 2.2 about here.]

The organisation of Table 2.2 mimics that of Figure 2.1 collecting the the results, organised into a 4×11 matrix of plots. The Table, in fact, reports the volatility and correlation whose value is change, the represented security, and the position (row and column) in which reading the BA/ODA cases. In a similar way, in Figure 2.1 from top to bottom, rows 1–2 pertain the BA, while rows 3–4 deal with the ODA. Plots are intended to be examined ad discussed by pairs: red–painted graphs always represent the evolution of the HR for the first security of the pair; green–coloured plots do the same but for the HR of the second security. Finally, columns 2 to 11 illustrate the cases of the above Table, while the first column shows the initial HRs, i.e. those derived by assuming the structure provided by (2.5) for the covariance matrix.

[Figure 2.1 about here.]

Starting from Column 1 in Figure 2.1, the starting HR depends on the weights assigned to the security in a proportional fashion: the greater the weight given to the first security, the highest the HR is (see the red lined in rows 1 and 3 of Col. 1); obviously the HR of the second security behaves exactly on the opposite way. At the first glance, the starting situation appears being the same for both the BA and the ODA; however, a deepest investigation of the plots suggests the existence of a small difference, as for small weights value the HR of the first security in the ODA lies below the zero line (blue–dashed), while for the BA it does not. Moving to the second column, we observe the effect of varying the covariance between S_1 and F_1: in this case, likewise in the other ones we assumed changes in the range $[0.5, 1]$: with the dashed–blu line we indicate the zero line, while the original position of the HR, as already discussed in Column 1, is highlighted by a dotted black line. We observe that the overall effect of changing σ_{S_1,F_1} is missed both by the BA and the ODA: as the weight of the first security increases, the related HR does the same, but the original behaviour of the HR (black–dotted line) works now as an upper bound for the HR values. Moreover, the increase is with a lower emphasis than before (i.e. as observed in Column 1, using the initial VC matrix): this is well represented by the red cone that now shows how the HR evolves (Column 2, rows 1 and 3). The effect is slightly enhanced in the ODA where it is evident not only for the first security (as in the BA), but also for the second one. Similar remarks, in a specular fashion, can be spent for the plots in Column 3, replacing S_1 by S_2 and F_1 by F_2. Turning to Column 4 to examine the effects in changing the covariance between S_1 and S_2, we observe that, as reasonably expected, this variation does not influence the HR that, in fact, maintains at the same levels as discussed in the original situation in Column 1. The situation is sensitively different when moving to the results in Column 5: while varying the covariance between F_1 and F_2 does not affect the HR in the BA, the same does not apply in the ODA where the HR values seem rolling up around the original levels represented by the black–dotted line. A similar trend is repeated when the changes operate between S_1 and F_2 (and specularly between S_2 and F_1).

Overall, we can then conclude that the BA is much less sensitive than the ODA to fluctuations in the values of the VC matrix.

2.6 Case Study – Hedging Crude Oil Portfolios

Consider an investor aimed at hedging a portfolio composed by more than one spot energy commodity. The problem can be taken on with two alternative approaches: (i) each commodity is considered as a stand–alone asset, and hence the hedge is calculated according to the BA, by weighting the hedge ratio of each portfolio component; (ii) the investor directly hedges the overall portfolio position (ODA), managing the hedge for each asset in a second time, depending on the portfolio weights calculated during the optimisation procedure.

This case study is focused on understanding which approach works at best, since there is a trade–off between hedging with these approaches. The main advantage of adopting the ODA, in fact, is in the embedding of more information: all the covariances among different futures contracts are considered, at the cost of estimating a richer and then more parametrised multivariate GARCH model. On the other hand, the BA uses less information, but is less demanding in the estimation of the parameters for the multivariate model.

We consider an investor who wants to equally allocate her wealth between WTI and Brent spot commodities and to hedge the exposure with related futures.

In order to provide a comprehensive discussion of the problem, we start with some exploratory data analysis, at first, including both the formal description of the involved securities, and the statistical characterisation of the dataset; second, we describe the hedging procedure, and finally we discuss the results.

2.6.1 Data Analysis

Table 2.3 illustrates basic features of observed spot and futures contracts, including also the abbreviations employed thorough the chapter. Selected securities are all listed in American

US Dollar; WTI is negotiated on the Intercontinental Exchange (ICE), while Brent is traded on the New York Mercantile Exchange (NYMEX).

[Table 2.3 about here.]

We extracted daily closing prices from the US Energy Information Administration (EIA) [2] for spot instruments, and from Bloomberg[3] for futures thus creating a sample of 1505 daily observations per time series, in the period: 03 January 2012 - 02 January 2018. The starting date of the investment (and hence of the associated hedge) is: 31 December 2015. The period January 2012 - December 2015 (almost four years, for an overall number of 1004 daily observations) was employed for estimating the multivariate GARCH models parameters. The period: January 2016 - January 2018 (501 daily observations) served as out-of-sample set for testing the HRs.

[Figure 2.2 about here.]

The selected time frame is characterised by a sharp decrease in prices during the 2014–2015 biennium due to a combination of severe over–capacity and weak demand, which forced oil prices to their lowest levels since the financial crisis. Moreover, at the end of 2016 the joint action of "Trumplflation" and the decision of non–Opec producers to cut supply generated a period of boom and bursts in the price levels. These effects are clearly observable in both spot and futures prices and returns time series as illustrated in Figure 2.2, displaying frequent shifts in prices and volatility clusters in returns.

[Table 2.4 about here.]

Table 2.4 gives a comprehensive overview about the statistical properties of spot and futures prices and returns. In detail, we provide basic statistics, as well as results for the Jarque–Bera (JB thereinafter) Normality test, the unconditional correlation between spot and futures (Unc. Corr. in the spot column only), the Ljung–Box (LB) test for serial correlation and the Augmented Dickey–Fuller test for stationarity. Test statistics and p–values (in rounded brackets) are provided: significance levels are labeled by $***$, $**$ and $*$ for 99%, 95% and 90% confidence interval, respectively. Table 2.4 Panel A displays main descriptive statistics for the data over the whole period. We observe the presence of high standard deviation; excess kurtosis and positive skewness, while the negative returns averages are the signal of the downtrend features of the period. The unconditional correlation is close to one, between WTI spot and futures (0.9668), while Brent is less correlated although correlation maintains higher (0.7168). The right asymmetry and the positive kurtosis pinpoint the departures from Normality of the empirical distributions, as stated by the JB test, which always rejects the null hypothesis of Normality at 99% significance level. Moreover, the LB test, performed on the demeaned returns, highlights the presence of serial correlation and ARCH effects in

[2] www.eia.gov

[3] www.bloomberg.com

all the time series both at 5 and 15 lags. Lastly, Table 2.4 Panel B always highlights the non–stationary of prices, while returns are always stationary at 99% significance level.

2.6.2 Implementing the hedge

Hedging Strategies

We considered two strategies: Fixed Hedging (FH) and Dynamically Rebalanced Hedging (DRH).

In a formal way, assume that $[0, T1]$ is the in–sample interval, where $T1$ is the date at which the investor begins to hedge; denote by $[T1 + 1, T2]$ the time frame representing the out–of–sample until the end of the hedging period ($T2$). The FH is the strategy that estimates the HR (HR0) just once, using data observed in $[0, T1]$ and then maintains unchanged long/short position until reaching $T2$. On the other hand, with the DRH strategy the hedged portfolio is dynamically rebalanced at each instant t, $T1 < t \leq T2$. The HR is estimated at every instant t: we apply a rolling window that spans over the data discarding at each time the oldest observation and adding a new one, thus computing and overall number of (T_2-T_1+1) HR estimations.

Evaluating the Hedge

We selected three indicators to assess the hedge performance: the Overall Returns (OR), the Hedging Efficiency (HE) and the Expected Shortfall (ES). The motivation is that they offer a full coverage for in–sample and out–of–sample performances of the hedge, focusing on the returns of the hedged portfolio (the OR), its risk (the ES) and the total variance reduction observed by switching from the unhedged to the hedged position (the HE). Moreover, they can be calculated in the same way for both the BA and the ODA, making easier their comparison. Finally, as these indicators are calculated for each t ($t = T1 + 1, \ldots, T2$), we are able to report some descriptive statistics (minimum, maximum, mean) for a clearer and more readable description.

Formally, the OR is the difference between the returns of the unhedged position and those of the hedging position, thus representing the returns of a long–short portfolio:

$$OR(t) = \Pi_{unhedged}(t) - \Pi_{hedging}(t). \tag{2.16}$$

where $T1 < t \leq T2$, $\Pi_{unhedged}(t)$ and $\Pi_{hedged}(t)$ are the unhedged and hedging portfolio returns, respectively. For the OR indicator only, we provided information about the standard deviation and the percentage of negative returns: when computing the OR from a hedged portfolio, investors are evaluating how much of the portfolio value the hedge is preserved. In this light, the OR standard deviation and the percentage of negative returns are proxies for the hedge stability in the out–of–sample.

For what it concerns the remaining indicators, the HE and the ES are both based on the variance of the hedged portfolio. To make them comparable across different (BA and ODA) working frameworks, they have been calculated using the unconditional hedged portfolio variance, σ^2_{hedged}[4].

A formal definition of the HE is given by [Ederington, 1979], who evaluates the hedging performance by measuring the variance reduction of the hedged portfolio in percentage terms:

$$HE = \frac{\sigma^2_{hedged} - \sigma^2_{unhedged}}{\sigma^2_{unhedged}} \tag{2.17}$$

where σ^2_{hedged} and $\sigma^2_{unhedged}$ are as defined at rows 9 and 10 in Table 2.1. Higher HE values correspond to better hedging performances. The HE is a good indicator for the MVHR, since the less the variance of the hedged portfolio, the better the performance of the hedging strategy.

The Expected Shortfall is given by:

$$ES^{P\&L}_{\alpha} = \Pi(t)\frac{\Phi(z_{1-\alpha})}{1-\alpha}\sqrt{\sigma^2_{hedged}T}. \tag{2.18}$$

Here $\Pi(t)$, α, $\Phi(z_{1-\alpha})$ are as in rows 13, 10 and 14 of Table 2.1; finally, T is the time frame length. This indicator gives information about the stability of the hedge to unexpected downfalls.

2.6.3 Volatility Estimation

We are now ready to apply both static and dynamic hedging strategies for the BA and ODA multivariate approach. Possible combinations are illustrated in Table 2.4; labels in the table are those employed for referring to each strategy.

[Table 2.5 about here.]

For each combinations, HRs are calculated with all the proposed MGARCH models, for an overall number of sixteen trials for each energy commodity.

Provided both the large econometric literature dealing with crude oil commodities and the results of the statistical analysis in Table 2.5, we decided to filter the data with an

[4]This is supported as our approach focuses on the outcome of hedged portfolio only, mimicking the interest of an investor.

ARMA(1,1)–GARCH(1,1); this is inline with [Chang et al., 2011] that built bivariate HRs on the same commodities adopted in our chapter. Filtering results are reported in Appendix 2.7.

[Table 2.6 about here.]

[Table 2.7 about here.]

Tables 2.6 and 2.7 display the estimates of the four MGARCH models in the case of both the BA and ODA, respectively. To ease the readability of the results, we do not include in the body of the chapter MGARCH estimates under the dynamically rebalanced hedging strategy. However, the results are available in the 2.7.

The MGARCH models have been always estimated in their (1,1) specification. As explained in the previous sections, the core of the chapter relies in discussing the results, because the estimates sensitively vary when moving from the BA to the ODA. A possible explanation is that the BA forces the investor to estimate as many GARCH models as the number of spot/futures couples involved in building the hedged portfolio. In our case, since we have two spot commodities and two related futures contracts, this means estimating two bi–variate GARCH models per pairs, as highlighted in Table 2.6. On the other hand, the ODA leads to estimate only a multivariate GARCH model: in the examined case, for instance, we must estimate a tri–variate GARCH, since the VC matrix H_t includes two futures and the unhedged portfolio considered as a financial asset. In general, with a portfolio made by k securities the BA requires restimating k bi–variate MGARCHs whereas the ODA demands only the estimation of a $k+1$–parameter GARCH.

2.6.4 Results

We start discussing the results in Table 2.8, displaying the HRs obtained for the BA and he ODA in both the FH and DRH cases. We preliminary highlight that while the FH strategy returns only one value estimated on the in–sample data, on the other hand, the HRs derived under the DRH strategy are an average over the $T2 - T1 + 1$ HRs estimates with the walk–forward approach described in Section 2.6.2. Clearly, in the BA the HRs are not affected by portfolio selection process, as on the other hand, it happens in the ODA. To make possible the comparison of the results derived from different approaches we rescaled the HRs computed in the ODA by the weights of the unhedged portfolio.

[Table 2.8 about here.]

In general, the HRs obtained in the BA are smaller than those in the ODA, i.e. the investor choosing to hedge with the BA should short a lower amount of futures.

With respect to the FH strategy, WTI hedges are close to one in the BA, and greater than one in the ODA. With the exception of DBEKK, all the MGARCHs show similar behavioural patterns, as moving from the BA towards the ODA the HRs tend to increase. Similar considerations applies to the Brent. This behaviour is replicated also in the case of the DRH strategies.

[Table 2.9 about here.]

Table 2.9 illustrates the hedging results in all examined cases [5], evaluated by way of the indicators introduced in SubSec. 2.6.2. Starting from overall returns (OR) for the FH strategy, we examine the standard deviation ("std" in Table 2.8) and the percentage of time being negative ("down" in Table 2.9). For both the statistics, the general rule is: the lower, the better. With respect to the standard deviation, results of BA and ODA are almost similar: while the DCC and the DRDCC estimates increase moving from the BA towards the ODA, the BEKK models do the opposite. In detail, the DBEKK has the lowest standard deviation. On the other hand, the percentage of time being negative is always smaller for ODA HRs. Moreover, models in the ODA gets always the lowest minima, and the mean maintains closer to zero for the soft approach and slightly negative in other cases. To summarise, the OR highlights that the ODA performs better than the BA, at least during the examined period.

Moving to the remaining indicators (HE and ES), the ODA seems enhancing the hedge in all the cases: the variance is by far lower than in the BA. In detail, the average HE is roughly 0.6 in the BA, and 0.85 for the ODA. This tendency is even more evident for the ES: the average HE is close to -0.7 and -0.3 for the BA and ODA, respectively. Best results are associated to DCC GARCH model for the HE and to the DRBEKK for the ES.

The conclusions drawn for the FH strategy also apply to the DRH case.

Overall, our results pinpoint that HRs generated under the ODA perform better out–of–sample. With respect to the MGARCH models employed to estimate the VC matrix, the results highlight the robustness of the methodology under different HR specifications and alternative hedging strategies. The DBEKK works at the very best for the BA, while DRBEKK obtains best performances at all in the ODA.

2.7 Conclusion

This chapter investigates the hedging performance of two alternative hedging approaches, namely block and object–driven hedge ratio (HR). The research question is whether it is preferable combining the HR of each portfolio component to derive the portfolio hedge position (Block Approach – BA) rather than incorporating investor objectives and preferences in terms of risk and return (Objective–Driven Approach – ODA), i.e. by directly hedging the

[5] For both the BA and the ODA, for the four volatility estimators and for the two hedging strategies (fixed and dynamically rebalances).

portfolio position. To such aim, we evaluated the Minimum Variance Hedge Ratio, assume to hedge spot commodities by shorting related futures contracts. Moreover, we assumed that the hedged portfolio was driven by conditional and time–varying multivariate volatility; this allowed us to employ multivariate GARCH models. To assure the robustness of our analysis with respect to changes in the volatility estimator for the HR and to the hedging strategy, we employed four multivariate GARCH models, namely: DCC, diagonal BEKK, diagonal RBEKK and diagonal RDCC and two hedging strategies, fixed hedging (FH) and dynamically rebalanced hedging (DRH). The performance of the two hedging approaches was assessed through an empirical case study, in which the investor faces the problem of equally allocating her wealth between spot WTI and Brent crude oil commodities and hedging the spot exposure by shorting with futures. The hedge was then evaluated considering the returns and risk dimension of the hedged portfolio, both in–sample and out–of–sample. To this extent, we examined the hedge with three indicators: the overall returns, the hedging efficiency and the expected shortfall. Our findings highlight the superior hedging ability of the ODA on the BA, at least when dealing with energy commodities. This is not only confirmed by the overall returns, which are more stable and less prone to negative downfalls, and but also considering the risk of the hedged positions. The hedged portfolio variance, in fact, is lower for the ODA HR, getting an higher hedging efficiency and a lower expected shortfall. Results are stable with respect to alternative multivariate GARCH models and to various hedging strategies.

Table 2.2: Changes in the volatilities and correlations between the securities in our toy example.

Element ID	Security	Row	Column	Type
σ_{S_1}	First/Second	First/Second	Second	BA
σ_{S_1}	First/Second	Third/Fourth	Second	ODA
σ_{S_2}	First/Second	First/Second	Third	BA
σ_{S_2}	First/Second	Third/Fourth	Third	ODA
σ_{F_1}	First/Second	First/Second	Fourth	BA
σ_{F_1}	First/Second	Third/Fourth	Fourth	ODA
σ_{F_2}	First/Second	First/Second	Fifth	BA
σ_{F_2}	First/Second	Third/Fourth	Fifth	ODA
ρ_{S_1,F_1}	First/Second	First/Second	Sixth	BA
ρ_{S_1,F_1}	First/Second	Third/Fourth	Sixth	ODA
ρ_{S_2,F_2}	First/Second	First/Second	Seventh	BA
ρ_{S_2,F_2}	First/Second	Third/Fourth	Seventh	ODA
ρ_{S_1,S_2}	First/Second	First/Second	Eighth	BA
ρ_{S_1,S_2}	First/Second	Third/Fourth	Eighth	ODA
ρ_{F_1,F_2}	First/Second	First/Second	Ninth	BA
ρ_{F_1,F_2}	First/Second	Third/Fourth	Ninth	ODA
ρ_{S_1,F_2}	First/Second	First/Second	Tenth	BA
ρ_{S_1,F_2}	First/Second	Third/Fourth	Tenth	ODA
ρ_{S_2,F_1}	First/Second	First/Second	Eleventh	BA
ρ_{S_2,F_1}	First/Second	Third/Fourth	Eleventh	ODA

Table 2.3: Spot and futures contracts employed in our study. For each asset we provide the abbreviation, the reference ticker and the contract type (spot or derivative).

Variable	**Abbreviation**	**Ticker**	**Type**
WTI crude oil	WTI	RWTC	Spot Price FOB
Brent crude	BR	RBRTE	Spot Price FOB
FU - WTI crude oil	FU–WTI	CL1:COM	Futures
FU - Brent crude	FU–BR	CO1:COM	Futures

Table 2.4: Descriptive statistics and Dickey–Fuller test for both spot and futures.

	WTI		**Brent**	
	Spot	**Futures**	**Spot**	**Futures**
Panel A: Descriptive statistics				
mean	-0.0003	-0.0003	-0.0003	-0.0003
std	0.0211	0.0207	0.0192	0.0195
skewness	0.233	0.1891	0.3463	0.1776
kurtosys	6.37	6.2499	6.4226	6.4645
Unc.Corr.	0.9668		0.7169	
	(0.0000)***		(0.0000)***	
JB	725.3249	670.8727	764.2037	760.1193
	(0.0010)***	(0.0010)***	(0.0010)***	(0.0010)***
LB(5)	10.5377	17.1355	8.7225	20.3191
	(0.0613)*	(0.0042)**	(0.1206)*	(0.0010)***
LB(15)	21.9612	23.8273	26.9207	32.9208
	(0.1088)	(0.0680)*	(0.0293)**	(0.0048)***
Panel B: Dickey–Fuller test				
Prices	-1.2716	-1.2848	-1.2862	-1.2874
	(0.1881)	(0.184)	(0.1836)	(0.1832)
Returns	-41.3009	-42.167	-36.646	-42.1311
	(0.001)***	(0.001)***	(0.001)***	(0.001)***

Table 2.5: Hedging Implementation: summary.

	Static Hedging	**Dynamically Rebalanced Hedging**
BA	FH	DRH
ODA	FH	DRH

Table 2.6: MGARCH estimation results for the Block Approach under the FH strategy; p–values are reported in round brackets.

	DCC_W	DCC_B	$DBEKK_W$	$DBEKK_B$	$DRBEKK_W$	$DRBEKK_B$	$DRDCC_W$	$DRDCC_B$
θ_1	0.1223	0.0211						
	(0.0606)	(0.0235)						
θ_2	0.8494	0.7835						
	(0.1327)	(0.3270)						
C			0.5258	0.1776				
			(0.0644)	(0.0439)				
			0.5030	0.1610				
			(0.0703)	(0.0246)				
			0.0593	0.0975				
			(0.0154)	(0.0234)				
A			0.4791	-0.0676				
			(0.0787)	(0.0367)				
			0.4638	0.0435				
			(0.0794)	(0.0306)				
B			0.7859	0.9820				
			(0.0484)	(0.0094)				
			0.8022	0.9814				
			(0.0501)	(0.0055)				
α_{11}					0.3158	0.0747	0.3752	-0.1310
					(0.0728)	(0.0585)	(0.0426)	(0.0684)
α_{22}					0.2896	0.0524	0.3139	-0.1638
					(0.0646)	(0.1379)	(0.0216)	(0.2390)
β_{11}					0.9298	0.9851	0.9269	0.8495
					(0.0328)	(0.0142)	(0.2180)	(0.1342)
β_{22}					0.9397	0.9830	0.9158	0.9314
					(0.0323)	(0.0081)	(0.1213)	(0.0141)

Table 2.7: MGARCH estimation results for the Objective–Driven Approach under the FH strategy; p–values are reported in round brackets.

	DCC	DBEKK	DRBEKK	DRDCC
θ_1	0.0089			
	(0.0199)			
θ_2	0.9695			
	(0.0318)			
C		0.1986		
		(0.0456)		
		0.1796		
		(0.0430)		
		0.2048		
		(0.0498)		
		0.0842		
		(0.0198)		
		0.0704		
		(0.0211)		
		0.0449		
		(0.0335)		
A		0.0832		
		(0.0313)		
		0.0729		
		(0.0272)		
		0.1237		
		(0.0396)		
B		0.9755		
		(0.0107)		
		0.9764		
		(0.0106)		
		0.9662		
		(0.0150)		
$\max\alpha_{ii}$			0.2917	0.3509
			(0.0865)	(0.0624)
$\min\alpha_{ii}$			0.0063	0.0353
			(0.1074)	(0.1560)
$\max\beta_{ii}$			0.9795	0.9672
			(0.0106)	(0.0732)
$\min\beta_{ii}$			0.9103	0.7795
			(0.0544)	(0.1465)
$\max\alpha_{ii}+\beta_{ii}$			0.9166	0.8149
$\min\alpha_{ii}+\beta_{ii}$			1.2712	1.3181

Table 2.8: Hedge Ratios obtained from BA and ODA multivariate approaches varying the hedging strategies (Fixed Hedging vs Dynamically Rebalanced Hedging) and the MGARCH models.

		Fixed Hedging				**Dynamically Rebalanced Hedging**			
		DCC	DBEKK	DRBEKK	DRDCC	DCC	DBEKK	DRBEKK	DRDCC
BA	WTI	0.9595	1.0449	0.9987	0.9635	0.9719	0.9755	0.9643	0.9735
	BRENT	0.7093	0.7222	0.7113	0.7092	0.7251	0.7320	0.7262	0.7274
ODA	WTI	1.1504	0.9451	1.0211	1.0656	1.0681	1.0898	1.1959	1.1180
	BRENT	0.7207	0.9402	0.8947	0.8012	0.7970	0.7813	0.6859	0.7580

Table 2.9: Hedging results in the BA and ODA cases, varying the hedging strategies and the MGARCH models

Hedge Type	Fixed Hedging				Dynamically Rebalanced Hedging			
MGARCH	DCC	DBEKK	DRBEKK	DRDCC	DCC	DBEKK	DRBEKK	DRDCC
OR–BA								
min	-0.0467	-0.0457	-0.0463	-0.0466	-0.0462	-0.0459	-0.0460	-0.0462
max	0.0590	0.0561	0.0578	0.0589	0.0604	0.0576	0.0583	0.0601
mean	0.0000	-0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
std	0.0122	0.0122	0.0122	0.0122	0.0122	0.0121	0.0121	0.0122
down	0.5110	0.5030	0.5110	0.5110	0.5070	0.5130	0.5130	0.5090
OR–ODA								
min	-0.0477	-0.0452	-0.0447	-0.0449	-0.0573	-0.0558	-0.0513	-0.0527
max	0.0583	0.0576	0.0586	0.0578	0.0605	0.0597	0.0585	0.0598
mean	-0.0001	-0.0001	-0.0002	-0.0001	-0.0002	-0.0001	-0.0001	-0.0002
std	0.0124	0.0113	0.0116	0.0120	0.0122	0.0122	0.0126	0.0123
down	0.4950	0.5050	0.5010	0.4950	0.5030	0.4990	0.4950	0.4950
HE–BA								
min	0.5638	0.5514	0.5583	0.5633	0.5596	0.5605	0.5628	0.5612
max	0.6493	0.6413	0.6455	0.6489	0.6566	0.6518	0.6515	0.6569
mean	0.6110	0.6010	0.6065	0.6105	0.6111	0.6117	0.6124	0.6112
HE–ODA								
min	0.8437	0.8412	0.8423	0.8432	0.8079	0.8306	0.8408	0.8023
max	0.8602	0.8557	0.8582	0.8590	0.8598	0.8605	0.8604	0.8611
mean	0.8540	0.8504	0.8523	0.8531	0.8495	0.8512	0.8534	0.8517
ES–BA								
min	-0.8433	-0.6866	-0.7784	-0.8374	-1.0566	-1.0449	-1.0364	-1.0760
max	-0.7667	-0.6070	-0.7007	-0.7607	-0.6254	-0.5417	-0.5441	-0.6251
mean	-0.8039	-0.6470	-0.7391	-0.7980	-0.7599	-0.7439	-0.7705	-0.7537
ES–ODA								
min	-0.4266	-0.4077	-0.3510	-0.4389	-0.5569	-0.5584	-0.5426	-0.6819
max	-0.3570	-0.3352	-0.2770	-0.3688	-0.2760	-0.2054	-0.1199	-0.1902
mean	-0.3896	-0.3690	-0.3117	-0.4012	-0.3950	-0.3865	-0.3659	-0.3774

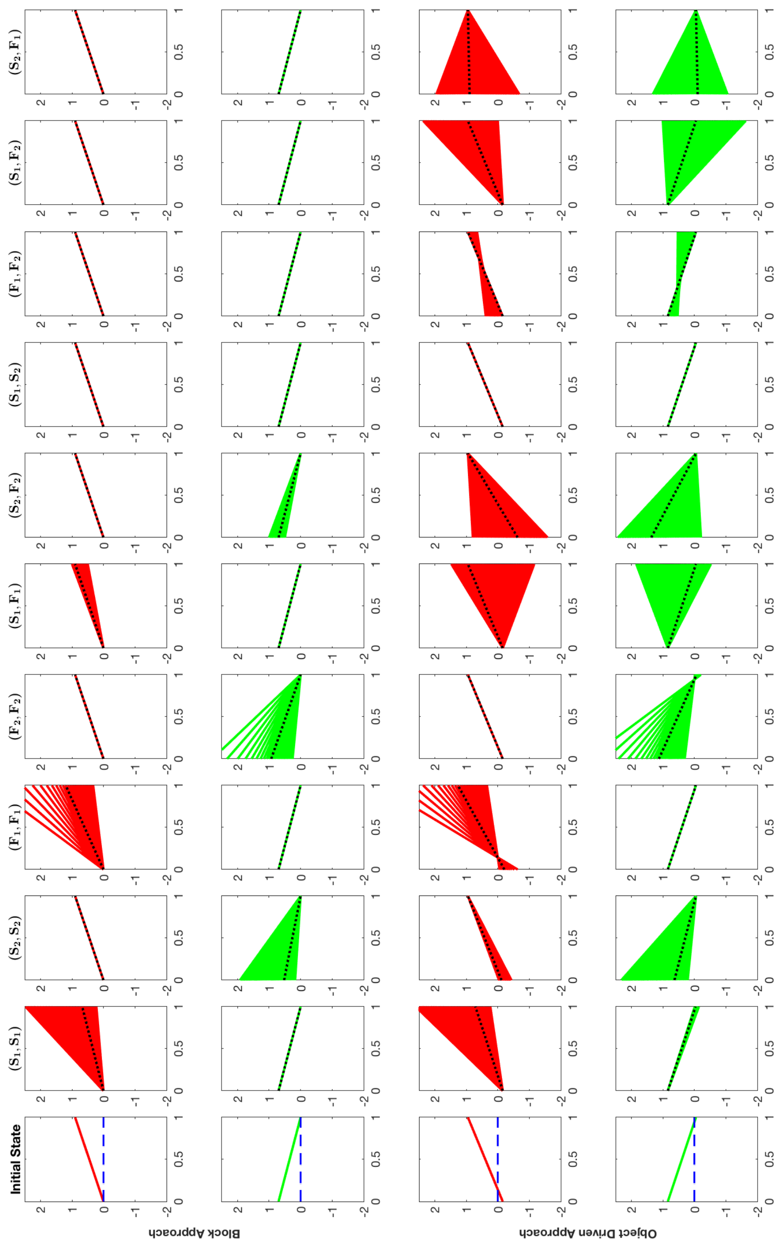

Figure 2.1: From top to bottom: S1 (red colour) and S2 (green colour) hedge ratios for the BA (first two rows) and ODA (last two rows). On the x-axis we display the S_1 weight varying from 0 to 1, while on the y-axis the hedge ratio values. From left to right, the first column displays the initial state, then we perturbate the volatilities of S_1, S_2, S_3, S_4 and the correlations between S_1, F_1, S_2, F_2, S_1, S_1, F_1, F_2, S_1, F_2, S_2, F_1, ceteris paribus. Blue dashed lines correspond to a constant zero line; black dotted lines reproduce the initial state as in column 1.

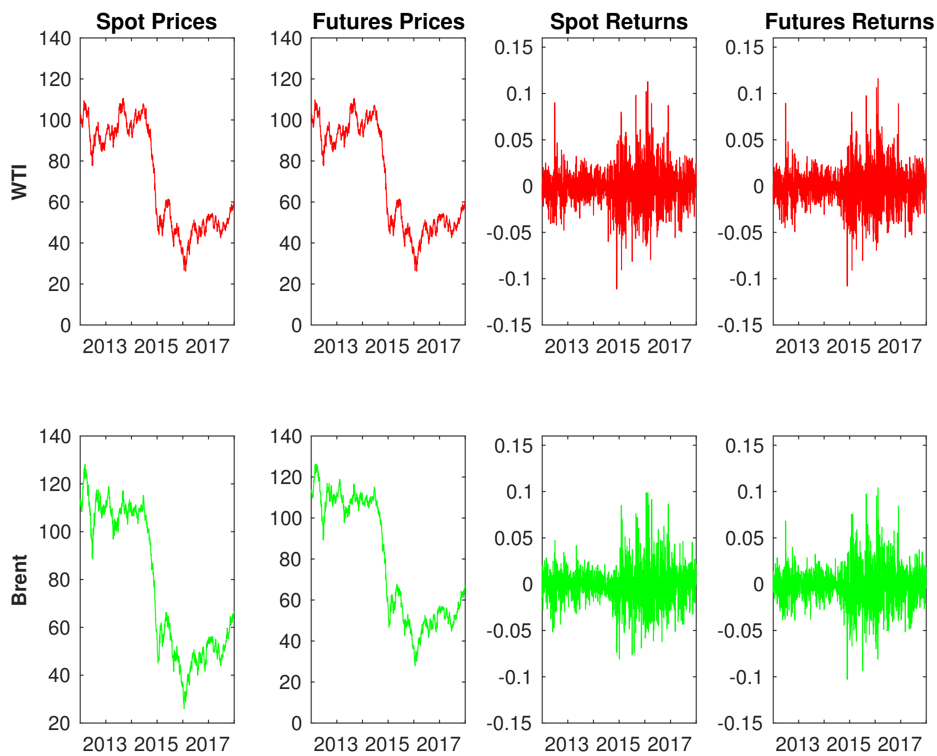

Figure 2.2: From left to right: spot and Futures prices and returns for the WTI (first row – red colour) and Brent (second row – green colour).

Appendix 2.A Filtering results

Table A2.1: Parameters of the ARMA(1,1)–GARCH(1,1) filtering for the BA.

		WTI		Brent	
		Spot	Futures	Spot	Futures
ARMA(1,1)	Intercept	-0.0003	-0.0003	-0.0003	-0.0003
	std	0.0005	0.0005	0.0005	0.0004
	AR	-0.4405	-0.4517	0.2601	-0.5910
	std	0.2335	0.1666	0.3075	0.1320
	MA	0.3808	0.3685	-0.2046	0.5090
	std	0.2402	0.1734	0.3070	0.1418
GARCH(1,1)	Constant	2.55E-06	2.36E-06	8.22E-07	1.44E-06
	std	1.09E-06	1.40E-06	0.0065	7.89E-07
	GARCH	0.9363	0.9307	0.9498	0.9474
	std	0.0077	0.0091	0.0065	0.0059
	ARCH	0.0589	0.0651	0.0492	0.0497
	std	0.0076	0.0081	0.0067	0.0058

Table A2.2: Parameters of the ARMA(1,1)–GARCH(1,1) filtering for the ODA

		Portfolio	WTI Futures	BRENT Futures
ARMA(1,1)	Intercept	-0.0003	-0.0003	-0.0003
	std	0.0005	0.0005	0.0004
	AR	0.2420	-0.45175	-0.59108
	std	0.2685	0.1666	0.1320
	MA	-0.17838	0.3685	0.5090
	std	0.2708	0.1734	0.1418
GARCH(1,1)	Constant	1.37E-06	2.36E-06	1.44E-06
	std	7.99E-07	1.40E-06	7.89E-07
	GARCH	0.9389	0.9307	0.9474
	std	0.0077	0.0091	0.0059
	ARCH	0.0584	0.0651	0.0497
	std	0.0078	0.0081	0.0058

Appendix 2.B MGARCH estimates for both Approaches under Dynamically Rebalanced Hedging strategies

Table B2.1: Parameters for the BA under the Dynamically Rebalanced Hedging framework.

	DCC_w	DCC_b	$DBEKK_w$	$DBEKK_b$	$DRBEKK_w$	$DRBEKK_b$	$DRDCC_w$	$DRDCC_b$
Θ_1	0.0502	0.0531						
std	0.0202	0.0303						
Θ_2	0.9416	0.7058						
std	0.0197	0.2376						
C			0.1777	0.1673				
std			0.0464	0.0889				
C			0.1847	0.1222				
std			0.0474	0.0280				
C			0.0298	0.1131				
std			0.0085	0.0231				
A			0.2484	0.0838				
std			0.0459	0.0703				
A			0.2433	0.0763				
std			0.0471	0.0454				
B			0.9589	0.9825				
std			0.0137	0.0200				
B			0.9581	0.9832				
std			0.0145	0.0055				
α_{11}					0.2466	-0.0774	0.2760	0.2185
std					0.0423	0.0358	0.0563	0.0546
α_{22}					0.2262	-0.0783	0.1871	0.2527
std					0.0301	0.0306	0.0174	0.0163
β_{11}					0.9590	0.9829	0.9575	0.7356
std					0.0144	0.0157	0.0816	0.0632
β_{22}					0.9664	0.9821	0.9803	0.9469
std					0.0103	0.0073	0.0624	0.0462

Table B2.2: Parameters for the ODA under the Dynamically Rebalanced Hedging frameworks.

	DCC	DBEKK	DRBEKK	DRDCC
Θ_1	0.0164			
	0.0050			
Θ_2	0.9799			
	0.0084			
C		0.0948		
		0.0256		
		0.0743		
		0.0191		
		0.0834		
		1.97E-02		
		0.0284		
		0.0106		
		0.0214		
		0.0087		
		0.0321		
		0.0098		
A		0.1008		
		0.0180		
		0.1086		
		0.0163		
		0.1141		
		0.0169		
B		0.9909		
		0.0036		
		0.9914		
		0.0028		
		0.9899		
		0.0033		
maxα_{ii}			0.1493	0.3295
			0.0597	0.0647
minα_{ii}			0.1114	0.1135
			0.0231	0.0265
maxβ_{ii}			0.9891	0.9899
			0.0034	0.0035
minβ_{ii}			0.8255	0.8290
			0.1084	0.0926
max$\alpha_{ii} + \beta_{ii}$			1.1384	1.3194
min$\alpha_{ii} + \beta_{ii}$			0.9369	0.9425

Chapter 3

Minimum Regularised Covariance Determinant Estimator: a novel application to Portfolio and Interest Rates modelling [1]

3.1 Introduction

In this chapter we introduce the Minimum Regularised Covariance Determinant estimator (MRCD), a robust estimator for the covariance matrix and its inverse (precision matrix), into the financial literature. We do this with a twofold aim: first, we look at portfolio selection, in particular at the Global Minimum Variance portfolio [Black and Litterman, 1992]. Here, we can make use of the MRCD closed form solution for estimating the precision matrix, which is the only input of the model. Existing research has not focused much on direct estimations for the covariance matrix inverse to enhance asset allocation. To the best of the author's knowledge, only [Senneret et al., 2016] and [Torri and Giacometti, 2017] discussed the direct estimation of the precision matrix within the GMV portfolio framework, proposing the use of the Graphical LASSO technique. Second, we introduce the MRCD into the analysis of interest rates variance and covariance. We run a comparison among other statistical estimators for the covariance to finally propose a case study to detect the main volatility drivers of the US term structure. This analysis is particularly useful for both portfolio construction and risk management in the fixed income space.

The chapter is organised as follows. Section 3.2 describes the application of MRCD to the Global Minimum Variance portfolio. After some introductory remarks and an extensive literature review, we perform two case studies. The first consists in a Monte Carlo experiment,

[1]The research paper *Portfolio Selection with Minimum Regularised Covariance Determinant*, co–authored with my PhD supervisors, Marina Resta and Maria Elena De Giuli, is based on the results in the first part of this chapter and has been submitted for publication. The paper has been presented at the XIX Quantitative Finance Workshop (University of Roma Tre, January 24-26 2018); at the 29th European Conference on Operational Research (EURO - Valencia, July 08-11 2018) and at the 42nd annual meeting of the AMASES (Naples, September 13-15 2018). The Research paper *A comparison of Estimation Techniques for the Covariance Matrix in a Fixed Income Framework*, co–authored with Marina Resta, is based on the results in the second part of this chapter and it has been published in *New Methods in Fixed Income Modelling*, pp. 99–115, Springer, Cham.

where we compare the out–of–sample performance and stability of two GMVP allocations, one under the MRCD and the other one under the sample covariance matrix. The second is an empirical illustration featuring five real data investment universes, where we run the same comparison to assess the benefits of MRCD even with real data. Section 3.2.7 traits the conclusion for this first MRCD application. Section 3.3 contains the application of MRCD to interest rates modelling. The main goal is to compare the MRCD against four alternative estimators for the covariance matrix for portfolio optimisation and risk management purposes in the fixed income space. Therefore, Section 3.3.1 explores this comparison, while Section 3.3.2 provides an empirical case study where we disentangle the US yield curve risk factors. Section 3.3.4 concludes the MRCD application to interest rates. Finally, Section 3.4 concludes the chapter.

3.2 Minimum Regularised Covariance Determinant for Portfolio Selection

Since its inception, the [Markowitz, 1952] mean-variance optimisation model suffered criticisms towards many directions, originating relevant research strands that have directly or indirectly contributed to build the Modern Portfolio Theory apparatus, see e.g. [Fabozzi et al., 2002, Kolm et al., 2014].

In this chapter we are interested in the issue concerning the estimation error: determining optimal mean-variance portfolio weights, in fact, requires the knowledge, for each security, of expected returns, variances and covariances. However, in practice those variables are unknowns and need to be estimated using historical information; the estimation error arises from this process, and directly affects resulting portfolio weights [Ledoit and Wolf, 2004b]. The Markowitz suggestion of calculating portfolio allocations with sample estimators has received severe objections: evidence outlines that weights may be biased, exhibiting too extreme values either positive or negative, with large fluctuations over time [DeMiguel et al., 2007].

Our work nests in a specific strand of the described literature, since we focus on the Global Minimum Variance Portfolio (GMVP) problem. The GMV portfolio is appealing for various reasons: first, it holds a closed form solution when short-selling is allowed; second, as it relies on the use of the covariance matrix only, it makes possible excluding the mean, i.e. the largest source of estimation error [Merton, 1980]; third, it carries on good and stable out–of–sample performances [Maillet et al., 2015]. We give at least two contributions to the debate: *(i)* we are aimed at reducing the effect of parameter uncertainty in the GMVP; *(ii)* we discuss the use of an explicit estimator for the inverse of the covariance matrix (i.e. precision matrix), which is well-fitted for dealing with high–dimensional and non-Normal issues. With respect to parameter uncertainty in the GMVP framework, the investigation vein has been particularly vivid, offering several alternatives to tackle on the problem: [Jagannathan and Ma, 2003] suggested to directly constrain portfolio weights, while [DeMiguel et al., 2009] constrained the portfolio norm. Another option consisted in shrinking the sample covariance matrix towards a more structure covariance estimator [Ledoit and Wolf, 2003] or double shrinking it [Candelon et al., 2012] applying the ridge regression methodology on shrunk weights. Under a similar spirit, [Maillet et al., 2015] proposed to use the robust regression framework, which resides between the worst-case GMVP solution (when the input is the sample covariance matrix) and the equally weighted portfolio. In addition, [Frahm and Memmel, 2010] suggested two shrinkage estimators for the Minimum Variance portfolio, finding that both dominate the sample one; [Kourtis et al., 2012] proposed two non-parametric shrinking methodologies and demonstrated that in this way it is possible to enhance the out–of–sample portfolio performance; [DeMiguel et al., 2013] shrunk the inverse of the sample covariance matrix with a novel calibration procedure for the shrinkage intensity; [Bodnar et al., 2018] used the random matrix theory to derive a feasible and simple estimator that works for both small and large sample sizes; [Sun et al., 2018] carried on the Cholesky decomposition of the covariance matrix deriving a Stein-type shrinkage strategy.

However, deriving the GMVP weights poses an additional challenge in comparison to other risk-based portfolios[2], since here the covariance matrix must be not only well-conditioned, but also invertible. The inverse covariance matrix (e.g. precision matrix), in fact, assumes a non-singular covariance matrix. Nevertheless, this condition depends on the dataset structure: a typical case of singularity involves the presence of intrinsic multicollinearity in the dataset that makes the covariance matrix not invertible[3]. Another case occurs when the number of observations is too small compared to the number of variables; [Bodnar et al., 2018] illustrated how the GMVP reacts in presence of low and high–dimensional datasets, controlling the concentration ratio c, i.e. the ratio between the number of assets and observations. With regards to the precision matrix estimation (see point *(ii)* of the listings given in previous rows), so far this issue has been examined and discussed only in a few number of contributions. A bespoke characterisation of the precision matrix for portfolio optimisation problem can be found in [Stevens, 1998]. However, only [Senneret et al., 2016] and [Torri and Giacometti, 2017] discussed the direct estimation of the precision matrix within the GMV portfolio framework, proposing the use of the Graphical LASSO technique. To such aim, our contribution resides in introducing the Minimum Regularised Covariance Determinant (MRCD) estimator of [Boudt et al., 2018] to determine the GMVP weights. The MRCD is an explicit estimator of the precision matrix; its added value relies in that it seems capable to cover both the high–dimensional and the non-Normality issues, hence

---

[2]The Inverse Volatility [Leote et al., 2012], the Equal-Risk-Contribution [Maillard et al., 2010] and the Maximum Diversification [Choueifaty and Coignard, 2008] portfolios just need the covariance matrix to derive weights.

[3]To avoid this, [Pantaleo et al., 2011] suggested using the Moore-Penrose inverse to assure the existence of the precision matrix.

allowing an overall improvement of weights misspecification. To the best of our knowledge, this is a first time application as the MRCD has been not yet applied to solve portfolio problems.

Overall, this chapter investigates the benefits for the GMVP out–of–sample performance when the MRCD is used against the sample estimator. To such aim, we work towards two directions. First, we included an extensive Monte Carlo simulation to check the effectiveness of the proposed approach, covering several cases for the concentration ratios c ($c < 1; c \approx 1; c > 1$) and different levels of departure from returns Normality. Second, we take on an empirical case study featuring five real investment universes with different stylised facts and dimensions. Both the simulated and empirical applications clearly demonstrate that the GMVP out–of–sample performance improves from the use of the MRCD estimator: results are a reduction in the portfolio turnover at no cost for the portfolio variance, and an increase in portfolio expected returns. Furthermore, the Monte Carlo analysis shows that applying the MRCD lowers the GMVP weights misspecification.

The remain of Section 3.2 is organised as follows. In Section 3.2.1, we start by providing an overview of the GMVP selection problem under parameter uncertainty, presenting the sample and the MRCD estimators. Section 3.2.2 analyses the properties of the MRCD via simulations. An empirical application is presented and discussed in Section 3.2.5. Section 3.2.7 concludes.

3.2.1 Global Minimum Variance Portfolio

Let us denote by $X = (\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n)'$, $\mathbf{x}_i \in R^n$ with $(i = 1, ..., p)$ the $n \times p$ matrix defining the investment universe of n returns for p assets. According to the Black–Litterman model [Black and Litterman, 1992], the optimal portfolio weights $\boldsymbol{\omega}^*$ solve the optimisation problem:

$$\boldsymbol{\omega} = \underset{\boldsymbol{\omega}}{\operatorname{argmin}}\, \boldsymbol{\omega}'\Sigma\boldsymbol{\omega}, \qquad s.t.\;\; \boldsymbol{\omega}'.\mathbf{1} = 1, \tag{3.1}$$

where $\boldsymbol{\omega}$ is the portfolio weights vector of length p; $\mathbf{1}$ is the vector whose p components are all equal to one and Σ is the square covariance matrix of asset returns; the dot operator (.) indicates the scalar product. The constraint $\boldsymbol{\omega}'.\mathbf{1} = 1$ simply means that all the available wealth is allocated, and therefore it must be interpreted as a budget constraint. In the absence of positivity constraints on portfolio weights, (3.1) admits a closed form solution:

$$\boldsymbol{\omega} = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}'\Sigma^{-1}\mathbf{1}}. \tag{3.2}$$

Since by construction the covariance population matrix is not observable, then (3.2) should not be feasible unless using a proper estimator for Σ. It is generally accepted the use of the Maximum Likelihood estimator $\hat{\Sigma}_s$ for the covariance matrix (sample covariance matrix):

$$\hat{\Sigma}_s = \frac{1}{n-1} X' \left(I - \frac{1}{n} \mathbf{1}\mathbf{1}' \right) X, \tag{3.3}$$

where I denotes the identity matrix of dimension p. The estimator in (3.3) is very sensitive to the underlying data structure and to its statistical properties.

When the precision matrix is needed, the use of the Maximum Likelihood estimator has been criticised under various aspects. First, using $\hat{\Sigma}_s$ can pose severe problems when solving for the optimal portfolio weights, depending on the relation between p and n. In fact, when $p \gg n$, i.e. when the number of assets in the investment universe is sensitively greater than the number of observations, the sample covariance matrix is singular and hence not invertible. This issue can be overcome by replacing $\hat{\Sigma}_s$ with the Moore-Penrose inverse:

$$\hat{\Sigma}_s^+ = \left(\hat{\Sigma}_s^* \hat{\Sigma}_s \right)^{-1} \hat{\Sigma}_s^*,$$

with $\hat{\Sigma}_s^*$ being the conjugate transpose of $\hat{\Sigma}_s$. This is line with [Pantaleo et al., 2011] [Bodnar et al., 2018], who replaced the sample covariance matrix $\hat{\Sigma}_s$ with its generalised inverse when the former it is not invertible.

Second, even when $p \approx n$, the sample covariance matrix carries on a large estimation error and its inverse is a poor estimator for *true* precision matrix.

Third, it has been objected that the sample covariance matrix requires the estimation of $\frac{1}{2}p(p+1)$ parameters, so that it carries on strong estimation error when the number of assets is very large. In addition, assuming the Normality of asset returns, the precision matrix is a biased estimator for its population counterpart, even if the sample covariance matrix is unbiased [Senneret et al., 2016]. This means that the solution of (3.1) might be not trustworthy.

Minimum Regularised Covariance Determinant estimator

The Minimum Regularised Covariance Determinant (MRCD) estimator [Boudt et al., 2018] assumes that data are mostly drawn from an elliptical distribution; the data not covered by this distribution are assumed coming from a deviation distribution and therefore they can reasonably considered as outliers. This working assumption seems perfectly aligned to the features of financial data, characterised by stylised facts such as non-Normality and presence of outliers [Cont, 2001].

The MRCD can be viewed as an extension of the Minimum Covariance Determinant (MCD) estimator [Rousseeuw, 1984] to model high–dimensional datasets. While the MCD returns the sample covariance matrix on the subset of the data, that is the covariance matrix minimising the dispersion among the observations, the MRCD differs in the selection principle of the subset which is now a convex combination between a target matrix T and the sample covariance itself. In this way, by construction, it is possible to let the MRCD working with high–dimension datasets, also when the ratio between the number of assets p and the number of observations n, i.e. the concentration ratio c is greater than one. The MRCD preserves the good properties of the MCD and it is always well-conditioned by construction.

The procedure leading to the MRCD estimator can be summarised in a few steps, and it is briefly explained in the next rows, where we use the same notations as at the beginning of this Section. We start by standardising each $\mathbf{x}_i \in X$ with:

$$\mathbf{u}_i = D_X^{-1}(\mathbf{x}_i - \boldsymbol{\nu}_i),$$

where $\mathbf{u}_i$ is the vector of standardised observations for the generic asset i collected in the $n \times p$ matrix U; $\mathbf{x}_i$ is the initial observation vector for the *i-th* asset; D_X^{-1} is the diagonal matrix of dimension p of estimated mean values and $\boldsymbol{\nu}_X$ is the $p \times 1$ vector of medians.

The MRCD works on subsets of size $h \leq n$ of the standardised data universe U, where we can identify a submatrix of U, say U_h, of dimension $h \times p$. Finally, replicating this task for each h, $[\frac{n}{2} + 1] \leq h \leq n$, where $[.]$ denotes the integer part, we get the collection of all possible U_h. For each subset h, the corresponding covariance matrix is a regularised matrix K:

$$K(h) = \rho T + (1 - \rho) c_\alpha S_U(h), \tag{3.4}$$

where ρ is the shrinkage weight or regularisation parameter, determining how much the sample covariance matrix is shrunk towards the target matrix T[4]. To ensure that $K(h)$ is positive definite and hence invertible, it must be $0 \leq \rho \leq 1$. In addition, $S_U(h)$ is the sample covariance matrix calculated in the subset h on U and z_α is the consistency factor as defined in [Croux and Haesbroeck, 1999], with $\alpha = (n - h)/n$ being the so-called trimming percentage. Before moving on, a remark on the target matrix T is required. [Boudt et al., 2018] proposed to use the identity matrix for general cases, or a regularised version of the identity itself. The latter is based on the average robust correlation among observations, and should be employed when an equal correlation structure is suspected. Both alternatives are conceived in such a way to ensure the positive definitiveness of the target.

Following [Boudt et al., 2018], we rewrite (3.4) with a factorisation that takes advantage of

[4]To ensure that the subsample is well–conditioned, [Boudt et al., 2018] select ρ following a data-driven procedure implying that the resulting matrix has condition number equals to 1000 at maximum.

the Singular Value Decomposition (SVD) of the target matrix: $T = Q\Lambda Q'$, with Q being the orthogonal matrix of eigenvectors from the target matrix and Λ the related diagonal matrix of the eigenvalues of T:

$$K(h_{MRCD}) = Q\Lambda^{\frac{1}{2}} \left[\rho I_p + (1-\rho) z_\alpha S_W(h)\right] \Lambda^{\frac{1}{2}} Q', \tag{3.5}$$

where:

$$S_W(h^*_{MRCD}) = \Lambda^{\frac{1}{2}} Q' S_U(h) Q \Lambda^{\frac{1}{2}},$$

with $\mathbf{w}_i = \Lambda^{\frac{1}{2}} Q' \mathbf{u}_i$, $(i = 1, ..., p)$. Note that setting $\rho = 0$, (3.5) returns the same estimator as the MCD, which is a particular case of the MRCD. Determining the MRCD covariance matrix means finding the subset of the original dataset that minimises the determinant of the regularised matrix $K(h)$:

$$h^*_{MRCD} = \underset{h}{\operatorname{argmin}} \left(det(K(h))^{\frac{1}{p}} \right),$$

so that final estimator is given by:

$$K_{MRCD} = K(h^*_{MRCD}) = D_X Q \Lambda^{\frac{1}{2}} \left[\rho I_p + (1-\rho) c_\alpha S_W(h)\right] \Lambda^{\frac{1}{2}} Q' D_X. \tag{3.6}$$

The fact that the MRCD allows deriving K as a matrices factorisation makes possible to have at hand the inversion procedure which is required to solve the GMVP. The inverse of the covariance matrix (i.e. the precision matrix) is in fact given by:

$$K^{-1}_{MRCD} = D_X^{-1} Q \Lambda^{\frac{1}{2}} \left[\rho I_p + (1-\rho) c_\alpha S_W(h)\right]^{-1} \Lambda^{\frac{1}{2}} Q' D_X^{-1}.$$

3.2.2 A Monte Carlo Experiment

In this Section, we employ simulated data to evaluate the effectiveness of the MRCD[5] to address the high–dimension and non–Normality issues in the estimation of GMVP weights and the related out–of–sample performance. Results are always compared to those of the sample covariance estimator, replaced by the Moore-Penrose inverse when $c \geq 1$. To such aim, following the framework defined in [Candelon et al., 2012], we simulate an initial investment universe composed by 100 assets and a total of 12000 monthly observations (corresponding to 1000 years); we assume that the returns are generated from the distribution G:

[5] To calculate the MRCD, we used the R code available at url: https://wis.kuleuven.be/stat/robust.

$$G = (1 - \xi)N + \xi A, \tag{3.7}$$

where N is a multivariate Normal distribution whose features will be discussed in the next rows; A is an alternative deviation distribution and $\xi = [0, 0.1]$ is the intensity controlling the deviation from Normality. By construction, G is an elliptical distribution. Setting $\xi = 0$, data are multivariate Normal. Normal returns are simulated using a one-factor model as data generating process:

$$\mathbf{r}_t = \beta \mathbf{f}_t + \boldsymbol{\epsilon}_t, \tag{3.8}$$

$\mathbf{r}_t$ is the monthly return vector; $\mathbf{f}_t$ is the $p \times 1$ vector of returns on the factor; β is the scalar representing the factor loadings and $\boldsymbol{\epsilon}_t$ is the $p \times 1$ vector of residuals. The asset factor loadings are drawn from a Uniform distribution in the range $[0.50, 1.50]$, while returns on the single factor are generated from a Normal distribution with monthly mean $0.05/12$ and variance $0.16/12$[6]. In addition, residuals are drawn from a Uniform distribution in the range $[0.15, 0.25]$ so that the related covariance matrix is diagonal with an average monthly volatility of $0.2/12$, as in [DeMiguel and Nogales, 2009]. On the other hand, the deviation alternative distribution A is generated via *substitutive contamination* [Perret-Gentil and Victoria-Feser, 2005]: for each asset of the investing universe a percentage ξ of observations is replaced by adding to the original values the asset mean plus five times the standard deviation.

To incorporate the dataset non–Normality, we use four different values for ξ, $\xi = \{0\%, 2.5\%, 5\%, 10\%\}$, to randomly replace the corresponding percentage of every asset time series. This design allows us investigate both the case of Normal returns ($\xi = 0$) and the cases of low, moderate and severe returns contaminations ($\xi = 2.5\%$, $\xi = 5\%$, $\xi = 10\%$, respectively). Regarding the dimension of the dataset, we consider five different in–sample lengths upon which calculating the asset returns covariance matrix, as illustrated in Table 3.1. Keeping constant the number of assets ($p = 100$), we are able to assess the goodness of MRCD for very different concentration ratios, monitoring the behaviour of the estimator for the low dimension case, i.e. when $c < 1$; in the case of $c \approx 1$, when p and n are approximately of the same size; and finally the high–dimension case, when $c > 1$.

[Table 3.1 about here.]

Our simulated study is organised on a few steps:

1. Step 1, we simulate returns according to the scheme defined in the previous rows, see (3.7) and (3.8);

[6] Parameters for the Normal distribution of the single factor are set following [DeMiguel and Nogales, 2009, Maillet et al., 2015]

2. Step 2, we estimate the asset returns covariance matrix on the initial m observations, where m is as described in Table 3.1;
3. Step 3, we derive the GMVP weights according to (3.2) and we correspondingly allocate the available wealth.
4. Step 4, we calculate the out–of–sample portfolio returns in the next period $m + 1$, which is part of the $n - m$ out–of–sample observations.

Steps 1-3 are repeated $n - m$ times using a rolling window to move across data: at each run, the m in–sample observations moves on by one observation, discarding the oldest that is replaced by the $m+1$ newest one, which is part of the $n-m$ out–of–sample observations. This procedure allows us to compute $n - m$ portfolio weights and to evaluate the out–of–sample performance upon $n - m$ Monte Carlo trials.

With respect to the performance, we monitor an overall number of four evaluation metrics: the turnover, the Global Portfolio standard deviation, the portfolio Sharpe ratio, and the L1 distance.

Recalling that $\mathbf{r}_{t+1}$ is the p–length vector of asset returns at time $t+1$, we define the portfolio mean and the portfolio variance as:

$$\hat{\mu} = \frac{1}{n-m} \sum_{t=m}^{n-1} \hat{\boldsymbol{\omega}}_t^{'} . \mathbf{r}_{t+1};$$

$$(\hat{\sigma})^2 = \frac{1}{n-m-1} \sum_{t=m}^{n-1} \left(\hat{\boldsymbol{\omega}}_t^{'} . \mathbf{r}_{t+1} - \hat{\mu} \right)^2 ;$$

respectively. The turnover and the portfolio Sharpe ratio are then given by:

$$Turnover = \frac{1}{n-m-1} \sum_{t=m}^{n-1} \sum_{j=1}^{p} \left(\left| \hat{\boldsymbol{\omega}}_{j,t+1}^{(i)} - \hat{\boldsymbol{\omega}}_{j,t}^{(i)} \right| \right) ;$$

$$\hat{SR} = \frac{\hat{\mu}}{\hat{\sigma}}$$

where m is the in–sample length. These measures evaluate the overall performance of the MRCD, while the L1 distance assesses how much the misspecification in the covariance matrix affects the estimated portfolio weights $\hat{\boldsymbol{\omega}}$:

$$\|\boldsymbol{\omega} - \hat{\boldsymbol{\omega}}\|_1 = \sum_{i=1}^{p} |\boldsymbol{\omega}_i - \hat{\boldsymbol{\omega}}_i|$$

At this point, using the L1 norm makes sense because, being this a simulation study, we know the *true* covariance matrix and hence the *true* GMVP weights $\boldsymbol{\omega}$.

3.2.3 MRCD estimation: a note

As discussed in Section 3.2, the MRCD assumes working on subsamples of length h. Under the spirit of giving a comprehensive discussion of this estimator, we run all the simulation for $h = \{0.5, 0.75, 0.9\}$, those values, in our opinion, well accomplish to the assumptions in [Boudt et al., 2018] who although suggesting a data–driven procedure to select h, recommend using $h = 0.75$ to represent datasets with outliers. Moreover, we selected to use the scaled version of the Identity as target matrix T to account for equicorrelation in asset returns.

We preliminary observe that the parameter ρ, which drives the regularisation towards the target matrix T changes varying the dimension of the dataset and the contamination level. Figure 3.1 show the behaviour of ρ for the selected values of h, varying the concentration ratio, at different levels of contamination. The data on which the graphs were built are given in Appendix 3.4.

[Figure 3.1 about here.]

When the dataset is low dimensional ($c < 1$), the MRCD results essentially agree with those of the MCD, with the exception of the case $h = 0.5$, $\xi = 10\%$, when $\hat{\rho}$ is greater than 0.01. On the other hand, when the numbers of observations and assets are closer one to each other ($c \approx 1$) the regularisation increases, although the final results are similar to those of the MCD. In addition, when $h = 0.5$ the strength of regularisation towards the target matrix increases for higher levels of contamination. Finally, when the dataset is high–dimensional ($c > 1$), the MRCD comes into play with a larger regularisation, which increases for larger data contamination. Again, the MRCD is more stable for $h \geq 0.75$. Overall, the MRCD converges to the MCD when data are low dimensional ($c < 1$); in this case, the estimates are stable and non-Normal, especially when $h \geq 0.75$. For higher concentration ratios, the MRCD needs a greater regularisation intensity. Again, when $h \geq 0.75$, $\hat{\rho}$ is more stable with respect to different levels of contamination.

3.2.4 Simulation Results

In accordance to what stated at the beginning of Section 3.2.2, we start discussing the performance of the MRCD with the aid of both financial and statistical indicators. Table 3.2 reports out–of–sample portfolio turnovers with simulated data for the five subsamples described in Table 3.1. In detail, the out–of–sample turnovers calculated with the sample covariance matrix are given in Panel A, while the results for the MRCD are shown in Panel B, for $h = 0.5$; in Panel C for $h = 0.75$; in Panel D for $h = 0.9$. The level of contamination is also reported for each Panel.

[Table 3.2 about here.]

Overall, the results in Table 3.2 suggest that in terms of misspecification of the portfolio turnovers the dimension of the dataset matters more than the non–Normality: in fact, the turnovers show the greater discrepancy if compared across different sizes, while it is lower for different contamination levels, ceteris paribus. For low levels of c ($c < 1$) the sample and the MRCD estimators perform similarly. However, the more the concentration ratio grows, the more convenient is using the MRCD. For example, when $c \approx 1$, the MRCD with $h = 0.75$ performs at the best, displaying a turnover which is half the one provided by the sample estimator. This consideration holds for all the contaminated samples. Moreover, this trend is confirmed when $c > 1$, with the MRCD showing similar turnovers for different values of h. Interestingly, the highest turnover is attained by both the estimators when $c \approx 1$, confirming that this case is particularly difficult to handle as noted in [Bodnar et al., 2018].

Next indicator is the out–of–sample portfolio variance which is useful to assess the stability of the MRCD. Results are given in Table 3.3 following the same organisation in panels as discussed for Table 3.2.

[Table 3.3 about here.]

For low–dimensional datasets, the similarity between the MRCD and the sample estimator is evident in the S1 and S2 subsamples, where the two competing estimators attain similar results. Switching to the case $c \approx 1$, the sample estimator generates the lowest portfolio out–of–sample variances at any levels of h: the MRCD shows closer estimates only for $h = 0.9$. These considerations hold also for all the considered contaminated samples. However, when $c > 1$, the MRCD stems for its good performance. For example, in the subsample S4 the lowest global variance is generated by the MRCD with $h = 0.75$ and $h = 0.9$; in the case of S5 the MRCD always beats the sample estimator. Overall, with high–dimension datasets, the MRCD shows to be stable with respect to various dataset contaminations, with only small discrepancies between the Normal and the most contaminated case ($\xi = 10\%$).

Lastly, Table 3.4 displays the Sharpe ratios for the various simulated datasets, using the conventions already adopted and discussed for Tables 3.2 and 3.3.

[Table 3.4 about here.]

The trend already highlighted for the turnover and the variance is therein confirmed: for low-dimensional datasets (S1 and S2), the sample and the MRCD perform in a similar fashion, and for $h = 0.9$ the MRCD gets the best results. On the other hand, when $c \approx 1$ or higher, the MRCD performs at the best, doubling the Sharpe ratio of the sample estimator. For example, for $h = 0.5$ and $h = 0.75$, the MRCD works at best in S3, while in both S4 and S5 the MRCD exhibits very good performances for every values of h. In general, the Sharpe ratio seems increasing with the level of contamination: this is probably due to the presence of positive outliers.

Lastly, we discuss the misspecification effect on the GMVP weights. As remarked in previous rows, this is possible because working with simulated data we know the *true* covariance matrix, and hence the true portfolio weights. Figure 3.2 displays the results for the L1 norm using the boxplot formalism: distance values lie on the vertical axis, while the concentration ratio is on the horizontal axis. Each boxplot illustrates the departure from the mean for the contamination levels examined.

[Figure 3.2 about here.]

The main point is that the sample covariance matrix performs the best for low dimensional datasets, by minimising the distance from the *true* weights and also being the more stable across contaminated samples. The MRCD at $h = 0.75$ and $h = 0.90$ follows. However, as the concentration ratio approaches 1, the sample covariance matrix is still quite stable to departures from the Normality, but it shows the greatest distance from the *true* portfolio weights. On the contrary, the MRCD performs in line with the results of low-dimension datasets, with the MRCD for $h = 0.5$ showing a very stable performance, robust to misspecification in the covariance matrix. Moreover, the MRCD is again the best estimator in the case of high–dimension datasets, without any notable differences varying h.

3.2.5 Empirical Illustration: Five Investment Universes

In this section, we explore the effectiveness of the MRCD estimator in a case study involving real data[7], focusing on five investment universes, as described in Table 3.5[8]. For replication purposes, we refer to data freely available on the Kenneth French's website for what it concerns datasets 1 to 4[9] (the site also contains a comprehensive description of the datasets), and to data provided by Bloomberg[10]. For the dataset 5; the complete listings of asset tickers is given in Appendix 3.4. As a general remark, the MRCD estimation settings are the same employed with in Section 3.2.2, while the inverse of the sample estimator is replaced with the Moore–Penrose inverse when $c \geq 1$.

[Table 3.5 about here.]

The selected datasets, all at monthly frequency, have different features: those from K. French's website represent a set of portfolios selected using different factors, such as the

[7]As noted in [DeMiguel and Nogales, 2009] earlier and in [Candelon et al., 2012] later, empirical analyses involving real data are prone to generate spurious results. However, we believe that insights on real data can be beneficial especially for those practitioners aimed at using the MRCD in practical applications.

[8]As described in [Ardia et al., 2017, Bertrand and Lapointe, 2018], an empirical analysis involving different investment universe should be run upon different investment universes, since the composition and the level of correlation of each universe matter in estimating risk-based portfolios weights.

[9]`http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

[10]`http://www.bloomberg.com`.

Operating Profitability and Investment or the Size and Momentum. These datasets include 30 components at maximum. On the other hand, dataset 5 is made by 300 assets, allowing to work with a high–dimension dataset. Universes' statistical properties are shown in Table 3.6.

[Table 3.6 about here.]

To manage the curse of dimensionality, we selected five in–sample lengths, $m = \{360, 180, 90, 60, 36\}$ for the datasets 1-4; the corresponding concentration ratios c show a minimum value of 0.0694 ($m = 360$) and a maximum of 0.6944 ($m = 36$) in the case of the 25OPI, the 25SBM and the 25SM. In the case of the 30I, the minimum c is equal to 0.0833 ($m = 360$) while the maximum corresponds to 0.8333 ($m = 36$). In the case of dataset 5 the in–sample lengths are four, because $m = 360$ was not feasible, as the overall length of the sample was lower. Moreover, dataset 5 is the only case in which we have $c > 1$: its lowest value is 0.8333 ($m = 180$) while the highest is 8.333 ($m = 36$).

In this case study, we apply the same working strategy as already used with simulated data: we estimate the covariance matrix with the sample and the MRCD methodologies to assess the out–of–sample performances of GMV portfolios. Since with real data we have no knowledge about the *true* covariance matrix, we only check the behaviour of three indicators, i.e. the portfolio turnover, the variance and the Sharpe ratio.

MRCD estimation: technical details

Replicating the same steps as in Section 3.2.2, the MRCD has been estimated again considering three values of the subsamples: $h = \{0.5, 0.75, 0.9\}$.

[Figure 3.3 about here.]

Figure 3.3 shows the evolution of the parameter ρ on the different investment universes and dimensionality.

Likewise in Section 3.2.2, the data from which the graphs were obtained are given in Appendix 3.4. As noted with simulated data, increasing the concentration ratio of the dataset, the MRCD tends to stronger shrink the sample covariance matrix of the subsample h. This evidence is supported also by comparing all the investment universes with the 300SPX: it is clear that when the concentration ratio is higher than 1, the MRCD regularises more towards the target matrix T. On the contrary, the level of h seems not to influence the MRCD regularisation towards the target matrix.

3.2.6 Empirical Results

We are now able to discuss the results on observable data obtained by monitoring, the out–of–sample turnovers, the portfolios variance and the Sharpe ratio.

Table 3.7 shows the out–of–sample turnovers: columns represent the five asset universes, ordered according to the number of assets, from the smallest to the largest. Unlike in the cases discussed in Section 3.2.2, the high–dimension issue is treated in five distinct Panels: for each of them, we list the turnover for the sample and for the MRCD estimators, varying h. As previously remarked, the results for the 300SPX are in four Panels only, because the overall number of observations is too small to perform the analysis when $m = 360$.

[Table 3.7 about here.]

In general, the MRCD attains the lowest turnover for high–dimensional datasets. When the concentration ratio is higher than 1, as in the SPX300 case, the MRCD works better than the sample estimator for all the examined in–sample lengths. Interestingly, the higher the subsample parameter h, the better the out–of–sample performance. On the other hand, the sample estimator performs better than the MRCD when the concentration ratio is lower than 1. This is clear looking at the turnovers for the investment universes where $p = 25$: here, the GMVP weights computed with the sample estimator outperform the MRCD. Similarly to Section 3.2.2, the non–Normality of the dataset does not carry on a level of misspecification as great as the one due to the dataset dimension. To conclude, the turnover for the investment universe 30I supports previous findings at maximum level: at greater concentration ratio values correspond better performance of the MRCD. We now turn to the out–of–sample portfolio variance whose results are listed in Table 3.8.

[Table 3.8 about here.]

The results go towards the same direction as the turnover. In fact, the MRCD minimises the variance when $c > 1$ and $h > 0.9$. On the contrary, the sample estimator performs at the best for low-dimension datasets. Finally, Table 3.9 shows the out–of–sample Sharpe ratios for the observable datasets.

[Table 3.9 about here.]

Overall the MRCD estimator performs better than the sample estimator, indicating that the expected returns of MRCD based portfolios are higher than those based on the sample estimator.

3.2.7 Conclusion

This chapter focuses on the estimation of the precision matrix for deriving the Global Minimum Variance Portfolio (GMVP) weights when the asset universe is characterised by high–dimension and non–Normality. In this scenario, the main issue concerns the fact that resulting GMVP weights can be misspecified because of the large estimation error associated to the estimation of the precision matrix, i.e. the inverse of the covariance matrix.

Our contribution mainly resides in introducing the Minimum Regularised Covariance Determinant (MRCD) estimator to determine the GMVP weights. To the best of our knowledge, this is a first time application as the MRCD has been not yet applied to solve portfolio problems. The MRCD, in fact, seems particularly suitable to manage the aforementioned issues: it is based on a direct characterisation of the precision matrix, so that it makes possible to consider both the high–dimension and non–Normality of the investing universe. To give evidence of those facts, we carried on a comparison between the sample estimator (replaced by the Moore–Penrose inverse when the number of assets is greater than the one of observations) and the MRCD, using artificial data generated via Monte Carlo simulations at first, and with various observable investment universes, later.

We demonstrated that the MRCD lets improving the out–of–sample performance of GMV portfolios. In the case of Monte Carlo simulations we examined four performance metrics: turnover, portfolio variance, Sharpe ratio and L1 norm. In this case the MRCD got better results than the sample estimator especially in case of investment universes characterised by a concentration ratio higher than one and non–Normality. Moreover, the comparison between estimated and true portfolio weights made via the L1 norm, suggested that the MRCD limits the misspecification in portfolio weights. When testing the MRCD with observable datasets, we were able to use only three of the aforementioned performance portfolio metrics (turnover, variance, and Sharpe ratio), because the comparison with the true weights values was no more possible. However, the results clearly highlight the superior results the MRCD on the sample estimator.

We can therefore conclude that our findings are appealing for a wide audience: on the one hand, in fact, academics can find room enough to refine the use of MRCD in managing portfolio selection problems; on the other hand, MRCD seems a relatively easy–to–use technique for practitioners aiming to improve current GMVP allocations in a quite stable way with mathematical soundness.

3.3 Minimum Regularised Covariance Determinant for Interest Rates Modelling

The estimation of the covariance matrix plays an important role in many financial applications, including risk management, option pricing and portfolio selection. In fact, the it conveys all the information related to the co–movements among a bunch of financial securities, and hence allows the investors to allocate resources following the principle of diversification. Such universal "consensus" around the role of the covariance matrix holds also when focusing on fixed income instruments. Here, it is key for at least three reasons. First, in the term structure analysis, the covariance matrix helps to estimate co–movements across different maturities. Second, it can assess the main factors driving fixed income securities prices for risk metrics calculation and stress testing [Litterman and Scheinkman, 1991]. Third, inside the portfolio allocation framework, covariance analysis allows to determine how fixed income instruments react in combination to other financial products, helping investors in assessing the risk profile of interest rates securities [Martellini et al., 2003].

The covariance matrix is generally calculated by an inference procedure: being the dataset in use just a mere sample of the real population, which it is not known, the covariance matrix estimates the true relationships among population components. However, far from being an easy task, inferring the covariance is affected by a trade–off between the estimation error and the model error. On the one hand, model–free approaches guarantee the unbiasedness of the estimator without controlling for estimation error; on the other hand, parametric approaches impose some structure (model) for the covariance, lowering the estimation error at the cost of enlarging the one from the model [Briner and Connor, 2008].

The more commonly applied estimation technique relies on the so-called Sample Covariance matrix (SCVm): the work of [Markowitz, 1952] is a key example of the use of the sample estimator for portfolio selection. The most appealing feature of this technique is that it does not require specifying any structure for the covariance, that is: it implies a model-free approach [Anderson, 1963]. Moreover, when the data sample is Normally distributed, the SCVm is an unbiased estimator for the population moments [Briner and Connor, 2008]. However, estimating the SCVm has also pernicious drawbacks, especially in portfolio optimisation [Ledoit and Wolf, 2004b]. To make an example, when the sample size is smaller than the number of considered variables, the sample covariance matrix is ill–conditioned by construction [Ledoit and Wolf, 2004a]. This issue has been variously addressed, by proposing alternative methodologies to estimate the covariance matrix. In the case of portfolio selection, the attempts of enhancing the covariance matrix include the contribution of [Jorion, 1986] who used a Stein–type estimator. Moreover, [Michaud, 1989] demonstrated that Bayes–Stein estimators can positively impact on portfolio selection procedure in presence of outliers in the assets time series; [Black and Litterman, 1992] tackled the Michaud's issue by proposing a global equilibrium extension of the Markowitz's model; [Jagannathan and Ma, 2003] demonstrated that constraints on portfolio optimisation based on the SCVm generates the same positive outcome as applying shrinkage on the SCVm. On the other hand, problems arise also when, on the opposite, the sample size is too wide. The main point is that when the sample dimension is $n \times p$, where p is the number of assets and n is the number of observations for each asset, the covariance matrix is of dimensions n. However, as the maximum rank of the sample covariance matrix should not exceed n-1 [Briner and Connor, 2008], when p is greater than n (and hence, a fortiori greater than n-1) the covariance cannot be inverted [Schäfer and Strimmer, 2005]. This represents a serious issue in financial applications, especially within the fixed income framework, where large bond portfolios are the rule more than the exception, then dealing with high–dimensional portfolios. Finally, as demonstrated in [Ledoit and Wolf, 2004b] the SCVm is particularly sensitive to outliers in the dataset: sample eigenvalues are systematically spiked upward or downward, depending on the observations are either too much large or small, respectively. As pointed out by [Fan et al., 2005], we therefore can not trust on the sample covariance matrix, since it is hazardous to estimate it without imposing any structure.

Over the years these problems have been variously faced in the financial literature. [Pantaleo et al., 2011] presented a review of the solutions that deal with this model–free curse, and classified the resulting estimators into three research strands, relying on: (*i*) spectral properties of the covariance, (*ii*) the hierarchical clustering approach, and (*iii*) statistical models. The spectral properties of the covariance matrix inspire a research strand in connection to factor modelling which is based, in turn, on the Arbitrage Pricing Theory [Ross, 1976] and the Capital Asset Price Model [Sharpe, 1964], [Lintner, 1975]. In this class we can also include the estimators based on Random Matrix Theory – RMT – [Mehta, 2004], which operate a mathematical transformation directly on each sample eigenvalue by adding or subtracting value whether the eigenvalue is above or below a selected threshold. Main contributions include the work of [Briner and Connor, 2008], who used the RMT to identify the leading factors for his factor model for portfolio applications; moreover, [Pafka et al., 2004] used an exponential moving average model together to the RMT to improve the portfolio optimisation methodology, while [Frahm and Jaekel, 2005] applied the concept of RMT to minimise the risk of a portfolio based on the S&P500. Finally, [Wolf, 2007], highlighted that the resampled efficiency [Michaud and Michaud, 2008] is a very similar tool to RMT, since the portfolio optimisation input are calculated via a Monte Carlo resampling. The second research strand is linked to the hierarchical clustering approach [Anderberg, 1973]. This assumes that data can be clustered in groups according to a convenient similarity measure. By changing this similarity measure, data can be grouped in several different ways, thus enhancing the results of the portfolio optimisation procedure. The approach has been widely discussed by [Tola et al., 2008], in comparison to the RMT for optimising a portfolio composed by highly capitalised stocks from NYSE, as well as in [Pantaleo et al., 2011] who used the hierarchical clustering in a comparative study among nine covariance estimators. Finally, the class of statistical estimators is perhaps the most representative one: it includes the Shrinkage (SH) technique [Ledoit and Wolf, 2004b], as well as other robust estimators like the Minimum Covariance Determinant – MCD – due to [Rousseeuw et al., 2004]. The MCD estimator is commonly employed to identify robust estimates for the parameters of multivariate distributions and it has applications in many scientific areas. To cite some examples, the MCD is of common use in multivariate data analysis, and it is frequently used as input for other procedures likewise multivariate linear regression [Rousseeuw et al., 2004]; discriminant analysis [Hawkins and McLachlan, 1997] and factorial analysis [Pison et al., 2003]. More recently, the MCD has been extended to high–dimensional problems by [Boudt et al., 2018], with the Minimum Regularised Covariance Determinant (MRCD). The MRCD is suitable for outlier detection, observations ranking and clustering analysis [Boudt et al., 2018]. However, while the MCD has been mainly used outside finance, the shrinkage (SH) has played a crucial role in fostering the Markowitz portfolio optimisation framework. In particular, the SH procedure [Ledoit and Wolf, 2003], [Ledoit and Wolf, 2004a] revisits the concept of Stein estimators [Stein, 1956], proposing a series of alternatives for the target matrix, including: the covariance matrix as derived from the Single Index model [Sharpe, 1964], likewise in [Jorion, 1986], and the constant correlation matrix, i.e. a matrix where the pairwise correlation is treated as a constant, as in [Ledoit and Wolf, 2004b] who used a five factors model for comparison purposes against the shrinkage methodology for the covariance among equities in portfolios of different size. Furthermore, [Schäfer and Strimmer, 2005] extended the approach of [Jorion, 1986] to small size samples, and considered additional types of target matrix for the shrinkage. Finally, [Ledoit and Wolf, 2012] introduced the Nonlinear Shrinkage (NSH) technique and opened a new way for enhancing the shrinkage covariance estimator. This Section benefits from the previous rows review to focus on the following research question: which statistical estimator works at best to estimate the covariance matrix among interest rates? In detail, the chapter is structured as follows. Section 3.3.1 gives some brief remarks and basic analytics for the methodologies we employed to calculate the covariance matrix, and namely: Sample, Shrinkage, Nonlinear Shrinkage, MCD and MRCD estimators. A case study is presented in Section 3.3.2: following the seminal work of [Litterman and Scheinkman, 1991], in light of the robust PCA introduced in [Hubert et al., 2005], we analyse the US term structure curve through a robust PCA based on Sampling, SH, NS, MCD and MRCD approaches. To the best of our knowledge, the extension of MCD and MRCD estimators to this branch of finance is quite new, as well as the case study under discussion. Section 3.3.3 shows the empirical results while 3.3.4 concludes.

3.3.1 A Comparison of Covariance Estimation Techniques

Notations

From now on, we denote by Σ the true covariance matrix, by S the sample covariance matrix and by: S_{Shrink}, $S_{NLShrink}$, S_{MCD} and S_{MRCD} the covariance matrices obtained with the Shrinkage, Nonlinear Shrinkage, MCD and MRCD methodologies, respectively. We also denote by MSE($\cdot$) the Mean Squared Error of behind introduced matrices: to make an example, the MSE for the true covariance matrix is:

$$MSE(\Sigma) = Var(\Sigma) + Bias(\Sigma)^2$$

where $Bias(\Sigma)$ represents the bias induced by the estimation error and $Var(\Sigma)$ is the variance of the CVm. The covariance matrix, by definition, is a squared, symmetric, positive (semi–)definite matrix, with the variances on the main diagonal and the covariances elsewhere. Thus, it should be invertible and well-conditioned [Fisher and Sun, 2011].

Sample estimator

Let us denote by R a $p \times n$ matrix containing securities observations, where p is the number of securities and n the number of observations. Following [Briner and Connor, 2008], the sample covariance matrix S is given by:

$$S = \frac{1}{n-1} R(I - \frac{1}{n}\mathbf{1}\mathbf{1}')R', \tag{3.9}$$

where I is the identity matrix of dimensions p, 1 is the $p \times 1$ unitary vector, and $'$ indicates the transposition operator. The maximum rank of S is $n-1$. As said in Section 1, the sample estimator is unbiased and easy to estimate but contains a huge amount of estimation errors, because residing on a model-free approach, it does not require any structure for R, and therefore it is very sensitive to outliers in the dataset.

Shrinkage and Nonlinear Shrinkage Estimators

The rationale behind the introduction of the shrinkage procedure within the portfolio framework relates to the low accuracy of traditional estimators when describing the very underlying features of stocks, since their efficiency decreases more and more, as the sample size decreases [Briner and Connor, 2008]. Using the Shrinkage estimator makes possible to limit the potential error on estimates by reducing the Mean Squared Error of the SCVm. Furthermore, even if the SCVm is an unbiased estimator, it does not minimise the MSE, since MSE(S) is formed only by the variance, and $Bias(S) = 0$. However, [Stein, 1956] demonstrated how shrinkage estimators can reduce the MSE, and [Ledoit and Wolf, 2003], [Ledoit and Wolf, 2004b] improved this assertion as they highlighted that reducing the MSE can be reached by imposing some structure to the sample covariance matrix via a proper target matrix, T. In other words, instead of using a model-free approach, their covariance estimator is based on a convex linear combination between the sample covariance matrix S and a target matrix T:

$$S_{Shrink} = \delta T + (1-\delta)S, \tag{3.10}$$

where $\delta \in [0,1]$ is the shrinkage intensity. For $\delta = 1$, S_Shrink equals the target matrix, while for $\delta = 0$, we have: $S_{Shrink} = S$, i.e. the SCVm. In order to calculate the optimal shrinkage intensity [Ledoit and Wolf, 2004b] derived an optimal value δ^* based on minimising the expected value of the loss function given by the Frobenius norm of the quadratic distance between the true and the shrinkage covariance matrix:

$$L(\delta) = \|(\delta T + (1-\delta)S - \Sigma)^2\|, \tag{3.11}$$

When p is fixed and n tends to infinity, the optimal value $\delta*$ asymptotically behaves like a constant [Ledoit and Wolf, 2004b]. The estimate of the covariance matrix obtained with the shrinkage technique is always positive definite and well-conditioned, making it a good candidate for computational implementations. Moreover, shrinking the SCVm towards a more structured matrix makes the covariance matrix less sensitive to estimation errors; however, it can be extremely biased if the assumptions of the underlying model diverge from those of the true covariance. This trade-off is carefully described in [Jagannathan and Ma, 2003], asserting that the SH estimator looks like a compromise between the bias of the target matrix and the variability of the traditional SCVm. As seen in Section 1, we can find various characterisations of the target matrix T. In this work, we followed the approach of [Ledoit and Wolf, 2004b], assuming the target matrix being equal to the constant correlation matrix. This choice can be easily motivated in the following way: let us consider a set of N perspective interest rates with different maturity and same tenor, daily observed along a time horizon of length n. Since variations among interest rates are quite slow, we can assume that the relationships among different maturities do not change daily. Then, the covariance and correlation between the securities i and j are given by:

$$Cov(i,j) = S_{i,j} = \frac{1}{n-1}\sum_{h=1}^{p}(i_h - \bar{i})(j_h - \bar{j}),$$

$$Cor(i,j) = \rho_{i,j} = \frac{s_{i,j}}{s_i, s_j}, \tag{3.12}$$

where $\bar{i}$, $\bar{j}$ are the mean value for assets i and j, respectively. The average of the sample correlation is then given by:

$$\bar{\rho} = \frac{2}{(n-1)n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\rho_{i,j} \tag{3.13}$$

Putting together the last equations, by assuming the target F being the constant correlation matrix, we have:

$$f_{ii} = s_{ii} = s_i^2 \quad \text{and} \quad f_{ij} = \rho_{ij}s_i s_j, \tag{3.14}$$

where f_{ii} is the variance of every asset, lying on the main diagonal of F, and f_{ij} is the average of correlations between each couple of assets, elsewhere. This procedure generates a CVm with a more imposed structure and a lower estimation error. Combining F with S in the shrinkage procedure gives then an improved estimator of the true covariance matrix. However, the linear Shrinkage is just a first order approximation to the nonlinear problem of calculating sample eigenvalues [Ledoit and Wolf, 2012], so that each sample eigenvalue is shifted towards the grand mean of all sample eigenvalues with the same intensity. On the contrary, the rationale behind Nonlinear Shrinkage (NSH) is that different sample eigenvalues should be differently moved. [Ledoit and Péché, 2011] expanded the shrinkage concept to the nonlinear case, yet. However, they improved the linear estimator by constructing a target matrix based on the distribution of the sample eigenvalues, only: in this way, the target matrix is independent from the structure of the true covariance matrix, but generates an *oracle* estimator, which is reliable only in a very limited number of cases. Conversely, the *bona–fide* estimator in [Ledoit and Wolf, 2004b] is basically an *oracle* estimator, but consistently estimated: Monte Carlo stress tests highlighted that this estimator is at least as good as the 2004 Shrinkage procedure of [Ledoit and Wolf, 2004b] or even better, making it an improved candidate for portfolio optimisation. From the analytic viewpoint, the NSH procedure aims at calculating:

$$S_{NLShrink} = YDY'$$

where Y is the matrix formed by the eigenvectors $y_{(i,p)}$ of the sample matrix S, and D is a diagonal matrix whose elements capture the link between the eigenvectors and the true covariance matrix. The procedure consists in two steps that are below summarised.

Algorithm 3.1: Nonlinear Shrinkage.

1 Find the matrix A that is closest to the true covariance matrix Σ according to the Frobenius norm.

2 Solve the minimisation problem:

$$\min_{D} \|YDY' - A\|,$$

to find $D^* = Diag(d_1^*, d_2^*, ..., d_p^*)$, with $d_i^* = \mathbf{y}_i' \Sigma \mathbf{y}_i, i = 1, ..., p$.

Minimum Covariance Determinant and Minimum Regularised Covariance Determinant

The Minimum Covariance Determinant is a robust covariance estimator that allows detecting outliers in multivariate data by calculating the Mahalanobis distances between every observation and the central value of the data [Rousseeuw and Yohai, 1984]. More precisely, the aim of the estimation process is to find a subset of the data with the lowest value of the determinant and containing h observations, with: $(p + n + 1)/2 < h < n$ where p and n are as defined in previous rows. The choice of h depends only on the covariance between the data, thus the MCD covariance matrix is given by the SCVm of the subset of the original data that minimises the dispersion of the observations. Being R the $p \times n$ matrix with p components $r_i \in \Re^n, i = 1, ..., p$, each vector $r_i(i = 1, \ldots, p)$ can be viewed as the random variable whose realisations are the n observations of the i–th asset. The procedure to obtain S_{MCD} can be summarised into five steps, as in Algorithm 3.2.

However, applying the MCD estimator has remained difficult for many years, because too computationally expensive. Branch and bounds algorithms [Candela, 1996], heuristics [Woodruff and Rocke, 1994] and relaxation techniques of the exact solution [Schyns, 2008] have been suggested. The FAST–MCD algorithm developed by [Rousseeuw and Driessen, 1999] tackled down the problem, making the MCD computation more efficient. FAST–MCD is a deterministic procedure which produces a good approximation of the MCD for both small and large datasets. The main idea consists in considering a small random subset of dimension N+1, instead of looking for the subset of h. The procedure replicates the steps already highlighted in the MCD algorithm, working on the new reduced dimensions subset. The resulting subset HFAST–MCD is then locally improved thanks to sequential concentration steps (C–steps) in which the mechanism of computation of the Mahalanobis distances and outlier elimination is repeated until any further reduction of the determinant of the covariance matrix becomes unfeasible.

In 2017, [Boudt et al., 2018] proposed an alternative estimator called Minimum Regularised Covariance Determinant estimator (MRCD). The MRCD is still based on subsets of the covariance matrix, this time selected through a convex combination between a target matrix and the SCVm. The main steps leading to the MRCD are illustrated in Algorithm 3.3.

Algorithm 3.2: Minimum Covariance Determinant.

---

1 Identify an initial subset H_0 of h observations, with $(p+n+1)/2 < h < n$:

$$H_0 = \text{argmin}(\det(\text{cov}(\mathbf{r}_i | i \in H_0))), \tag{3.15}$$

From H_0 we can evaluate the sample mean vector $\boldsymbol{\mu}_0$ and the sample covariance matrix S_0 than can be then employed as indicators of location and scatter.

2 Compute the Mahalanobis distances between each component of R and μ_0:

$$d_i = Dist_{S_0}(\mathbf{r}_i, \mu_0) = \sqrt{(\mathbf{r}_i - \mu_i)' S_0^{-1} (\mathbf{r}_i - \mu_i)}$$

where S_0^{-1} is the inverse of S_0, and $d_i \in \Re$.

3 The assets whose distance is behind the acceptance region are assigned a weight equal to zero and henceforth excluded, while those whose distance falls inside the acceptance region are kept and receive a weight equal to one. Weights are assigned with the following:

$$w_i = \begin{cases} 0, d_i > \sqrt{\chi^2_{N,0.975}} \\ 1, d_i \leq \sqrt{\chi^2_{N,0.975}} \end{cases} \tag{3.16}$$

where $\sqrt{\chi^2_{N,0.975}}$ represents the cut-off value for detecting outliers, and $\sqrt{\chi^2_{N,0.975}}$ is the 0.975 quantile of the χ^2_N distribution. Thus, a new subset H_1 is derived with centre μ_1 and scatter S_0.

4 Using the weights derived by 2.15, the estimators of position and scatter $\boldsymbol{\mu}_{MCD}$ and S_{MCD} are computed:

$$\mu_{MCD} = \frac{\sum_{i=1}^{n} w_i \mathbf{r_i}}{\sum_{i=1}^{n} w_i},$$

$$S_{MCD} = \frac{c_1 (\sum_{i=1}^{n} w_i (\mathbf{r}_i - \boldsymbol{\mu}_{MCD})(\mathbf{r}_i - \boldsymbol{\mu}_{MCD})')}{\sum_{i=1}^{n} w_i},$$

Here c_1 is a constant that ensures the asymptotic consistency towards a normal distribution (Croux and Haesbroeck, 1999), so that the robust Mahalanobis distances for each asset become:

$$Rd_i = Dist_{S_{MCD}}(\mathbf{r}_i, \mu_{MCD}) = \sqrt{(\mathbf{r}_i - \boldsymbol{\mu}_{MCD})' S_{MCD}^{-1} (\mathbf{r}_i - \boldsymbol{\mu}_{MCD})}.$$

5 The final step consists then in marking as outliers and excluding the observations outside the acceptance region, i.e. those for which we have:

$$w_{(Rd_i)} = \{\mathbf{r}_i | Rd_i \geq \sqrt{\chi^2_{p,0.975}}\}.$$

Algorithm 3.3: Minimum Regularised Covariance Determinant.

1 The dataset is standardised according to the:

$$\mathbf{u}_i = P_R^{-1}(\mathbf{r}_i - \boldsymbol{\nu}_R), (i = 1, \ldots, p),$$

where $P_R^(- 1)$ is the $p \times p$ diagonal matrix containing the scatter estimations, $\mathbf{r}_i$ are the components of R, and $\boldsymbol{\nu}_R$ is the $p \times 1$ median vector.

2 Compute:

$$K(H) = kF + (1-k)c_\alpha S_U(H)$$

where k is the shrinkage weight or regularisation parameter, F the target matrix, $S_U(H)$ the sample covariance matrix calculated on $U = \{u_i\}$ in the subset H, and c_α a consistency factor as defined in [Croux and Haesbroeck, 1999], with $\alpha = (n-h)/n$ being the so-called trimming percentage.

3 Find the subset of the original dataset minimising the determinant of the regularised matrix $K(H)$.

$$H_{MRCD} = \underset{H}{\text{argmin}}(det(K(H)^{1/p})),$$

4 Compute the MRCD covariance matrix:

$$K_{MRCD} = P_r Q \Lambda^{1/2}[kI + (1-k)S^*(H_{MRCD})]\Lambda^{1/2} Q P_r,$$

where S^* is diagonal matrix ad–hoc computed [a] that rescales the diagonal elements of the final covariance matrix $K_M RCD$, k is the regularisation parameter, Q is the orthogonal matrix of eigenvectors from the target matrix and $\Lambda^{1/2}$ is the square root eigenvalues matrix.

[a] As explained in [Boudt et al., 2018], S^* is a transformation of the sample covariance matrix $S_U(H)$ which undergoes first a multiplication for its eigenvectors and squared eigenvalues, then the resulting matrix is rescaled by its own diagonal elements.

3.3.2 Empirical Case Study: PCA of the US Yield Curve

It is generally acknowledged that bond prices are sensitive not only to parallel shifts in the yield curve, but also to non-parallel shifts; in particular, [Litterman and Scheinkman, 1991] found out that bond prices are mainly sensitive to three factors, which can explain almost the 99% of the total variance: level, steepness, and curvature. Evaluating the exposure to these three factors is the leading feature of Litterman and Scheinkman's (LS, thereinafter) approach to hedging, and it is generally ruled out by applying the Principal Component

Analysis (PCA) on the covariance matrix of changes in interest rates, usually estimated via the sample estimator.

Principal Component Analysis [Hotelling, 1933] is a dimension reduction technique that works on a covariance (or correlation) matrix identifying the volatility factors that drive the time series under investigation. The PCA relies on the spectral decomposition of the covariance matrix Σ:

$$\Sigma = G\Omega G',$$

where G is the square matrix of the eigenvalues of Σ, and Ω is a diagonal matrix filled with the eigenvalues of the covariance matrix. The principal components are given by the normalised eigenvectors, ranked in descendant order according to the size of related eigenvalues. This because the total variance is equal to the sum of all the eigenvalues, so that the size of a single eigenvalues is the percentage of total variance explained. As a limited number of eigenvalues is usually enough to explain at least the 99% of total variation, the reduction of the covariance matrix can be performed by retaining only the eigenvalues that explain a certain threshold of the variance, eliminating the others.

With PCA, the LS approach can be characterised as follows: the first Principal Component (PC) should equally affect all the maturities in the term structure, and should be regarded as the response in shifts of the term structure. The second PC response should look like an upward sloping curve: it should affect closer maturity with the same intensity, but with different sign. This should be regarded as the response in changes of the slope of term structure. The third PC should affect in the same way the extremes of the term structure, with a change of sign in the middle maturities. This should be regarded as the response in changes of the curvature of the term structure. Robust methodologies in estimating the covariance matrix might improve PCA results, hence fostering the reliability of LS hedging approach. To such aim, we checked the LS assertion considering the US term structure with 3–months tenor composed by the instruments highlighted in Table 3.10.

[Table 3.10 about here.]

The dataset consists of 24 time series composed by daily data for the period: 02/01/2014 – 08/09/2017, for an overall number of 962 observations. The behaviour of the 24 time series is depicted in Figure 3.4.

[Figure 3.4 about here.]

In order to find relationships among the term structure components, we estimated the covariance matrix of daily changes in the spot prices of the 24 curves under examination, testing five different covariance estimators: Sample, Shrinkage, Nonlinear Shrinkage, MCD and MRCD.

3.3.3 Empirical Results

The estimates highlighted how the quality of the sample covariance matrix tends to deteriorate increasing the number of interest rate curves from 1 to 24, as testified by looking at the condition number, i.e. the ratio between the largest and the smallest eigenvalue, as depicted in Figure 3.5.

[Figure 3.5 about here.]

Looking at the plots in Figure 3.5, we can highlight that the sample covariance method appears being very sensitive to the matrix size, when the number of examined interest rates is greater than eight; while the Shrinkage, MCD and MRCD adapt well in estimating the covariance for high–dimensional arrays. The Nonlinear Shrinkage tends to deteriorate the performance behind the same "magic" threshold as the sample approach, but it stabilises after the 15^{th} interest rate to converge to zero. As a preliminary conclusion, we can then argue that estimating the covariance matrix with a statistical estimator likewise SH, MCD and MRCD should lead to more robust results than in the case of both the sample covariance estimator and NS. We then run the PCA on the five covariance matrices, obtained with the above-mentioned estimators, and we monitored the percentage of explained variance by different and uncorrelated risk factors. According to the LS approach, this should lead to find 3 factors explaining the 99% of the overall variance; for this reason, we set the value three as a threshold for our analysis. Results are displayed in Table 3.11, where we list earlier five factors sorted by percentage of explained variance. Values represents the cumulative variance in percentage.

[Table 3.11 about here.]

Looking at the results, the sample and MCD estimators are the fastest to reach the 99% threshold of explained variance, with just 2 factors. Nonlinear Shrinkage is also very fast, with 3 factors, confirming the LS view. On the other hand, MRCD and Shrinkage are very slow, taking 7 and 15 factors to explain the threshold, respectively. The relation between the change in spot rates and different maturities should be captured by the number of the above listed factors for each methodology. However, there is no economic or financial motivation for explaining the total variation of the US term structure with more than three factors. Thus, in order to have a common ground of comparison with the LS model, in Figure 3.6 we plot the sensitivity of changes in interest rate to increasing maturity, as explained by earlier three PCs for all the covariance methodologies in use.

[Figure 3.6 about here.]

The slowest methodologies in explaining the total variation of the US term structure (SH and MRCD) clearly fail in delivering good insights about the risk factors afflicting the US term structure. On the other hand, the sample, NSH and MCD present results in line with the LS model.

A risk manager should find these results useful, because all those three methodologies (sample, Nonlinear Shrinkage and MCD) characterise in the same way the response to earlier two risk factors: the first PC (in blue), in fact, expresses the sensitivity to parallel shifts, while the second (orange) the change in the slope; moreover, the magnitude is the same across all the maturities but the sign changes. However, both sample and Nonlinear Shrinkage fail to deliver a good representation of the third component, which is well represented only by the MCD methodology. The US term structure, in fact, when the estimation is performed with either sample or Nonlinear Shrinkage estimator, shows a strange peak between 10 to 12 years maturity: this seems to be originated from some outliers in the term structure. The MCD methodology, on the other hand, confirms the LS assertion also for the third component: it has same sign for both closer and long-term maturities, while in the medium term the sign is different. This should be regarded as the response in changes of the curvature of the term structure.

In conclusion, we found that three out of five covariance methodologies, namely sample, Nonlinear Shrinkage and MCD, are able to represent the US term structure response with respect to the three factors introduced by the LS model. In particular, the MCD is the methodology that achieves more aligned results to the conclusions stated by the LS model. Looking in conjunction at the results of Table 3.11 and Figure 3.5 we can then argue that the MCD arise as the best estimation method among those examined, as it maintains well-conditioned the covariance matrix, and offers results well-fitting to the LS assertion.

3.3.4 Conclusion

In this work we compared various statistical methodologies aimed at providing a robust estimate of the covariance matrix, namely: Shrinkage, Nonlinear Shrinkage, Minimum Covariance Determinant (MCD) and Minimum Covariance Regularised Determinant (MRCD) estimators. Results were evaluated using the sample covariance matrix estimator as benchmark. These techniques were tested within the fixed income framework. To this extent, we analysed the performances of the examined estimators in evaluating the US term structure with the PCA technique, according to what stated by [Litterman and Scheinkman, 1991]. Strong evidences highlight the benefit of switching from the sample covariance matrix to robust variants. First, by looking at the condition numbers, we highlighted that while the sample covariance estimator is very sensitive to the matrix size, on the other hand, the Shrinkage, MCD and MRCD well adapt in estimating the covariance for high–dimensional arrays, which in our example are composed by 24 interest rates. Second, the PCA approach seems to give more precise results when applied on the robust estimates: MCD and Nonlinear Shrinkage explain the 99% of total variance in a slower way than the sample methodology, thus highlighting that the second and the third factor represent a not negligible percentage of the overall variation. In particular, the MCD is the only methodology that well characterised the third factor too: this is crucial especially for risk management applications, as it expresses the response in changes of the curvature of the term structure. To conclude, we can state that statistical covariance estimators can help in modelling the factors that drive the term structure curve, as they make possible to highlight relationships among different maturities in a fashion which is less likely to be biased by outlier observations. In detail, the PCA approach is enhanced, and this in turn should signify a better hedging power for financial products relying on fixed income instruments. Future works are planned to assess the impact of such robust estimators on a wider range of fixed income instruments applications including risk and portfolio management.

3.4 Conclusions

In this chapter we introduced the Minimum Regularised Covariance estimator in the financial literature. First, we employed its closed form solution for the precision matrix to enhancing the out–of–sample performances and weights stability in the Global Minimum Variance portfolio. This was accomplished trough an extensive Monte Carlo experiment, where we evaluated the benefits of employing the MRCD against the sample covariance matrix under the Global Minimum Variance portfolio. Moreover, we tested the empirical soundness of the MRCD even with real data, comparing its out–of–sample performance with an empirical illustration featuring five investment universes. Both the simulated and empirical applications clearly demonstrated that the out–of–sample performance of the GMVP benefited from the use of the MRCD estimator: results suggested a reduction in the portfolio turnover at no cost for the portfolio variance, whilst portfolio expected returns increased.

Second, we introduced the MRCD into the fixed income space. We compared various methodologies to estimate the interest rates covariance matrix. Adopting a statistical approach for the robust estimation of this object, we compared the sample covariance matrix, the Shrinkage, the Nonlinear Shrinkage, the Minimum Covariance Determinant against the MRCD methodology. The comparison revolved around a practical application aimed at individuating the principal risk factors of the US term structure curve. Results confirmed fixed income portfolio construction and risk management can benefit from the use of robust statistical methodologies for the estimation of the covariance matrix.

In conclusion, the Minimum Regularised Covariance estimator showed a great applicability for financial applications. Results in this chapter can be extended towards many directions for future research.

Table 3.1: Sample sizes considering $p = 100$ assets. Frequency is monthly. Total observations are $n = 12000$ everywhere.

Id.	**In-sample length**	**Out-of sample length** $(n - m)$	**Concentration ratio** $(c = p/n)$
S1	360	11640	0.27
S2	180	11820	0.55
S3	90	11910	1.11
S4	60	11940	1.66
S5	36	11964	2.77

Table 3.2: Out-of-sample portfolio turnovers with simulated data. Values are averages across the $n - m$ out-of-sample observations, as described in Table 3.1.

	S1	**S2**	**S3**	**S4**	**S5**
Panel A: Sample					
$\xi = 0$	0.0356	0.1263	1.5314	0.6047	0.5879
$\xi = 2.5\%$	0.0356	0.1248	1.5414	0.5936	0.6003
$\xi = 5\%$	0.0350	0.1248	1.5223	0.6213	0.5731
$\xi = 10\%$	0.036	0.1237	1.6005	0.6055	0.5728
Panel B: MRCD with $h = 0.5$					
$\xi = 0$	0.2858	1.4296	0.7526	0.4027	0.3028
$\xi = 2.5\%$	0.2085	1.0710	0.6311	0.3714	0.2375
$\xi = 5\%$	0.1927	0.9086	0.6112	0.3934	0.2272
$\xi = 10\%$	0.1515	0.8954	0.6734	0.3517	0.2428
Panel C: MRCD with $h = 0.75$					
$\xi = 0$	0.1382	0.7611	0.6233	0.5653	0.3844
$\xi = 2.5\%$	0.075	0.5042	0.4755	0.4237	0.3483
$\xi = 5\%$	0.0743	0.4776	0.4811	0.4072	0.3578
$\xi = 10\%$	0.1046	0.4672	0.5109	0.4606	0.3547
Panel D: MRCD with $h = 0.9$					
$\xi = 0$	0.0811	0.3695	0.7725	0.3708	0.3269
$\xi = 2.5\%$	0.0697	0.2528	0.6947	0.278	0.2899
$\xi = 5\%$	0.0643	0.2624	0.6802	0.3166	0.3076
$\xi = 10\%$	0.0776	0.2674	0.6929	0.344	0.3033

Table 3.3: Out-of-sample portfolio variances with simulated data. Values are averages across the $n - m$ out-of-sample observations, as described in Table 3.1.

	S1	**S2**	**S3**	**S4**	**S5**
Panel A: Sample					
$\xi = 0$	0.0105	0.0086	0.0049	0.014	0.0307
$\xi = 2.5\%$	0.0131	0.0111	0.006	0.0171	0.0386
$\xi = 5\%$	0.0156	0.0126	0.0072	0.0205	0.0429
$\xi = 10\%$	0.0199	0.015	0.0089	0.0248	0.0513
Panel B: MRCD with $h = 0.5$					
$\xi = 0$	0.0157	0.0239	0.0126	0.0105	0.0097
$\xi = 2.5\%$	0.0185	0.0307	0.0166	0.0129	0.0119
$\xi = 5\%$	0.0218	0.0344	0.0197	0.0159	0.0143
$\xi = 10\%$	0.026	0.0375	0.0232	0.0188	0.0177
Panel C: MRCD with $h = 0.75$					
$\xi = 0$	0.0118	0.0155	0.0121	0.0088	0.0074
$\xi = 2.5\%$	0.0154	0.0203	0.0162	0.0122	0.0084
$\xi = 5\%$	0.0184	0.023	0.0197	0.0155	0.0096
$\xi = 10\%$	0.0231	0.0275	0.0231	0.0174	0.0115
Panel D: MRCD with $h = 0.9$					
$\xi = 0$	0.011	0.011	0.0099	0.0063	0.005
$\xi = 2.5\%$	0.0145	0.0148	0.0144	0.0102	0.0063
$\xi = 5\%$	0.0173	0.0169	0.0176	0.0127	0.0075
$\xi = 10\%$	0.0218	0.0204	0.0208	0.0145	0.0088

Table 3.4: Out-of-sample Sharpe ratios with simulated data. Values are averages across the $n - m$ out-of-sample observations, as described in Table 3.1.

	S1	S2	S3	S4	S5
Panel A: Sample					
$\xi = 0$	0.4031	0.3528	0.2015	0.284	0.2365
$\xi = 2.5\%$	1.1214	0.9374	0.3918	0.6986	0.5888
$\xi = 5\%$	1.7172	1.3408	0.6654	0.9128	0.9521
$\xi = 10\%$	2.3889	1.99	0.9225	1.5307	1.4587
Panel B: MRCD with $h = 0.5$					
$\xi = 0$	0.3627	0.2219	0.3989	0.4353	0.4593
$\xi = 2.5\%$	0.9818	0.5991	1.0184	1.2319	1.3275
$\xi = 5\%$	1.5148	0.8897	1.5549	1.7298	1.9068
$\xi = 10\%$	2.2241	1.4545	2.2825	2.6832	2.7272
Panel C: MRCD with $h = 0.75$					
$\xi = 0$	0.4307	0.2487	0.3415	0.3884	0.4482
$\xi = 2.5\%$	1.1497	0.7289	0.8035	1.1193	1.2578
$\xi = 5\%$	1.6564	1.0742	1.2447	1.5689	1.7118
$\xi = 10\%$	2.3838	1.6301	1.9011	2.3542	2.5007
Panel D: MRCD with $h = 0.9$					
$\xi = 0$	0.4265	0.3226	0.2812	0.3767	0.4344
$\xi = 2.5\%$	1.1345	0.8773	0.6428	1.0566	1.2023
$\xi = 5\%$	1.6996	1.3007	0.9235	1.4349	1.6568
$\xi = 10\%$	2.4473	1.9429	1.465	2.1547	2.457

Table 3.5: List of observable datasets, monthly frequency. For each dataset we provided a short Description, the Id, the number of assets (p), the number of observations (n) the range (Time Interval) and the Source.

No.	Description	Id.	p	n	Time Period	Source
1	25 Operating Profitability and Investment portfolios	25OPI	25	661	01/07/1963 – 01/08/2018	K.French
2	25 Size and Book to Market portfolios	25SBM	25	1106	01/07/1926 – 01/08/2018	K.French
3	25 Size and Momentum portfolios	25SM	25	1106	01/11/1926 – 01/08/2018	K.French
4	30 Industry portfolios portfolios	30I	30	1106	01/07/1926 – 01/08/2018	K.French
5	300 stocks from S&P500	300SPX	300	269	01/01/1996 – 01/06/2018	Bloomberg

Table 3.6: Investment universes statistical properties.

Universe	**Min.**	**Max.**	**Mean**	**Median**	**St.Dev.**	**Skew.**	**Kurt.**
25OPI	-0.0878	0.0840	0.1349	0.1100	1.0444	-0.1024	9.1507
25SBM	-0.3000	0.3750	0.2127	0.2000	1.3813	0.4033	21.7706
25SM	-0.1664	0.2173	0.2231	0.2000	1.2918	0.8627	22.8392
30I	-0.2925	0.6108	0.1600	0.1300	1.3315	0.5294	25.9098
300SPX	-0.0186	0.0152	0.0071	0.0104	0.0947	-0.5710	7.6663

Table 3.7: Out-of-sample portfolio turnovers with empirical data.

	25OPI	25SBM	25SM	30I	300SPX
Panel A: m= 360					
Sample	0.1356	0.1648	0.1784	0.1063	-
MRCD $h = 0.5$	0.6144	0.6899	0.6976	0.4093	-
MRCD $h = 0.75$	0.4265	0.5084	0.4996	0.2883	-
MRCD $h = 0.9$	0.4526	0.5317	0.5005	0.2351	-
Panel B: m= 180					
Sample	0.2852	0.3193	0.3344	0.2046	1.6568
MRCD $h = 0.5$	1.1843	1.3381	1.3296	0.7453	1.2869
MRCD $h = 0.75$	0.8391	0.9609	0.9952	0.5386	1.2869
MRCD $h = 0.9$	0.7904	0.8135	0.8483	0.4392	0.7445
Panel C: m= 90					
Sample	0.6578	0.6845	0.729	0.4578	0.7703
MRCD $h = 0.5$	3.1556	2.9723	2.9167	2.4086	0.8086
MRCD $h = 0.75$	1.6775	1.6995	1.7406	1.2163	0.8086
MRCD $h = 0.9$	1.4163	1.2995	1.4451	0.8533	0.5716
Panel D: m= 60					
Sample	1.2195	1.2288	1.2911	0.9396	0.6756
MRCD $h = 0.5$	5.0113	4.1764	4.4666	3.4872	0.7003
MRCD $h = 0.75$	2.8484	2.5903	2.7128	2.2013	0.7003
MRCD $h = 0.9$	2.0792	1.9105	1.9506	1.3966	0.5205
Panel E: m= 36					
Sample	3.4767	3.6021	3.7485	4.381	0.599
MRCD $h = 0.5$	4.6199	4.2094	4.2259	2.2994	0.6373
MRCD $h = 0.75$	4.1951	3.6796	3.9437	2.8756	0.6373
MRCD $h = 0.9$	3.1234	2.6077	2.8495	2.1413	0.5381

Table 3.8: Out-of-sample portfolio variance with empirical data.

	25OPI	**25SBM**	**25SM**	**30I**	**300SPX**
Panel A: m= 360					
Sample	0.7212	0.557	0.5685	0.4975	-
MRCD $h = 0.5$	0.7948	0.638	0.6702	0.5958	-
MRCD $h = 0.75$	0.7858	0.627	0.6432	0.5741	-
MRCD $h = 0.9$	0.7771	0.6169	0.6321	0.5706	-
Panel B: m= 180					
Sample	0.6577	0.5241	0.5561	0.4507	0.013
MRCD $h = 0.5$	0.8159	0.6682	0.7304	0.6095	0.0088
MRCD $h = 0.75$	0.7748	0.6432	0.6928	0.5764	0
MRCD $h = 0.9$	0.7449	0.6217	0.6837	0.5617	0.0096
Panel C: m= 90					
Sample	0.5518	0.4563	0.5065	0.3865	0.0195
MRCD $h = 0.5$	0.9219	0.7576	0.8717	0.8184	0.0087
MRCD $h = 0.75$	0.7939	0.6692	0.7429	0.6626	0.0087
MRCD $h = 0.9$	0.728	0.6364	0.7026	0.6154	0.0046
Panel D: m= 60					
Sample	0.469	0.3918	0.4428	0.3262	0.0239
MRCD $h = 0.5n$	0.9853	0.8313	0.9287	0.8616	0.0076
MRCD $h = 0.75n$	0.8052	0.6824	0.7663	0.7058	0.0076
MRCD $h = 0.9n$	0.7167	0.6223	0.6972	0.6016	0.0038
Panel E: m= 36					
Sample	0.311	0.2623	0.2905	0.1682	0.0306
MRCD $h = 0.5$	0.8571	0.7706	0.8679	0.7055	0.0071
MRCD $h = 0.75$	0.7472	0.6908	0.7416	0.6912	0.0071
MRCD $h = 0.9$	0.6634	0.6055	0.6581	0.5877	0.0036

Table 3.9: Out-of-sample portfolio turnovers with empirical data.

	25OPI	25SBM	25SM	30I	300SPX
Panel A: m= 360					
Sample	0.0313	0.2126	0.1507	0.2006	-
MRCD $h = 0.5$	0.0508	0.2334	0.2082	0.2309	-
MRCD $h = 0.75$	0.0749	0.2362	0.2168	0.2341	-
MRCD $h = 0.9$	0.0643	0.2432	0.2189	0.2368	-
Panel B: m= 180					
Sample	0.0913	0.2508	0.1765	0.1963	0.1811
MRCD $h = 0.5$	0.1364	0.2224	0.2438	0.2278	0.2567
MRCD $h = 0.75$	0.1366	0.2593	0.2315	0.2418	0.2567
MRCD $h = 0.9$	0.1184	0.2618	0.2189	0.2265	0.2786
Panel C: m= 90					
Sample	0.0855	0.2411	0.1952	0.1968	0.1701
MRCD $h = 0.5$	0.1163	0.2042	0.2128	0.1601	0.1734
MRCD $h = 0.75$	0.119	0.2494	0.1851	0.1953	0.1734
MRCD $h = 0.9$	0.095	0.2634	0.2098	0.1757	0.1746
Panel D: m= 60					
Sample	0.0809	0.1697	0.1181	0.1788	0.1548
MRCD $h = 0.5$	0.1522	0.1489	0.1204	0.107	0.2169
MRCD $h = 0.75$	0.0936	0.1688	0.1311	0.135	0.2169
MRCD $h = 0.9$	0.1086	0.1923	0.1208	0.1676	0.1907
Panel E: m= 36					
Sample	0.0512	0.126	0.0831	0.0881	0.1489
MRCD $h = 0.5$	0.0954	0.1729	0.1666	0.1461	0.1583
MRCD $h = 0.75$	0.0639	0.1737	0.1098	0.1547	0.1583
MRCD $h = 0.9$	0.0588	0.1776	0.1121	0.1733	0.1412

Table 3.10: US term structure composition for the examined instruments.

Instrument	Maturity	Total
USD Libor	Overnight, 1 week, 1, 2, 3, 6, 12–months	7
USD Swap	2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25 and 30–years	17

Table 3.11: Earlier five factors explaining the overall variance, for various estimation methodologies of the covariance matrix.

Factor	**SCV**	**SH**	**NSH**	**MCD**	**MRCD**
Panel A					
F1	97.50	75.41	97.18	98.84	81.49
F2 A	99.28	78.03	98.95	99.07	95.88
F3 A	99.52	80.57	99.22	99.37	97.09
F4 A	99.68	82.96	99.37	99.60	97.97
F5 A	99.76	85.20	99.48	99.72	98.54
Panel B					
Factors	2	15	3	2	7

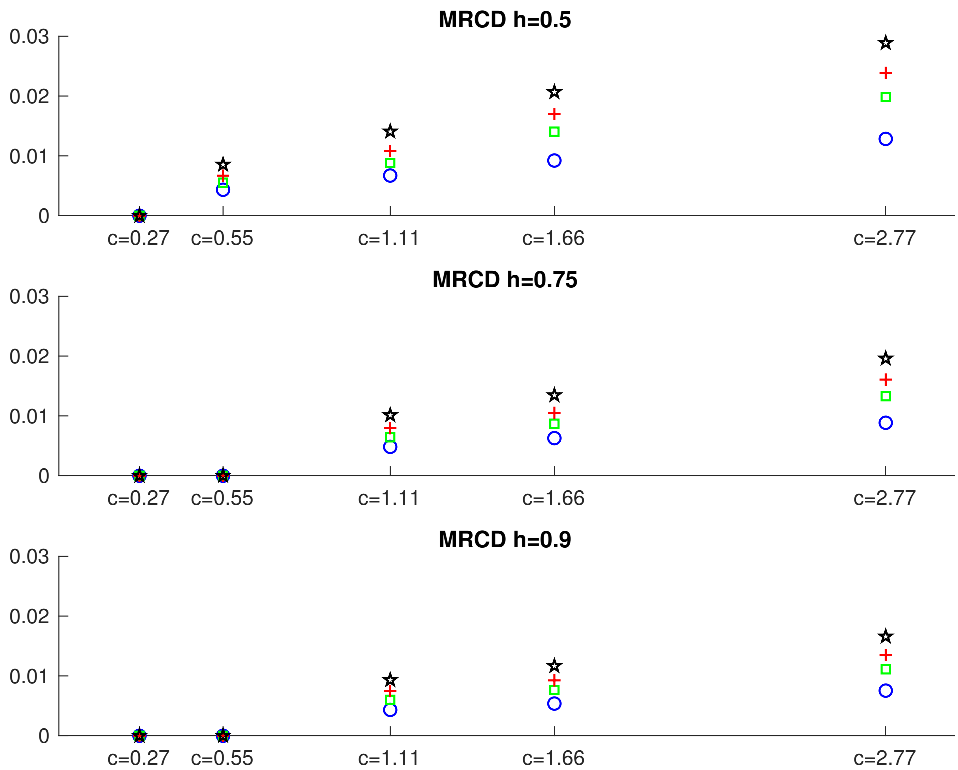

Figure 3.1: The behaviour of $\hat{\rho}$ (vertical axis) varying the concentration ratios (horizontal axis) with simulated data. Colours and shapes correspond to different levels of dataset contamination: blue circle ($\xi = 0\%$), green square ($\xi = 2.5\%$), red cross ($\xi = 5\%$) and black pentagram ($\xi = 10\%$). Values are averages along the $n - m$ out-of-sample observations, as described in Table 3.1. From the top to the bottom each panel corresponds to a different level of h in ascendant order

.

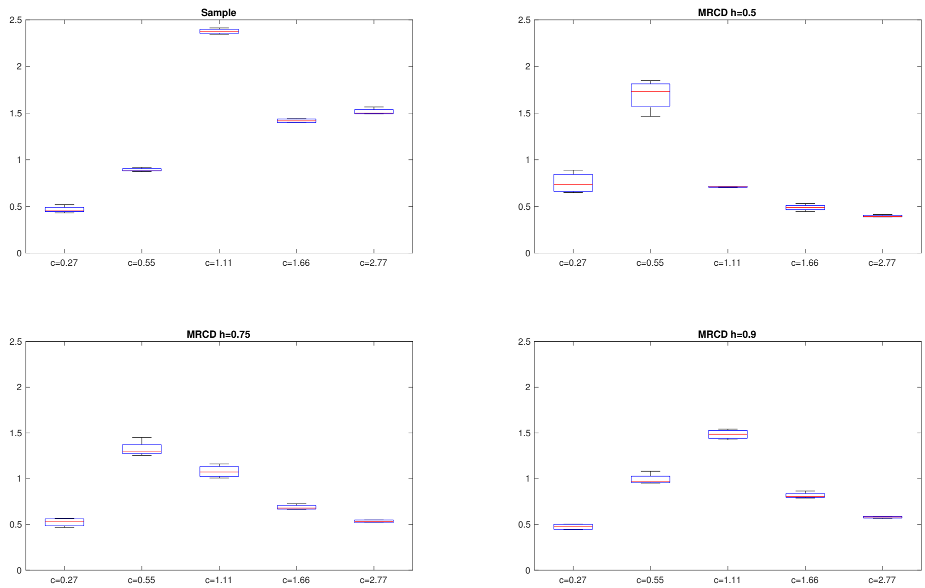

Figure 3.2: Weights distance with simulated data. Results are averages across the $n - m$ out-of-sample observations, as described in Table 3.1.

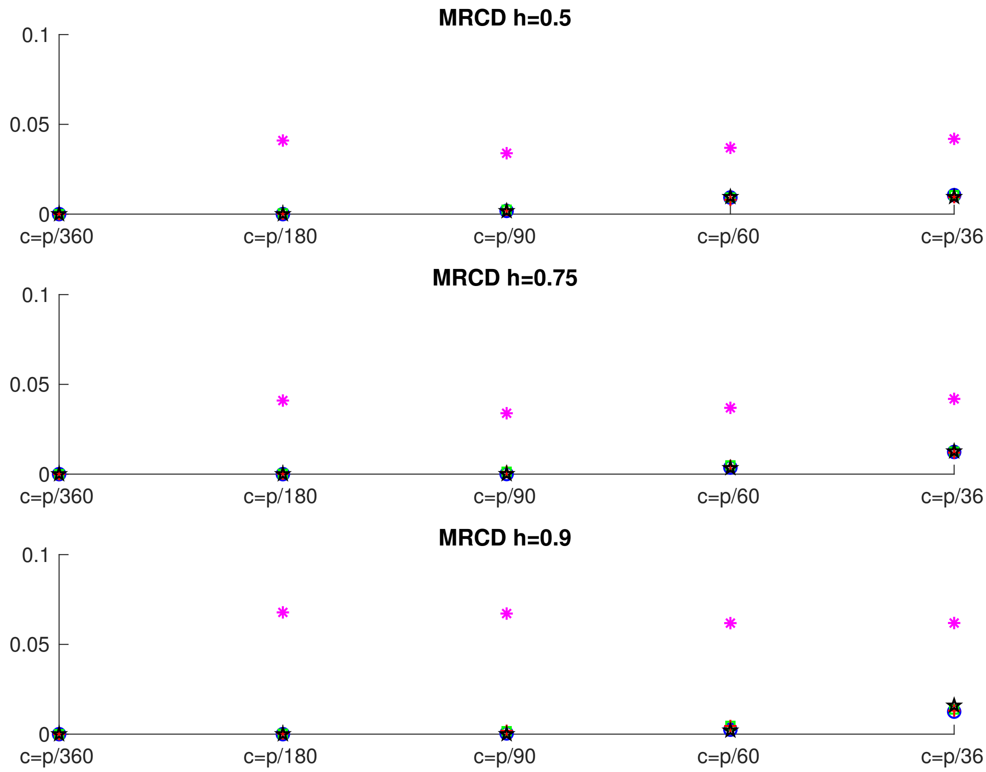

Figure 3.3: The behaviour of $\hat{\rho}$ (vertical axis) varying the concentration ratios (horizontal axis) with empirical data. Colours and shapes correspond to different asset universes: 25OPI is represented with the blue circle, 25SBM by the green square, 25SM with the red cross, 30I with the black pentagram and 300SPX by the magenta star. Each panel corresponds to a different level of h in ascendant order from the top to the bottom. Note that values at $m = 360$ are omitted for the 300SPX universe dataset as not applicable.

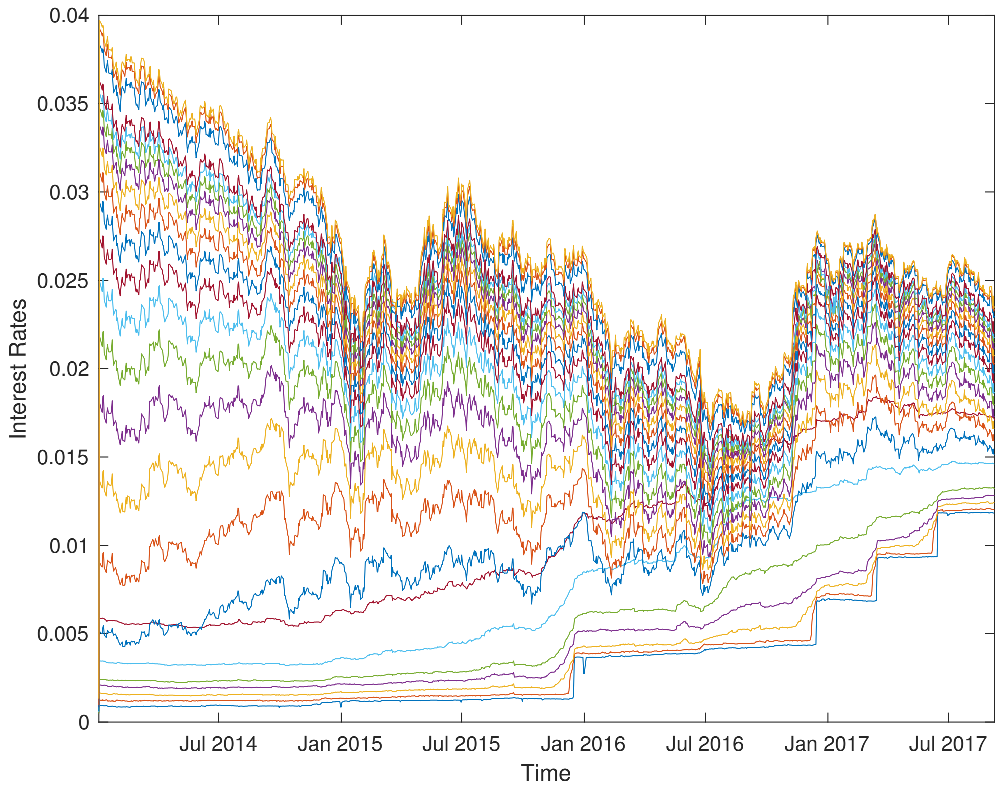

Figure 3.4: Daily spot interest rates for the period 02/01/2014 – 08/09/2017. From bottom to top, the plot depicts the behaviour of: USD LIBOR with maturity overnight, 1 week, 1, 2, 3, 6, 12 months, and USD Swap with maturity 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 ,15, 20, 25 and 30-year.

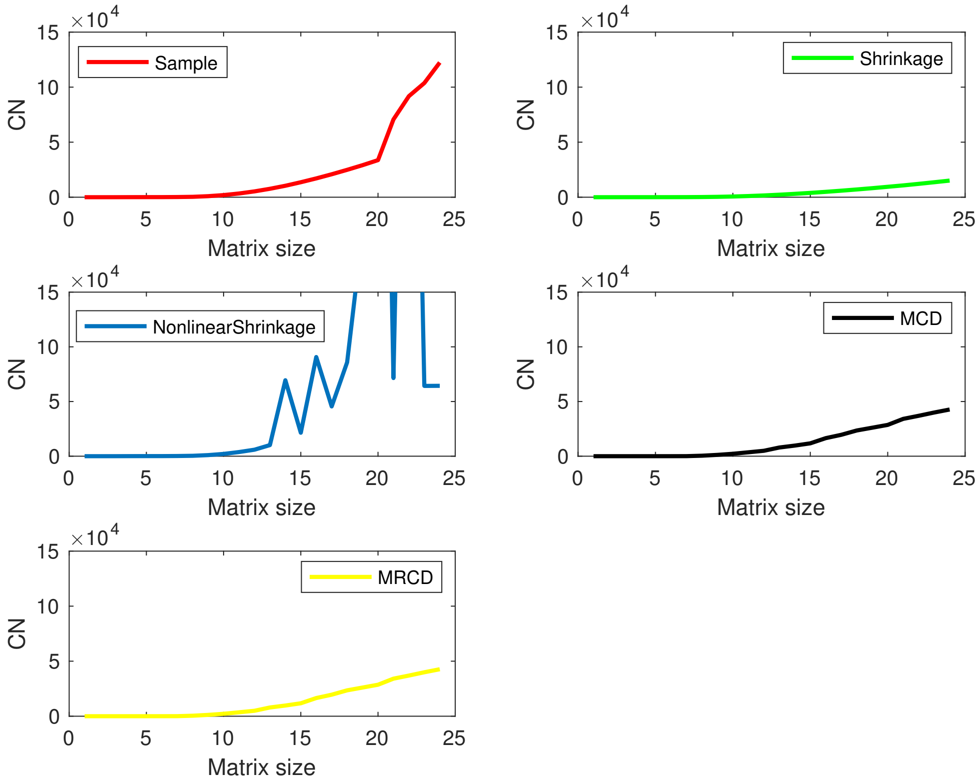

Figure 3.5: Condition Number (CN) varying the size of the covariance matrix. From left to right: CN behaviour for the Sample covariance matrix (SCV - top-left), Shrinkage (SH - top-right), Nonlinear Shrinkage (NSH - middle-left), MCD (middle-right) and for MRCD (bottom-left). On the x-axis we reported the matrix size, while the y-axis reports the value of the condition number.

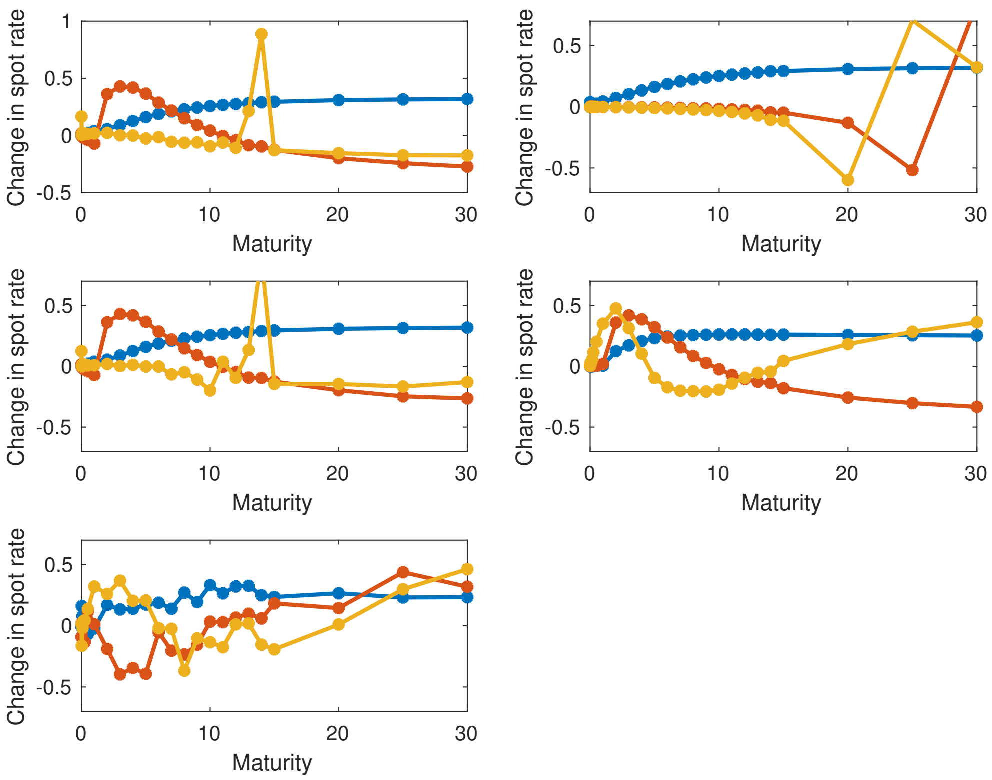

Figure 3.6: Impact on the US term structure of earlier three Principal Components (PC), varying the estimation procedure of the covariance matrix. Curves represent the sensitivity of changes (returns) in interest rate against increasing maturity. The sensitivity to the first PC is blue, the one to the second PC is orange and the last to the third PC is yellow. From left to right: PC for the Sample covariance matrix (SCV - top-left), Shrinkage (SH - top-right), Nonlinear Shrinkage (NSH - middle-left), MCD (middle-right) and for MRCD (bottom-left).

Appendix 3.A MRCD regularisation parameters with simulated data

Table A3.1: MRCD regularisation parameters in the case of simulated data.

	25OPI	**25SBM**	**25SM**	**30I**	**300SPX**
Panel A: MRCD with $h = 0.5$					
$\xi = 0$	0.0000	0.0043	0.0067	0.0092	0.0128
$\xi = 2.5\%$	0.0000	0.0055	0.0088	0.014	0.0198
$\xi = 5\%$	0.0000	0.0066	0.0108	0.0169	0.0238
$\xi = 10\%$	0.0000	0.0085	0.014	0.0206	0.0288
Panel B: MRCD with $h = 0.75$					
$\xi = 0$	0.0000	0.0000	0.0048	0.0062	0.0088
$\xi = 2.5\%$	0.0000	0.0000	0.0064	0.0086	0.0132
$\xi = 5\%$	0.0000	0.0000	0.0079	0.0105	0.016
$\xi = 10\%$	0.0000	0.0000	0.0101	0.0134	0.0195
Panel C: MRCD with $h = 0.9$					
$\xi = 0$	0.0000	0.0000	0.0043	0.0053	0.0075
$\xi = 2.5\%$	0.0000	0.0000	0.006	0.0076	0.0111
$\xi = 5\%$	0.0000	0.0000	0.0074	0.0092	0.0134
$\xi = 10\%$	0.0000	0.0000	0.0093	0.0116	0.0165

Appendix 3.B Dataset 5 composition

On following the complete listings of the assets composing the investment universe 5, taken from Bloomberg.

AAPL, ABMD, ABT, ADBE, ADM, ADSK, AEP, AFL, AIG, AIV, AJG, ALB, ALK, ALL, AMAT, AMGN, ANDV, AON, AOS, APA, APC, APD, ATVI, AVB, AVY, AXP, AZO, BA, BAC, BAX, BBT, BBY, BDX, BEN, BF/B, BHGE, BIIB, BK, BLL, BMY, BSX, BWA, C, CAG, CAH, CAT, CB, CCL, CELG, CERN, CHD, CI, CINF, CL, CLX, CMA, CMCSA, CMI, CMS, CNP, COF, COG, COO, COP, COST, CPB, CPRT, CSCO, CTAS, CTL, CTXS, CVS, CVX, D, DE, DHI, DHR, DIS, DISH, DLTR, DOV, DRE, DRI, DTE, DUK, DVA, DWDP, EA, ECL, ED, EFX, EIX, EL, EMN, EMR, EOG, EQR, EQT, ES, ESRX, ESS, ETN, ETR, EVRG, EXC, EXPD, F, FAST, FCX, FDX, FISV, FITB, FL, FLIR, FMC, FRT, GD, GE, GGP, GILD, GIS, GLW, GPC, GPS, GWW, HAL, HBAN, HCP, HD, HES, HIG, HOG, HOLX, HON, HP, HPQ, HRB, HRL, HRS, HSIC, HST, HSY, HUM, IBM, IDXX, IFF, INCY, INTC, INTU, IP, IPG, IR, ITW, IVZ, JBHT, JCI, JEC, JEF, JNJ, JPM, K, KEY, KIM, KLAC, KMB, KO, KR, KSS, KSU, L, LB, LEG, LEN, LH, LLY, LMT, LNC, LNT, LOW, LRCX, LUV, M, MAA, MAC, MAS, MCD, MCHP, MCK, MDT, MGM, MLM, MMC, MMM, MNST, MO, MRK, MRO, MS, MSFT, MSI, NBL, NEE, NEM, NFX, NI, NKE, NKTR, NOC, NSC, NTAP, NTRS, NUE, NWL, O, OKE, OMC, ORLY, OXY, PAYX, PBCT, PCAR, PCG, PEG, PFE, PG, PGR, PH, PHM, PKI, PNC, PNW, PPG, PPL, PSA, PSX, PVH, PX, QCOM, RCL, RE, REG, REGN, RHI, RJF, ROST, RTN, SBUX, SCG, SCHW, SHW, SIVB, SJM, SLB, SNA, SNPS, SO, SPG, SPGI, STI, STT, SWK, SWKS, SYMC, SYY, T, TGT, TIF, TJX, TMK, TMO, TROW, TRV, TSCO, TSS, TXT, UDR, UHS, UNH, UNM, UNP, USB, UTX, VAR, VFC, VMC, VNO, VRTX, VZ, WAT, WEC, WELL, WFC, WHR, WM, WMB, WY, XL, XLNX.

Appendix 3.C MRCD regularisation parameters with empirical data

Table C3.1: MRCD regularisation parameters in the case of empirical data.

	S1	**S2**	**S3**	**S4**	**S5**
Panel A: $m = 360$					
MRCD $h = 0.5$	0.0000	0.0000	0.0000	0.0000	-
MRCD $h = 0.75$	0.0000	0.0000	0.0000	0.0000	-
MRCD $h = 0.9$	0.0000	0.0000	0.0000	0.0000	-
Panel B: $m = 180$					
MRCD $h = 0.5$	0.0000	0.0000	0.0004	0.0001	0.0409
MRCD $h = 0.75$	0.0000	0.0000	0.0001	0.0001	0.0000
MRCD $h = 0.9$	0.0000	0.0000	0.0002	0.0002	0.0678
Panel C: $m = 120$					
MRCD $h = 0.5$	0.0016	0.0017	0.0024	0.0023	0.0339
MRCD $h = 0.75$	0.0000	0.0000	0.0013	0.0012	0.0000
MRCD $h = 0.9$	0.0000	0.0004	0.0016	0.0017	0.0671
Panel D: $m = 60$					
MRCD $h = 0.5$	0.0095	0.0092	0.0092	0.0079	0.0369
MRCD $h = 0.75$	0.0033	0.0036	0.0048	0.0045	0.0000
MRCD $h = 0.9$	0.002	0.0024	0.0046	0.0045	0.0618
Panel E: $m = 36$					
MRCD $h = 0.5$	0.0095	0.0105	0.0105	0.0094	0.0419
MRCD $h = 0.75$	0.0126	0.0123	0.0125	0.0119	0.0000
MRCD $h = 0.9$	0.0159	0.0125	0.0137	0.0131	0.0619

Chapter 4

Shrinkage estimator in Risk–Based Portfolios [1]

4.1 Introduction

The seminal contributions of [Markowitz, 1952, 1956] laid the foundations for his well–known portfolio building technique. Albeit elegant in its formulation and easy to be implemented in real-world applications, the Markowitz model relies on securities returns sample mean and sample covariance as inputs to estimate the optimal allocation. However, there is a large consensus on the fact that sample estimators perpetuate large estimation errors; this directly affects portfolio weights that often exhibit extreme values, fluctuating over time with very poor performance out-of-sample [DeMiguel and Uppal, 2009].

This problem has been tackled from different perspectives: [Jorion, 1986] and [Michaud, 1989] suggested Bayesian alternatives to the sample estimators; [Jagannathan and Ma, 2003] added constraints to the Markowitz model limiting the estimation error; [Black and Litterman, 1992] derived an alternative portfolio construction technique exclusively based on the covariance matrix among asset returns, avoiding estimating the mean value for each security and converging to the Markowitz Minimum Variance portfolio with no short-sales. This latter technique is supported by results in [Merton, 1980] and [Chopra and Ziemba, 1993], who clearly demonstrated how the mean estimation process can lead to more severe distortions than those in the case of the covariance matrix.

Following this perspective, estimation error can be reduced by considering risk-based portfolios: findings suggest they have good out-of-sample performance without much turnover [DeMiguel and Uppal, 2009]. There is a recent research strand focused on deriving risk-based portfolios other than the Minimum Variance one. In this context, [Qian, 2006] designed a way to select assets by assigning to each of them the same contribution to the overall portfolio risk; [Choueifaty and Coignard, 2008] proposed a portfolio where diversification is the key criterion in asset selection; [Maillard et al., 2010] offered a novel portfolio construction technique where weights perpetuate an equal risk contribution while maximising diversification.

[1] The research paper *Target matrix estimators in Risk-based portfolios* is based on the results in this chapter and it has been published in *Risks*, 6(4), 1–20, Special Issue Computational Methods for Risk Management in Economics and Finance.

These portfolios are largely popular among practitioners[1]: they highlight the importance of diversification, risk budgeting; moreover, they put risk management in a central role, offering a low computational burden to estimate weights. They are perceived as "robust" models since they do not require the explicit estimation of the mean. Unfortunately, limiting the estimation error in this way poses additional problems related to the ill-conditioning of the covariance matrix that occurs when the number of securities becomes sensitively greater than the number of observations. In this case, the sample eigenvalues become more dispersed than the population ones [Marčenko and Pastur, 1967], and the sample covariance matrix directly affects weight estimation. This means that for a high-dimensional dataset, the sample covariance matrix is not a reliable estimator.

To reduce misspecification effects on portfolio weights, more sophisticated estimators than the sample covariance have been proposed, for example the Bayes-Stein shrinkage technique [James and Stein, 1961], henceforth shrinkage stems for its practical implementation and related portfolio performance. This technique reduces the misspecification in the sample covariance matrix by shrinking it towards an alternative estimator. Here, the problem is to select a convenient target estimator as well as to find the optimal intensity at which to shrink towards the sample covariance matrix. The latter is usually derived by minimising a predefined loss function to obtain the minimum distance between the *true* and the shrunk covariance matrices [Ledoit and Wolf, 2003]. A comprehensive overview on shrinkage intensity parameters can be found in [DeMiguel and Nogales, 2013], where the authors proposed an alternative way of deriving the optimal intensity based on the smoothed bootstrap approach. On the other hand, the target matrix is often selected among the class of structured covariance estimators [Briner and Connor, 2008], especially because the matrix which shrinks is the sample one. As noted in [Candelon et al., 2012], the sample covariance matrix is the Maximum Likelihood Estimator (MLE) under the Normality of asset returns, hence it lets the data speak without imposing any structure. This naturally suggests it might be pulled towards a more structured alternative. Dealing with financial data, the shrinkage literature proposes six different models for the target matrix: the Single-Index market model [Ledoit and Wolf, 2003]; [Briner and Connor, 2008]; [Candelon et al., 2012]; [Ardia et al., 2017]); the Identity matrix [Ledoit and Wolf, 2004a]; [Candelon et al., 2012]; the Variance Identity matrix [Ledoit and Wolf, 2004a]; the Scaled Identity matrix [DeMiguel and Nogales, 2013]; the Constant Correlation model [Ledoit and Wolf, 2004b] and [Pantaleo et al., 2011]; the Common Covariance [Pantaleo et al., 2011]. All these targets belong to the class of more structured covariance estimators than the sample one, thus implying the latter is the matrix to shrink.

Despite the great improvements in portfolio weight estimation under the Markowitz portfolio building framework, the shrinkage technique has only been applied in one work involving

[1]The majority of papers on risk-based portfolios are published in journal aimed at practitioners, as the Journal of Portfolio Management.

risk-based portfolios: that of [Ardia et al., 2017], who comprehensively described the impacts of variance and covariance misspecifications in risk-portfolio weights. [Ardia et al., 2017] tested four alternative covariance estimators to reduce weight misspecification; among those, only one refers to shrinkage as in [Ledoit and Wolf, 2003], leaving room open for further research. In our work, we contribute to the existing literature, filling this gap and offering a comprehensive overview of shrinkage in risk-based portfolios. In particular, we study the effect of six target matrix estimators on the weights of four risk-based portfolios. To achieve this goal, we provide an extensive Monte Carlo simulation aimed at (1) assessing estimators' statistical properties and similarity with the *true* target matrix; (2) addressing the problem of how the selection of a specific target estimator affects the portfolio weights. We find out that the Identity and Variance Identity hold the best statistical properties, being well conditioned even in a high-dimensional dataset. These two estimators also represent the more efficient target matrices towards which to shrink the sample one. In fact, portfolio weight derived shrinking towards the Identity and Variance Identity minimise the distance from their *true* counterparts, especially in the case of Minimum Variance and Maximum Diversification portfolios.

The rest of the chapter is organised as follows. Section 4.2 introduces the risk-based portfolios employed in the study. Section 4.3 illustrates the shrinkage estimator, the moves to the six target matrix estimators and provides useful insights into misspecification when shrinkage is applied to risk-based portfolios. In Section 4.4, we run an extensive Monte Carlo analysis for describing how changes in the target matrix affect risk-based portfolio weights. Section 4.5 concludes.

4.2 Risk–Based Portfolios

Risk-based portfolios are particularly appealing since they rely only on the estimation of a proper measure of risk, i.e., the covariance matrix between asset returns. Assume an investment universe made by p assets:

$$X = \left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p\right) \quad (1)$$

is a $n \times p$ containing a history of n log-returns for the i-th asset, where $i = 1, \ldots, p$. The covariance matrix among asset log-returns is the symmetric square matrix Σ[2] of dimension $p \times p$, and the unknown optimal weights form the vector $\boldsymbol{\omega}$ of dimension $p \times 1$. Our working framework assumes to consider four risk-based portfolios: the Minimum Variance (MV), the

[2]With this we refer to the population covariance matrix, which by definition is not observable and then unfeasible. Hence, Σ is estimated taking into account the observations stored in X: we will deeply treat this in the next section.

Inverse Volatility (IV), the Equal-Risk-Contribution (ERC), and the Maximum Diversification (MD) upon two constraints; no short-selling ($\boldsymbol{\omega} \in \Re^p_+$) and full allocation of the available wealth ($\boldsymbol{\omega}'\mathbf{1}_p = 1$, where $\mathbf{1}_p$ is the vector of ones of length p).

The Minimum Variance portfolio [Markowitz, 1952] derives the optimal portfolio weights by solving this minimisation problem with respect to $\boldsymbol{\omega}$:

$$\boldsymbol{\omega}_{MV} \equiv \underset{\boldsymbol{\omega}}{\text{argmin}} \left\{ \boldsymbol{\omega}'\Sigma\boldsymbol{\omega} \;\middle|\; \boldsymbol{\omega} \in \Re^p_+, \boldsymbol{\omega}'\mathbf{1}_p = 1 \right\} \tag{2}$$

where $\boldsymbol{\omega}'\Sigma\boldsymbol{\omega}$ is the portfolio variance.

In the Inverse Volatility portfolio, also known as the equal-risk-budget portfolio [Leote et al., 2012], a closed form solution is available. Each element of the vector $\boldsymbol{\omega}$ is given by the inverse of the i–th asset variance (denoted by $\Sigma^{-1}_{i,i}$) divided by the inverse of the sum of all asset variances:

$$\boldsymbol{\omega}_{IV} \equiv \left(\frac{\sum^{-1}_{1,1}}{\sum_{i=1}^{p} \sum^{-1}_{i,i}}, \ldots, \frac{\sum^{-1}_{p,p}}{\sum_{i=1}^{p} \sum^{-1}_{i,i}} \right)' \tag{3}$$

In the Equal-Risk-Contribution portfolio, as the name suggests, the optimal weights are calculated by assigning to each asset the same contribution to the whole portfolio volatility, thus originating a minimisation procedure to be solved with respect to $\boldsymbol{\omega}$:

$$\boldsymbol{\omega}_{ERC} \equiv \underset{\boldsymbol{\omega}}{\text{argmin}} \left\{ \sum_{i=1}^{p} \left(\%RC_i - \frac{1}{p} \right)^2 | \boldsymbol{\omega} \in \Re^p_+, \boldsymbol{\omega}'\mathbf{1}_p = 1 \right\} \tag{4}$$

Here, $\%RC_i \equiv \frac{\boldsymbol{\omega}_i cov_{i,\pi}}{\sqrt{\boldsymbol{\omega}'\Sigma\boldsymbol{\omega}}}$ is the percentage risk contribution for the i-th asset, $\sqrt{\boldsymbol{\omega}'\Sigma\boldsymbol{\omega}}$ is the portfolio volatility as earlier defined, and $\boldsymbol{\omega}_i cov_{i,\pi}$ provides a measure of the covariance of the i-th exposure to the total portfolio π, weighted by the corresponding $\boldsymbol{\omega}_i$.

Turning to the Maximum Diversification, as in [Choueifaty and Coignard, 2008] we preliminary define $DR(\boldsymbol{\omega})$ as the portfolio's diversification ratio:

$$DR(\boldsymbol{\omega}) \equiv \frac{\boldsymbol{\omega}' \sqrt{diag(\Sigma)}}{\sqrt{\boldsymbol{\omega}'\Sigma\boldsymbol{\omega}}}$$

where $diag(\Sigma)$ is a $p \times 1$ vector which takes all the asset variances $\Sigma_{i,i}$ and $\boldsymbol{\omega}'\sqrt{diag(\Sigma)}$ is the weighted average volatility. By construction, it is $DR(\boldsymbol{\omega}) \geq 1$, since the portfolio volatility is sub-additive [Ardia et al., 2017]. Hence, the optimal allocation is the one with the highest DR:

$$\boldsymbol{\omega}_{MD} \equiv \underset{\boldsymbol{\omega}}{\operatorname{argmax}} \left\{ DR\left(\boldsymbol{\omega}\right) \middle| \, \boldsymbol{\omega} \in \Re^{p}_{+}, \boldsymbol{\omega}'\mathbf{1}_{p} = 1 \right\}. \tag{5}$$

4.3 Shrinkage Estimator

The shrinkage technique relies upon three ingredients: the starting covariance matrix to shrink, the target matrix towards which the former is shrunk, and the shrinkage intensity, or roughly speaking the strength at which the starting matrix must be shrunk.

In financial applications, the starting matrix which is to shrink is always the sample covariance matrix. This is a very convenient choice that helps in the selection of a proper shrinkage target: being the sample covariance a model-free estimator that completely reflects the relationships among data[3], it becomes natural to select a target in the class of more structured covariance estimators [Briner and Connor, 2008]. In addition, this strategy allows direct control over the trade-off between estimation error and model error in the resulting shrinkage estimates. In fact, the sample covariance matrix is usually affected by a large amount of estimation error. This is reduced when shrinking towards a structured target which minimises the sampling error at the cost of adding some misspecification by imposing a specific model. At this point, the shrinkage intensity is crucial because it must be set in such a way to minimise both errors.

To define the shrinkage estimator, we start from the definition of sample covariance matrix S. Recalling Equation (1), S is given by:

$$S = \frac{1}{n-1} X'(I_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n')X, \tag{6}$$

where I_n denotes the $n \times n$ identity matrix and $\mathbf{1}_n$ is the ones column vector of length n. The shrinkage methodology enhances the sample covariance matrix estimation by shrinking S towards a specific target matrix T:

$$\Sigma_s = \delta T + (1-\delta)\, S \tag{7}$$

where Σ_s is the shrinkage estimator; δ the shrinkage parameter and T the target matrix. In this work, we focus on the problem of selecting the target matrix. After a review of the literature on target matrices, in the following rows we present the target estimators considered in this study and we assess through a numerical illustration the impact of misspecification in the target matrix for the considered risk-based portfolios.

[3] The sample covariance matrix is the Maximum Likelihood Estimator (MLE) under Normality, therefore it lets data speaks without imposing any structure.

4.3.1 Literature Review

The target matrix should fit a desirable number of requirements: First, it should be structured much enough to lower the estimation error of the sample covariance matrix while not bringing too much error from model selection. Second, it should reflect the important features of the *true* covariance matrix [Ledoit and Wolf, 2004b]. The crucial question is: how much structure should we impose to fill in the requirements? Table 4.1 shows the target matrices employed so far in the literature, summarising information about the formula for the shrinkage intensity, the wealth allocation rule, and the addressed research question. Not surprisingly, all the papers shrink the sample covariance matrix. What surprises is that only six target matrices have been examined: the one relying on the Single-Index market model, the Identity matrix, the Scaled Identity, and the Variance Identity, the Constant Correlation model and the Common Covariance model. Previously, four were proposed by Ledoit and Wolf in separate works [Ledoit and Wolf, 2003, 2004a,b] and were again proposed in subsequent works, while the Common Covariance appears only in [Pantaleo et al., 2011] and the Scaled Identity only in [DeMiguel and Nogales, 2013].

[Table 4.1 about here.]

In Table 4.1, papers have been listed taking into account their contribution to the literature as regards the adoption of a novel target matrix estimator, the re-examination of a previously proposed target, and the comparison among different estimators. Ledoit and Wolf popularised the shrinkage methodology in portfolio selection: in [Ledoit and Wolf, 2003], they were also the first to compare the effects of shrinking towards different targets in portfolio performance. Shrinking towards the Variance Identity and shrinking towards the Market Model are two out of the eight estimators for the covariance matrix compared with respect to the reduction of estimation error in portfolio weights. They found significant improvements in portfolio performance when shrinking towards the Market Model. [Briner and Connor, 2008] well described the importance of selecting the target matrix among the class of structured covariance estimators, hence proposing to shrink the asset covariance matrix of demeaned returns towards the Market model as in [Ledoit and Wolf, 2003]. [Candelon et al., 2012] compared the effect of double shrinking the sample covariance either towards the Market Model and the Identity, finding that both estimators carry on similar out-of-sample performances. [DeMiguel and Nogales, 2013] compared the effects of different shrinkage estimators on portfolio performance, highlighting the importance of the shrinkage intensity parameter and proposing a scaled version of the Identity Matrix as a target. Another important comparison among target matrices is that of [Pantaleo et al., 2011], who compared the Market and Constant Correlation models as in [Ledoit and Wolf, 2003, 2004b] with the Common Covariance of [Schäfer and Strimmer, 2005] implemented as target matrix for the first time in finance. The authors assessed the effects on portfolio performances while controlling for the dimensionality of the dataset, finding that the Common Covariance should not be used when the number of observations is less than the number of assets. Lastly, [Ardia et al., 2017] is the only work to implement shrinkage in risk-based portfolios. They shrunk the sample covariance matrix as in [Ledoit and Wolf, 2003], finding that the Minimum Variance and the Maximum Diversification portfolios are the most affected from covariance misspecification, hence they benefit the most from the shrinkage technique.

4.3.2 Estimators for the Target Matrix

We consider six estimators for the target matrix: the Identity and the Variance Identity matrix, the Single-index, the Common Covariance, the Constant Correlation and the Exponential Weighted Moving Average (EWMA) models. They are all structured estimators, in the sense that the number of parameters to be estimated is far less the $\frac{1}{2}p\,(p+1)$ required in the sample covariance case. Compared with the literature, we take into account all the previous target estimators,[4] adding to the analysis the EWMA: this estimator well addresses the problem of serial correlation and heteroskedasticity in asset returns.

The identity is a matrix with ones on the diagonal and zeros elsewhere. Choosing the Identity as the target is justified by the fact that is shows good statistical properties: it is always well conditioned and hence invertible [Ledoit and Wolf, 2003]. Besides the identity, we also consider a multiple of the identity, named the Identity Variance. This is given by:

$$T_{VId} \equiv I_p diag\,(S)\, I_p, \tag{8}$$

where $diag\,(S)$ is the main diagonal of the sample covariance matrix (hence the assets variances) and I_p the identity matrix of dimension p.

The Single Index Model [Sharpe, 1963] assumes that the returns $\boldsymbol{r}_t$ can be described by a one-factor model, resembling the impact of the whole market:

$$\boldsymbol{r}_t = \boldsymbol{\alpha} + \boldsymbol{\beta} r_{mkt} + \varepsilon_t, \text{ with } t = 1, \ldots, n,$$

where r_{mkt} is the overall market returns; $\boldsymbol{\beta}$ is the vector of factor estimates for each asset; $\boldsymbol{\alpha}$ is the market mispricing, and ε_t the model error. The Single-Index market model represents a practical way of reducing the dimension of the problem, measuring how much each asset is affected by the market factor. The model implies the covariance structure among asset returns is given by:

[4] In reality, we exclude the Scaled Identity of [DeMiguel and Nogales, 2013] because of its great similarity with the Identity and Variance Identity implemented in our study.

$$T_{si} \equiv s^2_{mkt}\boldsymbol{\beta}\boldsymbol{\beta}' + \Omega, \tag{9}$$

where s^2_{mkt} is the sample variance of asset returns; $\boldsymbol{\beta}$ is the vector of beta estimates and Ω contains the residual variance estimates.

The Common Covariance model is aimed at minimising the heterogeneity of assets variances and covariances by averaging both of them [Pantaleo et al., 2011]. Let $\text{var}_{ij,i=j}$ and $\text{covar}_{ij,\ i\neq j}$ being the variances and covariances of the sample covariance matrix, respectively, their averages are given by:

$$\overline{\text{var}} = \frac{1}{p}\sum_{k=1}^{p} \text{var}_{k,i=j};$$

$$\overline{\text{covar}} = \frac{1}{p\,(p-1)\,/2}\sum_{k=1}^{p(p-1)/2} \text{covar}_{k,\ i\neq j};$$

where p is the number of securities. The resulting target matrix T_{cv} has its diagonal elements all equal to the average of the sample variance, while non-diagonal elements are all equal to the average of sample covariances.

In the Constant Correlation model the main diagonal is filled with sample variances, and elsewhere a constant covariance parameter which is equal for all assets. The matrix can be written according to the following decomposition:

$$T_{cc} \equiv P\ diag\,(S)\,P, \tag{10}$$

where P is the lower triangular matrix filled with the constant correlation parameter $\overline{\rho} = \frac{1}{p(p-1)/2}\sum_{i=1}^{p}\rho_{ij}$ for $i < j$ and ones in the main diagonal. Here, $diag\,(S)$ represents the main diagonal of the sample covariance matrix.

The EWMA model [RiskMetrics, 1996] was introduced by JP Morgan's research team to provide an easy but consistent way to assess portfolio covariance. RiskMetrics EWMA considers the variances and covariance driven by an Integrated GARCH process:

$$T_{\text{EWMA},t} \equiv X'X + \lambda T_{\text{EWMA},t-1}, \tag{11}$$

with $T_{\text{EWMA},0} = \ I_p\ T_{\text{EWMA},t-1}$ is the target matrix at time $t-1$ and λ is the smoothing parameter: the higher λ, the higher the persistence in the variance.

4.3.3 The Impact of Misspecification in the Target Matrix

We are now going to show to which extent risk-based portfolios can be affected by misspecification in the target matrix. To do so, we provide a numerical illustration, merely inspired by the one in [Ardia et al., 2017]. Assume an investment universe made by three securities: a sovereign bond (Asset-1), a corporate bond (Asset-2), and equity (Asset-3), we are able to impose an arbitrary structure to the related 3×3 *true* covariance matrix[5]. We preliminary recall that Σ can be written according to the following decomposition:

$$\Sigma \equiv \left(diag\left(\Sigma\right)\right)^{1/2} P_{\Sigma} \left(\mathrm{diag}\left(\Sigma\right)\right)^{1/2}$$

where $\left(diag\left(\Sigma\right)\right)^{1/2}$ is a diagonal matrix with volatilities on the diagonal and zeros elsewhere and P_{Σ} is the related correlation matrix, with ones on the diagonal and correlations symmetrically displaced elsewhere. We impose

$$(\Sigma_{1,1}^{1/2},\ \Sigma_{2,2}^{1/2},\ \Sigma_{3,3}^{1/2}) = (0.1,\ 0.1,\ 0.2)$$

and

$$(\mathrm{P}_{\Sigma;1,2},\ \mathrm{P}_{\Sigma;1,3},\ \mathrm{P}_{\Sigma;2,3}) = \ (0.1,\ 0.2,\ 0.7)$$

hence, the *true* covariance matrix is:

$$\Sigma \equiv \begin{bmatrix} 0.010 & 0.001 & 0.004 \\ 0.001 & 0.010 & 0.014 \\ 0.004 & 0.014 & 0.040 \end{bmatrix}$$

Now assume that the *true* covariance matrix Σ is equal to its shrunk counterpart when $\delta = \dfrac{1}{2}$:

$$\Sigma \equiv \Sigma_s = \frac{1}{2}S + \frac{1}{2}T$$

[5][Ardia et al., 2017] imposes Asset-1 and Asset-2 to have 10% annual volatility; Asset-3 to have 20% annual volatility; correlations between Asset-1/Asset-2 and Asset-1/Asset-3 are set as negative and correlation between corporate bonds and equities (Asset-2/Asset-3) is set as positive. However, to better resemble real data, specifically the S&P500, the US corporate index and the US Treasury Index total returns, we assume all three correlation parameters to be positive.

that is both the sample covariance matrix S and the target matrix T must be equal to $\frac{1}{2}\Sigma$ and the *true* target matrix is:

$$\mathrm{S} \equiv \mathrm{T} \equiv \begin{bmatrix} 0.005 & 0.0005 & 0.002 \\ 0.0005 & 0.005 & 0.007 \\ 0.002 & 0.007 & 0.020 \end{bmatrix}$$

with few algebraic computations, we can obtain the volatilities and correlations simply by applying the covariance decomposition, ending up with

$$\left(T_{1,1}^{1/2},\ T_{2,2}^{1/2},\ T_{3,3}^{1/2}\right) = (0.0707,\ 0.0707,\ 0.1414)\,;$$

$$\left(P_{T;1,2}\ ,\ P_{T;1,3},\ P_{T;2,3}\right) = (0.1,\ 0.2,\ 0.7)\,.$$

In this case, we can conclude that the target matrix T is undervaluing all the covariance and correlation values.

At this point, some remarks are needed. First, as summarised in Table 4.2, we work out the *true* risk-based portfolio weights. Weights are differently spread out: the Minimum Variance equally allocates wealth to the first two assets, excluding equities. This because it mainly relies upon the asset variance, limiting the diversification of the resulting portfolio. The remaining portfolios allocate wealth without excluding any asset; however, the Maximum Diversification overvalues Asset-1 assigning to it more than 50% of the total wealth. The Inverse Volatility and Equal-Risk-Contribution seem to maximise diversification under a risk-parity concept, similarly allocating wealth among the investment universe.

[Table 4.2 about here.]

Second, assuming Σ as the *true* covariance matrix allows us to simulate misspecification both in the volatility and in the correlation components of the target matrix T by simply increasing or decreasing the imposed *true* values. Since we are interested in investigating misspecification impact on the *true* risk-based portfolio weights, we measure its effects after each shift with the Frobenius norm between the *true* weights and the misspecified ones:

$$\|\widetilde{\boldsymbol{\omega}}\|_F^2 = \sum_{i=1}^{p} \widetilde{\omega}_i^2$$

where $\widetilde{\boldsymbol{\omega}} = \boldsymbol{\omega} - \hat{\boldsymbol{\omega}}$.

Third, turning the discussion onto the working aspects of this toy example, we will separately shift the volatility and the correlation of Asset-3, as done in [Ardia et al., 2017]. The difference with them is that we modify the values in the *true* target matrix T. Moreover, in order to also gauge how shrinkage intensity affects the portfolio weights, we perform this analysis for 11 values of δ, spanning from 0 to 1 (with step 0.1). This allows us to understand both extreme cases, i.e., when the *true* covariance matrix is only estimated with the sample estimator ($\delta = 0$) and only with the target matrix ($\delta = 1$). Remember that the *true* shrinkage intensity is set at $\delta = \dfrac{1}{2}$.

Moving to the core of this numerical illustration, we proceed as follows. First, for what is concerning the volatility, we let $T_{3,3}^{1/2}$ vary between 0 and 0.5, ceteris paribus. Results are summarised in Figure 4.1, row 1. As expected, there is no misspecification in all the risk-based portfolio at the initial state $T_{3,3}^{1/2} = 0.1414$, i.e., the *true* value. All the portfolio weights are misspecified in the range $[0; 0.1414)$, with the Minimum Variance portfolio showing the greatest departure from the *true* portfolio weights when the Asset-3 volatility is undervalued below 0.12. The absence of misspecification effects in its weights is due to the initial high-risk attributed to Asset-3; in fact, it is already excluded from the optimal allocation at the initial non–perturbed state. Regarding the other portfolios, their weights show a similar behaviour to the one just described: the Maximum Diversification weights depart from the non–misspecified state to reach the maximum distance from the *true* weights of 0.4; however, this effect dissipates as soon as the shrinkage intensity grows. The same applies for the Inverse Volatility and the Equal-Risk-Contribution. On the contrary, when volatility is overvalued in the range $(0.1414; 0.5]$, the Minimum Variance is not misspecified, since Asset-3 is always excluded from the allocation. This fact helps to maintain the stability of its weights: this portfolio is not affected by shifts in the shrinkage intensity when there is over–misspecification. All the remaining portfolios show low levels of misspecification due to diversification purposes. In particular, they react in the same way to shrinkage intensity misspecification, showing an increase in the Frobenius norm especially for low values of Asset-3 variance. A common trait shared by all the considered portfolios is that when weights are estimated with the sample covariance, only the distance from *true* portfolios is at the maximum.

[Figure 4.1 about here.]

Second, we assess the correlation misspecification impact. We let the correlation between Asset-3 and Asset-2 ($\mathrm{P}_{\mathrm{T};2,3}$) vary from 0 to 1, ceteris paribus. In this case, the greatest signs of perturbation are in the Minimum Variance and in the Maximum Diversification portfolios, while the Equal-Risk-Contribution shows far less distortions, as presented in Figure 4.1, row 2. The Minimum Variance portfolio is again misspecified in one direction: when the correlation parameter is undervalued and the sample covariance matrix dominates the target matrix in the shrinkage. On the other hand, the Maximum Diversification shows the highest departure from no-misspecification levels in both senses. However, as the Maximum Diversification it seems to benefit from high values of shrinkage intensity. The Equal-Risk-Contribution reacts similarly to the Maximum Diversification, but with a far lower level of misspecification, the Inverse Volatility is not affected at all by misspecification in the correlation structure of the target matrix T. This is due to the specific characteristics of Asset-3 and the way in which the Inverse Volatility selects to allocate weights under a risk-parity scheme.

In conclusion, with this numerical illustration, we assess the effects of target matrix misspecification in risk-based portfolios: the four risk-based portfolios react similarly to what previously found in [Ardia et al., 2017], even if in our case, shifts originated in the target matrix. The Minimum Variance and the Maximum Diversification portfolios are the most impacted: the weights of the former are severely affected by volatility and covariance shifts undervaluing the *true* values; the latter shows perturbations in weights when shifts are more extreme. Both portfolios benefit from a higher level of shrinkage intensity. On the other hand, the Inverse Volatility and the Equal-Risk-Contribution weights suffer less from both sources of misspecification. Overall, weights are affected by shifts in the shrinkage intensity: when sample covariance is the estimator ($\delta = 0$), the distance from the *true* weights stands at the maximum level in all the considered portfolios.

4.4 Case Study – Monte Carlo Simulation

This section offers a comprehensive comparison of the six target matrix estimators by means of an extensive Monte Carlo study. The aim of this analysis is twofold: (1) assessing estimators' statistical properties and similarity with the *true* target matrix; (2) addressing the problem of how selecting a specific target estimator impacts on the portfolio weights. This investigation is aimed at giving a very broad overview about (1) and (2) since we monitor both the p/n ratio and the whole spectrum of shrinkage intensity. We run simulations for 15 combinations of p and n, and for 11 different shrinkage intensities spanning in the interval $[0\,;1]$, for an overall number of 165 scenarios.

The Monte Carlo study is designed as follows. Returns are simulated assuming a factor model is the data generating process, as in [MacKinlay and Pastor, 2000]. In detail, we impose a one-factor structure for the returns generating process:

$$r_t = \xi f_t + \varepsilon_t;$$

$$\text{with } t = 1, \dots, n,$$

where f_t is the $k \times 1$ vector of returns on the factor, ξ is the $p \times k$ vector of factor loadings, and ε_t the vector of residuals of p length. Under this framework, returns are simulated implying multivariate normality and absence of serial correlation. The asset factor loadings are drawn from a uniform distribution and equally spread, while returns on the single factor are generated from a Normal distribution. The bounds for the uniform distribution and the mean and the variance for the Normal one are calibrated on real market data, specifically on the empirical dataset "49-Industry portfolios" with monthly frequency, available on the Kennet French website[6]. Residuals are drawn from a uniform distribution in the range $[0.10; 0.30]$ so that the related covariance matrix is diagonal with an average annual volatility of 20%.

For each of the 165 scenarios, we apply the same strategy. First, we simulate the $n \times p$ matrix of asset log-returns, then we estimate the six target matrices and their corresponding shrunk matrices $\hat{\Sigma}_s$. Finally, we estimate the weights of the four risk-based portfolios. Some remarks are needed. First, we consider the number of assets as $p = \{10,\ 50,\ 100\}$ and number of observations as $n = \{60,\ 120,\ 180,\ 3000,\ 6000\}$ months, which correspond to 5, 10, 15, 250 and 500 years. Moreover, the shrinkage intensity is allowed to vary between their lower and upper bounds as $\delta = \{0,\ 0.1,\ 0.2,\ 0.3,\ 0.4,\ 0.5,\ 0.6,\ 0.7,\ 0.8,\ 0.9,\ 1\}$. For each of the 165 scenarios we run 100 Monte Carlo trials[7], giving robustness to the results.

We stress again the importance of Monte Carlo simulations, which allow us to impose the *true* covariance Σ and hence the *true* portfolio weights $\boldsymbol{\omega}$. This is crucial because we can compare the *true* quantities with their estimated counterparts.

With respect to the point (1), we use two criteria to assess and compare the statistical properties of target matrices: the reciprocal 1-norm condition number and the Frobenius Norm. Being the 1-norm condition number defined as:

$$CN\,(A) = \ \kappa\,(A) = \|A^{-1}\|$$

for a given A. It measures the matrix sensitivity to changes in the data: when it is large, it indicates that a small shift causes important changes, offering a measure of the ill-conditioning of A. Since $CN\,(A)$ takes value in the interval $[0\ ; +\infty)$, it is more convenient to use its scaled version, the $RCN\,(A)$:

$$RCN\,(A) = 1/\ \kappa\,(A) \tag{12}$$

It is defined in the range $[0\ ; 1]$: the matrix is well-conditioned if the reciprocal condition number is close to 1 and ill-conditioned vice–versa. Under the Monte Carlo framework, we will study its Monte Carlo estimator:

[6] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

[7] Simulations were done in MATLAB setting the random seed generator at its default value, thus ensuring the full reproducibility of the analysis. Related code available at the GitHub page of the author: https://github.com/marconeffelli/Risk-Based-Portfolios.

$$E\left[CN\left(A_m\right)\right] = \frac{1}{M}\sum_{m=1}^{M} CN\left(A_m\right) \tag{13}$$

where M is the number of Monte Carlo simulations. On the other hand, the Frobenius norm is employed to gauge the similarity between the estimated target matrix and the *true* one. We define it for the $p \times p$ symmetric matrix Z as:

$$FN\left(Z\right) = \|Z\|_F^2 = \sum_{i=1}^{p}\sum_{j=1}^{p} z_{ij}^2$$

In our case, $Z = \hat{\Sigma}_s - \Sigma$. Its Monte Carlo estimator is given by the following:

$$E\left[FN\left(A_m\right)\right] = \frac{1}{M}\sum_{m=1}^{M} FN\left(A_m\right) \tag{14}$$

Regarding (2), we assess the discrepancy between *true* and estimated weights again with the Frobenius norm. In addition, we report the values at which the Frobenius norm attains its best results, i.e., when the shrinkage intensity is optimal.

4.4.1 Main Results

Figure 4.2 summarises the statistical properties of the various target matrices.

[Figure 4.2 about here.]

Figure 4.2 illustrates the reciprocal 1–norm condition number: the matrix is well–conditioned when the value is closer to 1, vice–versa is ill–conditioned the more it tends to zero. Overall, the Identity and the Variance Identity stem for being always well-conditioned: across all the combinations of p and n their reciprocal condition number is always one. Therefore, we focus our analysis on the remaining target matrices. In the case where $p = 10$, the Common Covariance dominates the other three alternatives, who perform poorly. As the number of assets increases, the reciprocal condition number deteriorates, especially for the EWMA, which now performs worse than the others, and for the Common Covariance, which is now aligned to the Single-Index and the Constant Correlation model. In conclusion, excluding the Identity and the Variance Identity, the considered targets show poor statistical properties.

Then, we turn to the study of similarity among *true* and estimated target matrices. Figure 4.3 represents the Monte Carlo Frobenius norm between the *true* and the estimated target matrices. The surfaces give a clear overview about the relation among the Frobenius norm itself, the p/n ratio and the shrinkage intensity. Overall, the Frobenius norm is minimised by the Single-Index and the Common Covariance: in these cases, the target matrices are not particularly affected by the shrinkage intensity, while their reactions to increases in the p/n ratio are controversial. In fact, quite surprisingly the distance between *true* and estimated weights diminishes as both p and n increases. For $p = 50$ and $p = 100$, there is a hump for small p/n values; however, the Frobenius norm increases when $\frac{p}{n} \geq 1$. Despite the low condition number, the EWMA shows a similar behaviour to the Single-Index and the Constant Correlation target matrices, especially with respect to p/n values. On the other hand, it is more affected by shifts in the shrinkage parameters; the distance from the *true* weights increases moving towards the target matrix. Lastly, the Identity and the Variance Identity show a similar behaviour: their distances from the *true* target matrix increase for higher values of δ and p/n. Lastly, the Common Covariance is the most far away from the *true* target matrix, being very sensitive both to high shrinkage intensity and p/n values.

[Figure 4.3 about here.]

To conclude, the Identity and the Variance Identity are the most well-conditioned matrices, being stable across all the examined p/n combinations. Nevertheless, the Single-Index and the Common Covariance target matrices show the greatest similarity with the *true* target matrix minimising Frobenius norm, while both the Identity and the Variance Identity seem less similar to the *true* target.

Results on Portfolio Weights

Tables 4.3 and 4.4 present the main results of the Monte Carlo study: for each combination of p and n, we report the Monte Carlo estimator of the Frobenius norm between the *true* and estimated weights. In particular, Table 4.3 reports averaged Frobenius norm along with the shrinkage intensity (excluding the case $\delta = 0$, which corresponds to the sample covariance matrix), while Table 4.4 lists the minimum values for the optimal shrinkage intensity.

In both tables, we compare the six target matrices by examining one risk-based portfolio at a time and the effect of increasing p for fixed n. Special attention is devoted to the cases when $p > n$: the high-dimensional sample. We have this scenario only when $p = 100$ and $n = 60$. Here, the sample covariance matrix becomes ill-conditioned [Marčenko and Pastur, 1967], thus it is interesting to evaluate gains obtained with shrinkage. The averaged Frobenius norm values in Table 4.3 give us a general overview about how target matrices perform across the whole shrinkage intensity spectrum in one goal. We aim to understand if, in average terms, shrinking the covariance matrix benefits risk-portfolio weights. On the other hand, the minimum Frobenius norm values help us understanding to what extent the various target matrices can help reproducing the *true* portfolio weights: the more intensity we need, the better the target is. In both tables, sample values are listed in the first row of each Panel.

[Table 4.3 about here.]

Starting from Table 4.3, Panel A, the Minimum Variance allocation seems better described by the Identity and the Variance Identity regardless of the number of assets p. In particular, we look at the difference between the weights calculated entirely on the sample covariance matrix and those of the targets: the Identity and the Variance Identity are the only estimator to perform better. In fact, shrinking towards the sample is not as bad as shrinking towards the Common Covariance. By increasing n and moving to Panel B, similar results are obtained. This trend is confirmed in Panel C, while in the cases of $n = 3000$ and $n = 6000$, all the estimators perform similarly. Hence, for the Minimum Variance portfolio the Identity matrix works best at reproducing portfolio weights very similar to the *true* ones. The same conclusions apply for the Maximum Diversification portfolio: when p and n are small, the Identity and the Variance Identity outperform other alternatives. On the other hand, we get very different results for the Inverse Volatility and Equal-Risk-Contribution. Both portfolios seem not gaining benefits from the shrinkage procedure, as the Frobenius norm is very similar to that of the sample covariance matrix for all the target matrices under consideration. This is *true* for all pairs of p and n. In the high-dimensional case ($p = 100$; $n = 60$), the Identity matrix works best in reducing the distance between *true* and estimated portfolio weights, both for the Minimum Variance and Maximum Diversification portfolios. On average, shrinkage does not help too much when alternative target matrices are used; only in the case of Common Covariance is shrinking worse than using the sample covariance matrix. All these effects vanish when we look at the Inverse Volatility and Equal-Risk-Contribution portfolios: here, shrinkage does not help too much, whatever the target is.

[Table 4.4 about here.]

Overall, results are in line with the conclusions of the numerical illustrations in Section 4.3. Indeed, the Minimum Variance portfolio shows the highest distance between *true* and estimated weights, similar to the Maximum Diversification. Both portfolios are affected by the dimensionality of the sample: shrinkage always help in reducing weights misspecification; it improves in high-dimensional cases. On contrary, estimated weights for the remaining portfolios are close to the *true* ones by construction, hence, shrinkage does not help too much.

Switching to Table 4.4, the results illustrate again that the Identity and the Variance Identity attain the best reduction of the Frobenius norm for the Minimum Variance and Maximum Diversification portfolios. If results are similar to those of Table 4.3 for the former, results for the latter show an improvement in using the shrinkage estimators. The Identity, Variance Identity, Common Covariance, and Constant Correlation target matrices outperform all the alternatives, including the sample estimator, minimising the Frobenius norm in a similar fashion. This is true also for the high-dimensional case. On the contrary, the other two portfolios do not benefit from shrinking the sample covariance matrix, even in high-dimensional samples, confirming the insights from Table 4.3. Lastly, we look at the shrinkage intensity at which target matrices attain the highest Frobenius norm reduction. Those values are displayed in Figure 4.4. The intensity is composed of the interval $[0; 1]$: the more it is close to 1, the more the target matrix helps in reducing the estimation error of the sample covariance matrix. Interestingly, the Identity and the Variance Identity show shrinkage intensities always close to 1, meaning that shrinking towards them is highly beneficial, as they are fairly better than the sample covariance matrix. This is verified either for the high-dimensional case and for those risk portfolios (Inverse Volatility and Equal-Risk-Contribution), who do not show great improvements when shrinkage is adopted.

[Figure 4.4 about here.]

Sensitivity to Shrinkage Intensity

To have a view on the whole shrinkage intensity spectrum (i.e., the interval $(0;1)$) we refer to Figure 4.5, where we report the Frobenius Norms for the weights (y-axis) with regard to the shrinkage intensity (x-axis). Each column corresponds to a specific risk-based portfolio: from left to right, the Minimum Variance, the Inverse Volatility, the Equal-Risk-Contribution, and the Maximum Diversification, respectively. Each row corresponds to the p/n ratio in n ascending order. For each subfigure, the Identity is blue circle-shaped, the Variance Identity is green square-shaped, the Single-Index is red hexagram-shaped, the Common Covariance is black star-shaped, the Constant Correlation is cyan plus-shaped, and the EWMA is magenta diamond-shaped.

[Figure 4.5 about here.]

Figure 4.5 illustrates the case $p = 100$, as to include the high-dimensional scenario. Starting from the latter (first row, $n = 60$), the Variance Identity is the only target matrix to always reduce weight misspecification for all the considered portfolios, for all shrinkage levels. The Identity do the same, excluding the ERC case where it performs worse than the sample covariance matrix. The remaining targets behave very differently across the four risk-based portfolios: the Common Covariance is the worst in both the Minimum Variance and Maximum Diversification and the EWMA is the worst in both remaining portfolios. The Market Model and the Constant Correlation do not improve much from the sample estimator across all portfolios.

Looking at the second row ($n = 120$), the Identity is the most efficient target, reducing the distance between estimated and *true* portfolio weights in all the considered portfolios. The Variance Identity is also very efficient in Minimum Variance and Maximum Diversification portfolios, while the remaining targets show similar results as in the previous case. The same conclusions apply for the case $n = 180$.

When the number of observations is equal to or higher than $n = 3000$, results do not change much. The Identity, the Variance Identity, the Market model, and the Constant Correlation are the most efficient target matrices towards which to shrink, while the EWMA is the worst for both Inverse Volatility and Equal-Risk-Contribution portfolios and the Common Covariance is the worst for the Minimum Variance and Maximum Diversification ones.

In conclusion, for the Minimum Variance portfolio, the Common Covariance should not be used, since it always produces weights very unstable and distant from the *true* ones. At the same time, the EWMA should not be used to shrink the covariance matrix in the Inverse Volatility and Equal-Risk-Contribution portfolios. The most convenient matrices towards which to shrink are the Identity and the Variance Identity. Overall, the Minimum Variance and the Maximum Diversification portfolio weights gain more from shrinkage than those of the Inverse Volatility and Equal-Risk-Contribution allocations.

4.5 Conclusion

In this article, we provide a comprehensive overview of shrinkage in risk-based portfolios. Portfolios solely based on the asset returns covariance matrix are usually perceived as “robust” since they avoid estimating the asset returns mean. However, they still suffer from estimation error when the sample estimator is used, causing misspecification in the portfolio weights. Shrinkage estimators have been proved to reduce the estimation error by pulling the sample covariance towards a more structured target.

By means of an extensive Monte Carlo study, we compare six different target matrices: the Identity, the Variance Identity, the Single-index model, the Common Covariance, the Constant Correlation, and the Exponential Weighted Moving Average, respectively. We do so considering their effects on weights for the Minimum Variance, Inverse Volatility, Equal-risk-contribution, and Maximum diversification portfolios. Moreover, we control for the whole shrinkage intensity spectrum and for dataset size, changing observation length and number of assets. Therefore, we are able to (1) assess estimators' statistical properties and similarity with the *true* target matrix; (2) address the problem of how selecting a specific target estimator affects the portfolio weights.

Regarding point (1), the findings suggest the Identity and the Variance Identity matrices hold the best statistical properties, being well conditioned across all the combinations of observations/assets, especially for high-dimensional datasets. Nevertheless, these targets are both not very similar to the *true* target matrix. The Single-Index and the Constant Correlation target matrices show the greater similarity with the *true* target matrix, minimising the Frobenius norm, albeit being poorly conditioned when observations and assets share similar sizes. Turning to point (2), the Identity attains the best results in terms of distance reduction between the *true* and estimated portfolio weights for both the Minimum Variance and Maximum Diversification portfolio construction techniques. The Variance Identity shows a similar performance. Both estimators are also stable against shifts in the shrinkage intensity.

Overall, selecting the target matrix is very important, since we verified that there are large shifts in the distance between *true* and estimated portfolio weights when shrinking towards different targets. In risk-based portfolio allocations the Identity and the Variance Identity matrices represent the best target among the six considered in this study, especially in the case of Minimum Variance and Maximum Diversification portfolios. In fact, they are always well conditioned and outperform their competitor in deriving the most similar weights to the *true* ones.

Lastly, the findings confirm that the Minimum Variance and Maximum Diversification portfolios are more sensitive to misspecification in the covariance matrix, therefore, they benefit the most when the sample covariance matrix is shrunk. These findings are in line with what was previously found in [Ardia et al., 2017]: the Inverse Volatility and the Equal-Risk-Contribution are more robust to covariance misspecification; hence, allocations do not improve significantly when shrinkage is used.

Table 4.1: Literature review of target matrices. "SCVm" = sample covariance matrix. "N.A." = not available. "GMVP" = Global Minimum Variance Portfolio.

Reference	Matrix to Shrink	Target Matrix	Shrinkage Intensity	Portfolio Selection Rule	Research Question
[Ledoit and Wolf, 2003]	SCVm	Market Model and Variance Identity	Risk-function minimisation	Classical Markowitz problem	Portfolio Performance comparison
[Ledoit and Wolf, 2004a]	SCVm	Identity	Risk-function minimisation	N.A.	Theoretical paper to gauge the shrinkage asymptotic properties
[Ledoit and Wolf, 2004b]	SCVm	Constant Correlation Model	Optimal shrinkage constant	Classical Markowitz problem	Portfolio Performance comparison
[Briner and Connor, 2008]	Demeaned SCVm	Market Model	Same as [Ledoit and Wolf, 2004b]	N.A.	Analysis of the trade-off estimation error and model specification error
[Pantaleo et al., 2011]	SCVm	Market Model, Common Covariance and Constant Correlation Model	Unbiased estimator of [Schäfer and Strimmer, 2005]	Classical Markowitz problem	Portfolio Performance comparison
[Candelon et al., 2012]	SCVm	Market Model and Identity	Same as [Ledoit and Wolf, 2003]	Black-Litterman GMVP	Portfolio Performance comparison
[DeMiguel and Nogales, 2013]	SCVm	Scaled Identity	Expected quadratic loss and bootstrapping approach	Classical Markowitz problem	Comprehensive investigation of shrinkage estimators
[Ardia et al., 2017]	SCVm	Market Model	Same as [Ledoit and Wolf, 2003]	Risk-based portfolios	Theoretical paper to assess effect on risk-based weights

Table 4.2: *True* weights of the four risk-based portfolios.

Asset	Minimum Variance (MV)	Inverse Volatility (IV)	Equal–Risk–Contribution (ERC)	Maximum Diversification (MD)
Asset-1	0.500	0.400	0.448	0.506
Asset-2	0.500	0.400	0.374	0.385
Asset-3	0.000	0.200	0.177	0.108

Table 4.3: Frobenius norm for the portfolio weights. Values are averaged along with the shrinkage intensity (excluding the case $\delta = 0$). For each n, the first line reports the Frobenius norm for the sample covariance matrix. Abbreviations in use are: S for sample covariance; Id for identity matrix; VId for Variance Identity; SI for Single-Index; CV for Common Covariance; CC for Constant Correlation and EWMA for Exponentially Weighted Moving Average.

	$P = 10$				$P = 50$				$P = 100$			
	MV	IV	ERC	MD	MV	IV	ERC	MD	MV	IV	ERC	MD
Panel A: $n = 60$												
S	0.834	0.1585	0.1736	0.5842	0.7721	0.0573	0.0637	0.4933	0.7555	0.0409	0.0447	0.4565
Id	0.6863	0.1425	0.1528	0.5045	0.6215	0.0559	0.0631	0.3873	0.4967	0.0404	0.0451	0.3652
VId	0.6935	0.1583	0.1732	0.5176	0.5999	0.0567	0.0634	0.4092	0.5901	0.0404	0.0445	0.3686
SI	0.838	0.1585	0.1736	0.5678	0.7685	0.0573	0.0637	0.4709	0.75	0.0409	0.0447	0.4288
CV	1.2438	0.1583	0.1731	1.011	1.1484	0.0567	0.0628	0.9381	1.1386	0.0404	0.0438	0.9185
CC	0.8353	0.1585	0.1733	0.5361	0.7808	0.0573	0.0635	0.4328	0.7663	0.0409	0.0445	0.3922
EWMA	0.8473	0.1593	0.1745	0.595	0.7811	0.0575	0.064	0.5142	0.7325	0.0411	0.045	0.4431
Panel B: $n = 120$												
S	0.9064	0.0877	0.0989	0.4649	0.7814	0.059	0.0656	0.5065	0.6519	0.0424	0.0472	0.4332
Id	0.8157	0.087	0.0983	0.4256	0.6259	0.0613	0.0688	0.4354	0.6307	0.0389	0.0431	0.328
VId	0.8235	0.0871	0.0985	0.4284	0.6259	0.0613	0.0688	0.4354	0.489	0.0421	0.0471	0.3712
SI	0.9097	0.0877	0.0989	0.4563	0.7777	0.059	0.0656	0.4925	0.6458	0.0424	0.0472	0.419
CV	1.3269	0.0871	0.0982	0.9667	1.1806	0.0587	0.0651	1.0138	1.0974	0.0421	0.0467	0.8951
CC	0.905	0.0877	0.0988	0.4357	0.7822	0.059	0.0655	0.4636	0.6566	0.0424	0.0471	0.3856
EWMA	0.9281	0.0883	0.0996	0.4859	0.7994	0.0592	0.0658	0.5246	0.6788	0.0427	0.0475	0.4601
Panel C: $n = 180$												
S	0.7989	0.1311	0.1423	0.5007	0.7932	0.0564	0.0627	0.4631	0.6905	0.0404	0.044	0.4065
Id	0.7206	0.1308	0.142	0.4736	0.6705	0.0562	0.0625	0.405	0.5477	0.0375	0.0399	0.3748
VId	0.7273	0.1308	0.1421	0.4757	0.6838	0.0562	0.0626	0.4127	0.5754	0.0402	0.044	0.3556
SI	0.8001	0.1311	0.1423	0.4954	0.7904	0.0564	0.0627	0.4545	0.6873	0.0404	0.044	0.3982
CV	1.2715	0.1308	0.1419	0.9961	1.2073	0.0562	0.0624	0.9988	1.1422	0.0402	0.0437	0.8705
CC	0.7957	0.1311	0.1423	0.4803	0.792	0.0564	0.0626	0.4259	0.692	0.0404	0.044	0.3672
EWMA	0.8415	0.1322	0.1435	0.526	0.8284	0.0567	0.0631	0.5005	0.7206	0.0408	0.0445	0.4429
Panel D: $n = 3000$												
S	0.7504	0.1476	0.1596	0.3957	0.734	0.049	0.0539	0.3988	0.513	0.0384	0.0428	0.3259
Id	0.7441	0.1477	0.1597	0.3946	0.7009	0.049	0.0539	0.3872	0.4615	0.0384	0.0428	0.3096
VId	0.7437	0.1477	0.1596	0.3945	0.7043	0.049	0.0539	0.3886	0.4673	0.0384	0.0428	0.312
SI	0.7516	0.1476	0.1596	0.3955	0.7339	0.049	0.0539	0.3984	0.5123	0.0384	0.0428	0.3252
CV	1.2864	0.1477	0.1597	0.963	1.2281	0.049	0.0538	0.9954	1.1041	0.0384	0.0428	0.6822
CC	0.7488	0.1476	0.1596	0.3949	0.7316	0.049	0.0539	0.3904	0.5096	0.0384	0.0428	0.3143
EWMA	0.8563	0.1489	0.1611	0.4452	0.8161	0.0497	0.0547	0.4652	0.6244	0.0389	0.0435	0.4076
Panel E: $n = 6000$												
S	0.9672	0.1302	0.1409	0.4821	0.5737	0.0539	0.0589	0.3481	0.5772	0.0402	0.0437	0.3436
Id	0.9496	0.1301	0.1408	0.4813	0.6095	0.0575	0.0639	0.4076	0.5449	0.0402	0.0437	0.3342
VId	0.951	0.1301	0.1409	0.4815	0.5419	0.054	0.0589	0.3401	0.5483	0.0402	0.0437	0.3354
SI	0.9688	0.1302	0.1409	0.482	0.574	0.0539	0.0589	0.3479	0.5772	0.0402	0.0437	0.3434
CV	1.4142	0.1301	0.1408	1.0034	1.1436	0.054	0.0589	0.9706	1.1422	0.0402	0.0437	0.7031
CC	0.9656	0.1302	0.1409	0.4814	0.5709	0.0539	0.0589	0.3415	0.575	0.0402	0.0437	0.3368
EWMA	1.0432	0.1312	0.1422	0.5232	0.6946	0.0547	0.0599	0.4319	0.681	0.0407	0.0444	0.4229

Table 4.4: Frobenius norm for the portfolio weights. Values corresponds to the optimal shrinkage intensity, listed after the Frobenius norm for each portfolio. We report values for the sample covariance matrix ($\delta = 0$) separately in the first row of each panel. For each n, the first line reports the Frobenius norm for the sample covariance matrix. Abbreviations used are: S for sample covariance; Id for identity matrix; VId for Variance Identity; SI for Single-Index; CV for Common Covariance; CC for Constant Correlation and EWMA for Exponentially Weighted Moving Average.

	$P = 10$				$P = 50$				$P = 100$			
	MV	IV	ERC	MD	MV	IV	ERC	MD	MV	IV	ERC	MD
Panel A: $n = 60$												
S	0.8340	0.1585	0.1736	0.5842	0.7721	0.0573	0.0637	0.4933	0.7555	0.0409	0.0447	0.4565
Id	0.6778	0.1424	0.1525	0.501	0.5997	0.0558	0.0624	0.3704	0.471	0.0403	0.0446	0.3462
VId	0.6689	0.1581	0.173	0.5084	0.5539	0.0565	0.0627	0.3795	0.5428	0.0402	0.0437	0.3331
SI	0.8345	0.1585	0.1735	0.558	0.7666	0.0573	0.0637	0.4633	0.7479	0.0409	0.0447	0.4195
CV	1.2392	0.1581	0.1729	0.509	1.117	0.0565	0.0627	0.3795	1.1068	0.0402	0.0437	0.3331
CC	0.8335	0.1585	0.1731	0.5081	0.7733	0.0573	0.0634	0.3795	0.757	0.0409	0.0444	0.3332
EWMA	0.8331	0.1586	0.1737	0.5852	0.7706	0.0573	0.0637	0.4953	0.7213	0.0409	0.0447	0.4395
Panel B: $n = 120$												
S	0.9064	0.0877	0.0989	0.4649	0.7814	0.059	0.0656	0.5065	0.6519	0.0424	0.0472	0.4332
Id	0.8121	0.087	0.0981	0.4241	0.6119	0.0613	0.0685	0.4255	0.613	0.0388	0.0428	0.3111
VId	0.8121	0.087	0.0982	0.4242	0.6119	0.0613	0.0685	0.4255	0.4425	0.042	0.0467	0.3445
SI	0.907	0.0877	0.0989	0.4526	0.776	0.059	0.0656	0.4872	0.6431	0.0424	0.0472	0.414
CV	1.3269	0.087	0.0981	0.4245	1.1756	0.0586	0.0651	0.4302	1.0916	0.042	0.0467	0.3445
CC	0.9043	0.0877	0.0987	0.4241	0.781	0.059	0.0654	0.4302	0.6527	0.0424	0.0471	0.3446
EWMA	0.9052	0.0876	0.0988	0.4651	0.7797	0.0589	0.0655	0.5056	0.6554	0.0424	0.0472	0.4331
Panel C: $n = 180$												
S	0.7989	0.1311	0.1423	0.5007	0.7932	0.0564	0.0627	0.4631	0.6905	0.0404	0.044	0.4065
Id	0.7177	0.1307	0.1419	0.4724	0.6613	0.0562	0.0624	0.3977	0.534	0.0375	0.0398	0.3645
VId	0.718	0.1307	0.1419	0.4724	0.6614	0.0562	0.0624	0.3979	0.5428	0.0402	0.0437	0.3331
SI	0.799	0.1311	0.1423	0.4929	0.7897	0.0564	0.0627	0.4515	0.6863	0.0404	0.044	0.3955
CV	1.2715	0.1307	0.1418	0.4724	1.2073	0.0562	0.0624	0.3979	1.1422	0.0402	0.0437	0.3331
CC	0.7942	0.1311	0.1422	0.4725	0.7912	0.0564	0.0626	0.3977	0.6904	0.0404	0.0439	0.3331
EWMA	0.8035	0.1312	0.1424	0.5008	0.7951	0.0564	0.0626	0.4653	0.6938	0.0404	0.044	0.4074
Panel D: $n = 3000$												
S	0.7504	0.1476	0.1596	0.3957	0.734	0.049	0.0539	0.3988	0.513	0.0384	0.0428	0.3259
Id	0.7425	0.1477	0.1596	0.3941	0.6988	0.049	0.0538	0.3859	0.4573	0.0384	0.0428	0.3072
VId	0.7426	0.1476	0.1596	0.3941	0.6988	0.049	0.0538	0.3859	0.4573	0.0384	0.0428	0.3072
SI	0.7506	0.1476	0.1596	0.3953	0.7339	0.049	0.0539	0.3983	0.512	0.0384	0.0428	0.325
CV	1.2864	0.1476	0.1596	0.3951	1.2281	0.049	0.0538	0.3859	1.1041	0.0384	0.0428	0.3072
CC	0.7477	0.1476	0.1596	0.3946	0.7299	0.049	0.0539	0.386	0.5073	0.0384	0.0428	0.3072
EWMA	0.7615	0.1477	0.1597	0.3981	0.7439	0.0491	0.0539	0.4043	0.5263	0.0384	0.0429	0.3346
Panel E: $n = 6000$												
S	0.9672	0.1302	0.1409	0.4821	0.5737	0.0539	0.0589	0.3481	0.5772	0.0402	0.0437	0.3436
Id	0.9486	0.13	0.1408	0.4811	0.6085	0.0575	0.0639	0.4072	0.5428	0.0402	0.0437	0.3331
VId	0.9486	0.13	0.1408	0.4811	0.5365	0.054	0.0589	0.3381	0.5428	0.0402	0.0437	0.3331
SI	0.9675	0.1302	0.1409	0.482	0.5738	0.0539	0.0589	0.3478	0.5772	0.0402	0.0437	0.3433
CV	1.4142	0.13	0.1408	0.4811	1.1436	0.054	0.0589	0.3381	1.1422	0.0402	0.0437	0.3331
CC	0.9644	0.1302	0.1409	0.4812	0.5687	0.0539	0.0589	0.3381	0.5733	0.0402	0.0437	0.3331
EWMA	0.9765	0.1302	0.1409	0.4832	0.5901	0.054	0.059	0.3561	0.59	0.0402	0.0438	0.3524

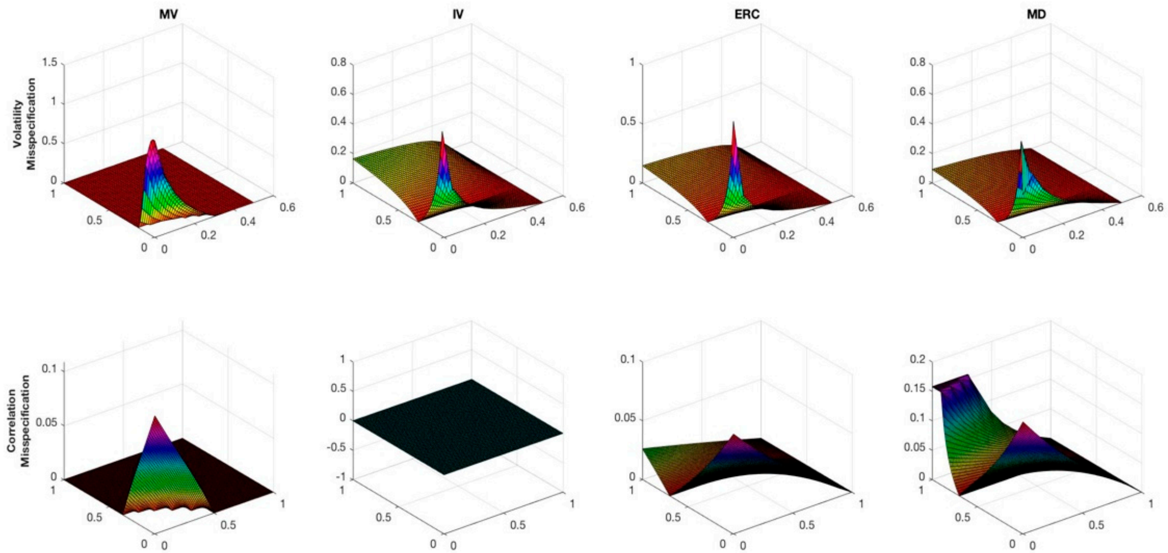

Figure 4.1: Frobenius norm between *true* and estimated weights; first row reports misspecification in volatility, while second row in correlation. The surfaces' three dimensions are: the shrinkage intensity in y axis (from 0 to 1); the misspecification in the volatility (from 0 to 0.5) or in the correlation (from 0 to 1) in x axis and the Frobenius norm in z axis. Each column refers to a specific risk-based portfolio. From the left to the right: Minimum Variance (MV), Inverse Volatility (IV), Equal-Risk-Contribution (ERC), Maximum Diversification (MD), respectively.

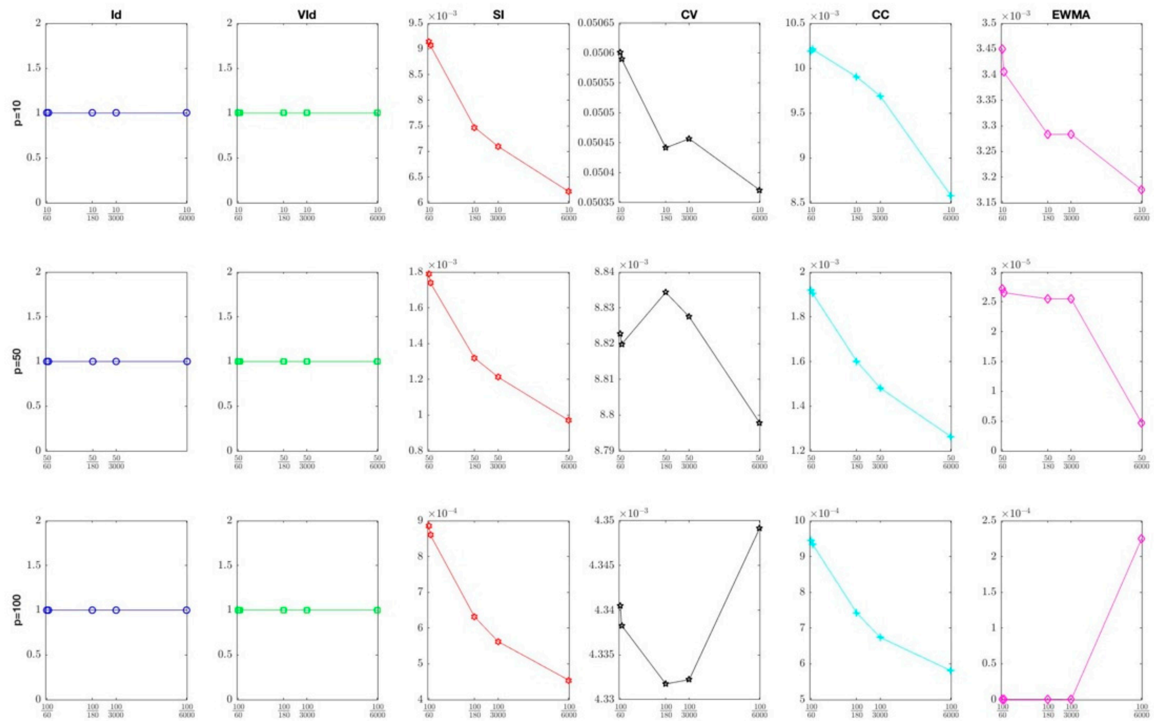

Figure 4.2: The reciprocal 1-norm condition number (y-axis) as the p/n ratio moves from $\frac{p}{60}$ to $\frac{p}{6000}$ (x-axis). Each column corresponds to a specific target matrix: from left to right, the Identity (Id): blue circle-shaped; the Variance Identity (VId): green square-shaped; the Single-Index (SI): red hexagram-shaped; the Common Covariance (CV): black star-shaped; the Constant Correlation (CC): cyan plus-shaped; and the Exponential Weighted Moving Average (EWMA): magenta diamond-shaped, respectively. Each row corresponds to a different p: in ascendant order from 10 (first row) to 100 (third row).

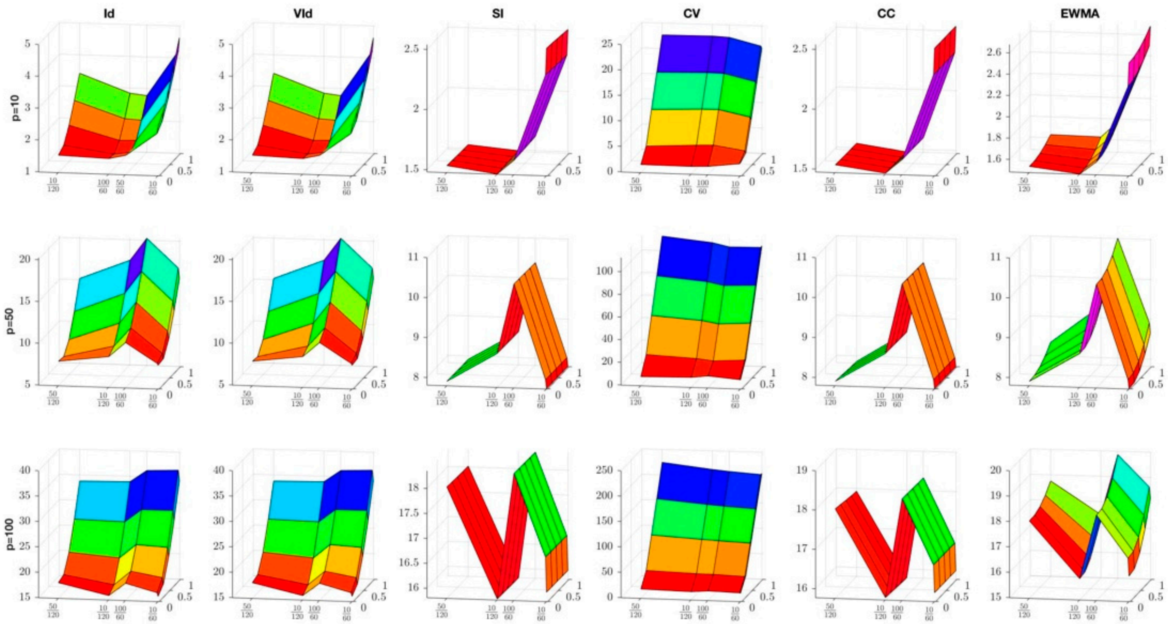

Figure 4.3: Surfaces representing the Frobenius norm (z-axis) between the *true* and the estimated target matrices, considering the shrinkage intensity (y-axis) and the p/n ratio (x-axis). Each column corresponds to a specific target matrix: from left to right, the Identity (Id), the Variance Identity (VId), the Single-Index (SI), the Common Covariance (CV), the Constant Correlation (CC), and the EWMA, respectively. Each row corresponds to a different p: in ascendant order from $p = 10$ (first row) to $p = 100$ (third row).

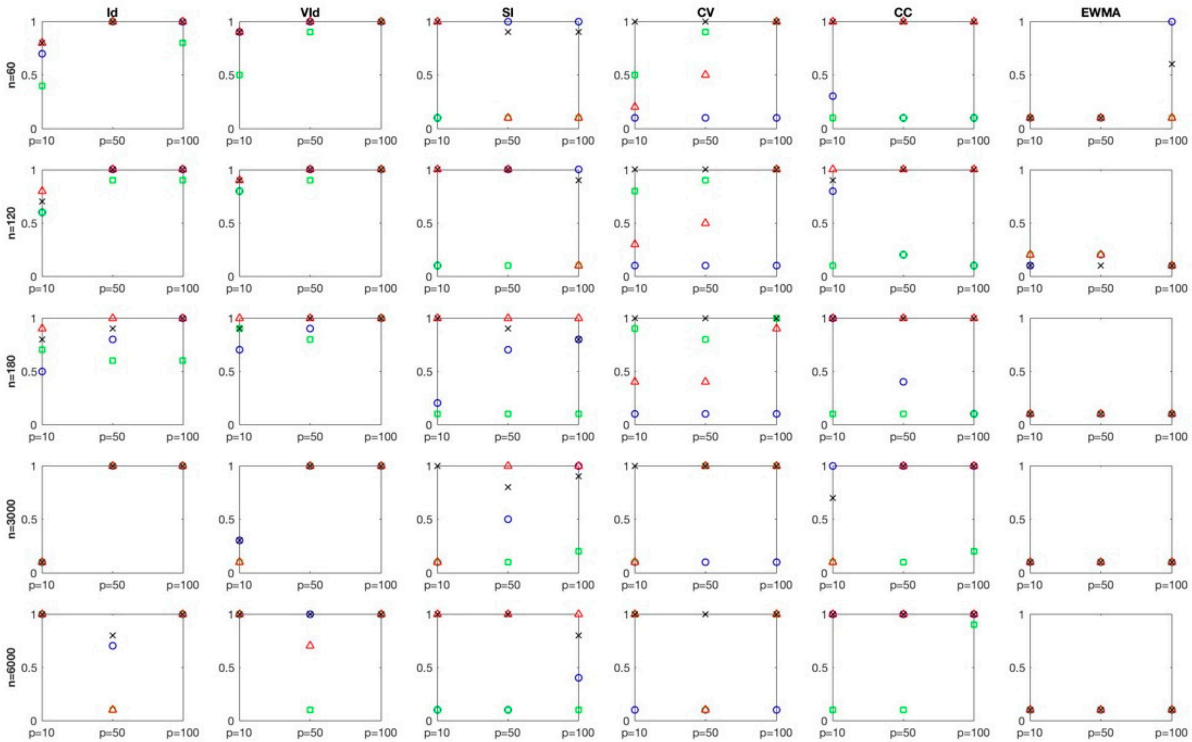

Figure 4.4: Optimal shrinkage intensity parameters for which the Frobenius norm is minimised. Each column corresponds to a specific target matrix: from left to right, the Identity, the Variance Identity, the Single-Index, the Common Covariance, the Constant Correlation, and the EWMA, respectively. Each row corresponds to a different n: in ascendant order from $n = 60$ (first row) to $p = 6000$ (fifth row). In each subplot, the MV portfolio is blue circle-shaped; the IV is green-square shaped; the ERC is red-triangle shaped; and the MD is black-cross shaped.

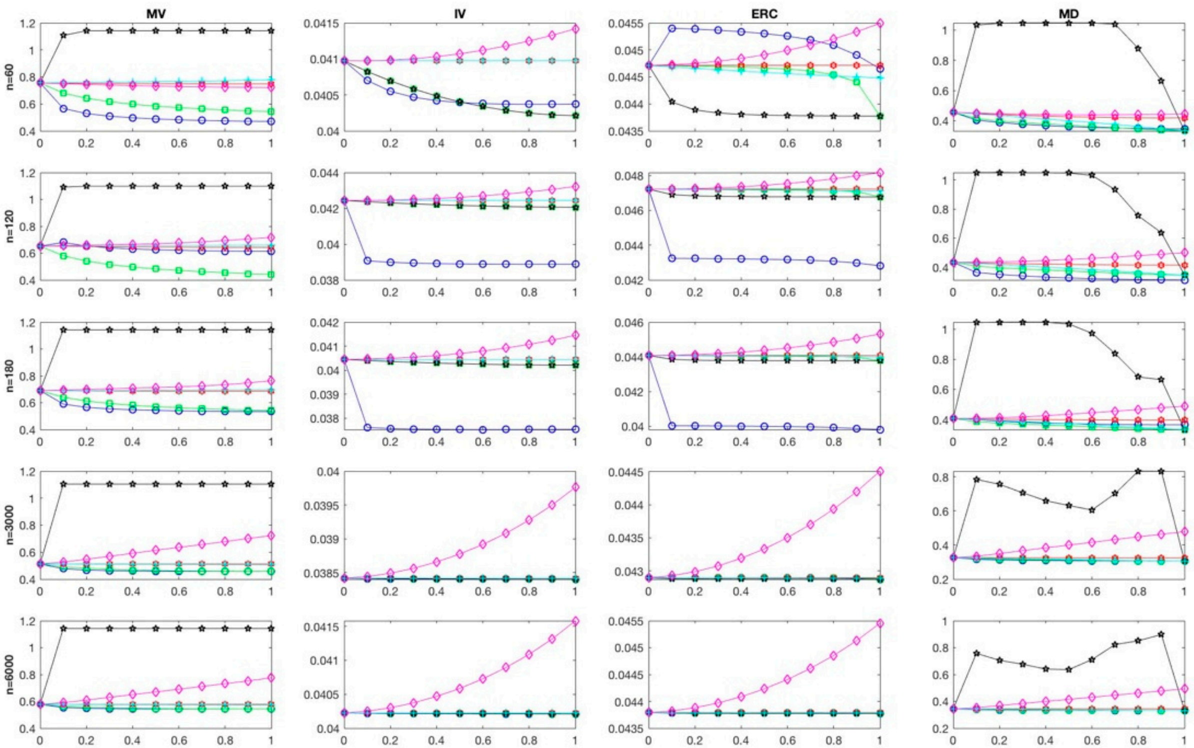

Figure 4.5: Frobenius norm for portfolio weights with regard to the shrinkage intensity parameter, when $p = 100$. Curves represent the six target estimators: the Identity (Id): blue circle-shaped; the Variance Identity (VId): green square-shaped; the Single-Index (SI): red hexagram-shaped; the Common Covariance (CV): black star-shaped; the Constant Correlation (CC): cyan plus-shaped; and the Exponential Weighted Moving Average (EWMA): magenta diamond-shaped, respectively.

Chapter 5

Conclusions and Future Research

The main motivation of this research was to improve the estimation of the covariance matrix for portfolio applications. To such aim, we built on existing econometric methodologies and proposed the adoption of new (in our field) ones, but there is still enough space left to consider this as an ongoing research for future times. In fact, the reviewed literature clearly highlighted that misspecification effects in the covariance matrix cannot be neglected. We reckon this as the need of improving existing solutions.

In details, our contributions to the literature can be summarised as follows. In Chapter 2, we developed and compared two novel methodologies for hedging in the multivariate GARCH framework. This filled a lack in the literature about hedging energy commodity portfolios, where hedging when the covariance matrix is conditional and heteroskedastic was limited to singular spot positions. The numerical illustration gave a comprehensive overview about how volatility and covariance misspecification impacts on the hedged portfolio, disentangling how these effects work under both approaches. Through the empirical case study, we were able to compare the two approaches, while observing interesting insights even for investors willing to hedge energy commodities. The empirical case study still leaves room to further investigation: a more comprehensive analysis covering different periods, commodities or even simulated datasets could help in generalising the results. In Chapter 3, we introduced the Minimum Regularised Covariance Determinant estimator in our field. Given the versatility of this estimator, it served a twofold aim: first, we have enriched the literature about estimation error in the precision matrix, improving asset allocation under the Global Minimum Variance portfolio. In this case, we offered a comparison against the sample covariance matrix estimator which is very comprehensive. In fact, we either tested the two with an extensive Monte Carlo study, and then we used five real investment universes to show how the out–of–sample performance and general stability of Global Minimum Variance portfolio weights are improved. Second, we have contributed to the interest rates literature by analysing five alternative estimators for the covariance matrix with the aim of giving insights about risk management and portfolio building for fixed income instruments. To this extent, a practical case study analysing the main factors that drive the volatility of the US yield curve has been proposed. Several directions for future research can be identified from each application of the MRCD. First, it would be then particularly interesting to test the MRCD against alternative estimators for the covariance matrix inverse. Recently proposed estimators as the shrinkage and the LASSO have performed well in enhancing the out–of–sample performance of the Global Minimum Variance portfolio. Second, interest rate modelling should be enhanced providing a wider array of applications and datasets – both real and simulated – where to test the benefits of MRCD covariance matrix. In Chapter 4, we focused our attention on misspecification errors in risk-based portfolios. To such aim, we improved the shrinkage technique by focusing on the target matrix. Through a numerical illustration, we assessed the impacts of a misspecificated shrinkage target matrix in the resulting risk–based portfolio weights. We were able to disentangling the impact produced by a shift in the volatility or in the covariance, finding that the Minimum Variance and the Maximum Diversification portfolios are more sensitive to these kinds of errors than the Inverse Volatility and the Equal–Risk–Contribution portfolios. Moreover, we compared six alternative estimators for the target matrix in a comprehensive Monte Carlo study. Evidence suggested that the Identity and the Variance Identity target matrix yield superior statistical properties than their competitors, allowing estimated risk-based portfolio weights to be more close to their population counterparts. We found room to improve this chapter at least in two directions. One by employing the same investigation approach to other portfolios rather than risk–based ones. Additionally, it would be interesting to assess their effect on several real investment universe. We leave these developments to future research.

Bibliography

Michael Anderberg. Cluster analysis for applications. Technical report, Office of the Assistant for Study Support Kirtland AFB N MEX, 1973.

Theodore Anderson. Asymptotic theory for principal component analysis. *Annals of Mathematical Statistics*, 34(1):122–148, 1963.

David Ardia, Guido Bolliger, Kris Boudt, and Jean-Philippe Gagnon-Fleury. The impact of covariance misspecification in risk-based portfolios. *Annals of Operations Research*, 254 (1-2):1–16, 2017.

Richard Baillie and Robert Myers. Bivariate GARCH estimation of the optimal commodity futures hedge. *Journal of Applied Econometrics*, 6(2):109–124, 1991.

Luc Bauwens, Sébastien Laurent, and Jeroen VK Rombouts. Multivariate garch models: a survey. *Journal of applied econometrics*, 21(1):79–109, 2006.

Simon Benninga, Rafael Eldor, and Itzhak Zilcha. The optimal hedge ratio in unbiased futures markets. *Journal of futures markets*, 4(2):155–159, 1984.

Philippe Bertrand and Vincent Lapointe. Risk-based strategies: the social responsibility of investment universes does matter. *Annals of Operations Research*, 262(2):413–429, 2018.

Fischer Black and Robert Litterman. Global portfolio optimization. *Financial Analysts Journal*, 48:28–43, 1992.

Taras Bodnar, Nestor Parolya, and Wolfgang Schmid. Estimation of the global minimum variance portfolio in high dimensions. *European Journal of Operational Research*, 266(1): 371–390, 2018.

Tim Bollerslev. On the correlation structure for the generalized autoregressive conditional heteroskedastic process. *Journal of Time Series Analysis*, 9(2):121–131, 1988.

Tim Bollerslev. Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. *The review of economics and statistics*, pages 498–505, 1990.

Kris Boudt, Peter Rousseeuw, Steven Vanduffel, and Tim Verdonck. The minimum regularized covariance determinant estimator. Available at SSRN: https://ssrn.com/abstract=2905259 or http://dx.doi.org/, 2018.

Beat Briner and Gregory Connor. How much structure is best? a comparison of market model, factor model and unstructured equity covariance matrices. *Journal of Risk*, 10: 3–30, 2008.

José Agulló Candela. Exact iterative computation of the multivariate minimum volume ellipsoid estimator with a branch and bound algorithm. In *COMPSTAT*, pages 175–180. Springer, 1996.

Bertrand Candelon, Christophe Hurlin, and Sessi Tokpavi. Sampling error and double shrinkage estimation of minimum variance portfolios. *Journal of Empirical Finance*, 19(4): 511–527, 2012.

Stephen Cecchetti, Robert Cumby, and Stephen Figlewski. Estimation of the optimal futures hedge. *The Review of Economics and Statistics*, pages 623–630, 1988.

Wing Chan and Denise Young. Jumping hedges: an examination of movements in copper spot and futures markets. *Journal of Futures Markets*, 26(2):169–188, 2006.

Chia-Lin Chang, Michael McAleer, and Roengchai Tansuchat. Crude oil hedging strategies using dynamic multivariate GARCH. *Energy Economics*, 33(5):912–923, 2011.

Chia-Lin Chang, Lydia Gonzãlez-Serrano, and Juan-Angel Jimenez-Martin. Currency hedging strategies using dynamic multivariate GARCH. *Mathematics and Computers in Simulation*, 94:164–182, 2013.

Chiao-Yi Chang, Jing-Yi Lai, and I-Yuan Chuang. Futures hedging effectiveness under the segmentation of bear/bull energy markets. *Energy Economics*, 32(2):442–449, 2010.

Sheng-Syan Chen, Cheng-Few Lee, and Keshab Shrestha. On a Mean Generalized Semi-variance Approach to Determining the Hedge Ratio. *Journal of Futures Markets*, 21(6): 581–598, 2001.

Sheng-Syan Chen, Cheng-few Lee, and Keshab Shrestha. Futures hedge ratios: a review. In *Encyclopedia of Finance*, pages 871–890. Springer, 2013.

Sherman Cheung, Clarence Kwan, and Patrick Yip. The hedging effectiveness of options and futures: A mean-Gini approach. *Journal of Futures Markets*, 10(1):61–73, 1990.

Vijay Kumar Chopra and William Ziemba. The effect of errors in means, variances, and covariances on optimal portfolio choice. *The Journal of Portfolio Management*, 19:6–11, 1993.

Taufiq Choudhry. Short-run deviations and time-varying hedge ratios: evidence from agricultural futures markets. *International Review of Financial Analysis*, 18(1):58–65, 2009.

Yves Choueifaty and Yves Coignard. Toward maximum diversification. *The Journal of Portfolio Management*, 35:40–51, 2008.

Rama Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2):223—-236, 2001.

Christophe Croux and Gentiane Haesbroeck. Influence function and efficiency of the minimum covariance determinant scatter matrix estimator. *Journal of Multivariate Analysis*, 71(2): 161–190, 1999.

Alberto Martin-Utrera DeMiguel, Victor and Francisco J. Nogales. Size matters: Optimal calibration of shrinkage estimators for portfolio selection. *Journal of Banking and Finance*, 37:3018–34, 2013.

Lorenzo Garlappi DeMiguel, Victor and Raman Uppal. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies*, 22:1915–53, 2009.

Victor DeMiguel and Francisco J Nogales. Portfolio selection with robust estimation. *Operations Research*, 57(3):560–577, 2009.

Victor DeMiguel, Lorenzo Garlappi, and Raman Uppal. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The Review of Financial Studies*, 22(5): 1915–1953, 2007.

Victor DeMiguel, Lorenzo Garlappi, Francisco J Nogales, and Raman Uppal. A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science*, 55(5):798–812, 2009.

Victor DeMiguel, Alberto Martin-Utrera, and Francisco J Nogales. Size matters: Optimal calibration of shrinkage estimators for portfolio selection. *Journal of Banking & Finance*, 37(8):3018–3034, 2013.

Louis Ederington. The hedging performance of the new futures markets. *The Journal of Finance*, 34(1):157–170, 1979.

Robert Engle. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350, 2002.

Robert Engle and Kenneth Kroner. Multivariate simultaneous generalized ARCH. *Econometric theory*, 11(01):122–150, 1995.

Robert Engle and Kevin Sheppard. Theoretical and empirical properties of dynamic conditional correlation multivariate garch. Technical report, National Bureau of Economic Research, 2001.

Frank Fabozzi, Francis Gupta, and Harry M Markowitz. The legacy of modern portfolio theory. *Journal of Investing*, 11(3):7–22, 2002.

Jianqing Fan, Heng Peng, and Tao Huang. Semilinear high-dimensional model for normalization of microarray data: a theoretical analysis and partial consistency. *Journal of the American Statistical Association*, 100(471):781–796, 2005.

Thomas Fisher and Xiaoqian Sun. Improved stein-type shrinkage estimators for the high-dimensional multivariate normal covariance matrix. *Computational Statistics & Data Analysis*, 55(5):1909–1918, 2011.

Christos Floros and Dimitrios Vougas. Hedge ratios in Greek stock index futures market. *Applied Financial Economics*, 14(15):1125–1136, 2004.

Gabriel Frahm and Uwe Jaekel. Random matrix theory and robust covariance matrix estimation for financial data. *arXiv preprint physics/0503007*, 2005.

Gabriel Frahm and Christoph Memmel. Dominating estimators for minimum-variance portfolios. *Journal of Econometrics*, 159(2):289–302, 2010.

Lawrence Glosten, Ravi Jagannathan, and David Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, 48(5):1779–1801, 1993.

Douglas Hawkins and Geoffrey McLachlan. High-breakdown linear discriminant analysis. *Journal of the American Statistical Association*, 92(437):136–143, 1997.

Harold Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of educational psychology*, 24(6):417, 1933.

Charles Howard and Louis D'Antonio. A risk-return measure of hedging effectiveness. *Journal of Financial and Quantitative Analysis*, 19(01):101–112, 1984.

Mia Hubert, Peter Rousseeuw, and Karlien Vanden Branden. Robpca: a new approach to robust principal component analysis. *Technometrics*, 47(1):64–79, 2005.

Ravi Jagannathan and Tongshu Ma. Risk reduction in large portfolios: Why imposing the wrong constraints helps. *Journal of Finance*, 58:1651–84, 2003.

Willard James and Charles Stein. Estimation with quadratic loss. In *Proceedings of the 4th Berkeley Symposium on Probability and Statistics, Volume 1*. Paper presented at Fourth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA,, June 20–July 30 1961.

Leland Johnson. The theory of hedging and speculation in commodity futures. *The Review of Economic Studies*, 27(3):139–151, 1960.

Philippe Jorion. Bayes-stein estimation for portfolio analysis. *The Journal of Financial and Quantitative Analysis*, 21:279–92, 1986.

Sang Hoon Kang and Seong-Min Yoon. Modeling and forecasting the volatility of petroleum futures prices. *Energy Economics*, 36:354–362, 2013.

Petter Kolm, Reha Tütüncü, and Frank Fabozzi. 60 years of portfolio optimization: Practical challenges and current trends. *European Journal of Operational Research*, 234(2):356–371, 2014.

Apostolos Kourtis, George Dotsis, and Raphael Markellos. Parameter uncertainty in portfolio selection: Shrinking the inverse covariance matrix. *Journal of Banking & Finance*, 36(9): 2522–2531, 2012.

Kenneth F Kroner and Jahangir Sultan. Time-varying distributions and dynamic hedging with foreign currency futures. *Journal of financial and quantitative analysis*, 28(04): 535–551, 1993.

Jason Laws and John Thompson. Hedging effectiveness of stock index futures. *European Journal of Operational Research*, 163(1):177–191, 2005.

Olivier Ledoit and Sandrine Péché. Eigenvectors of some large sample covariance matrix ensembles. *Probability Theory and Related Fields*, 151(1-2):233–264, 2011.

Olivier Ledoit and Michael Wolf. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance*, 10:603–21, 2003.

Olivier Ledoit and Michael Wolf. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88:365–411, 2004a.

Olivier Ledoit and Michael Wolf. Honey, I shrunk the sample covariance matrix. *Journal of Portfolio Management*, 30:110–19, 2004b.

Olivier Ledoit and Michael Wolf. Nonlinear shrinkage estimation of large-dimensional covariance matrices. *The Annals of Statistics*, 40(2):1024–1060, 2012.

Sergio H Lence, Kevin L Kimle, and Marvin L Hayenga. A dynamic minimum variance hedge. *American Journal of Agricultural Economics*, 75(4):1063–1071, 1993.

Raul Leote, Xiao Lu, and Pierre Moulin. Demystifying equity risk-based strategies: A simple alpha plus beta description. *Journal of Portfolio Management*, 38(3):56–70, 2012.

Donald Lien, Yiu Kuen Tse, and Albert KC Tsui. Evaluating the hedging performance of the constant-correlation GARCH model. *Applied Financial Economics*, 12(11):791–798, 2002.

John Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. In *Stochastic Optimization Models in Finance*, pages 131–155. Elsevier, 1975.

Robert Litterman and Josè Scheinkman. Common factors affecting bond returns. *The Journal of Fixed Income*, 1(1):54–61, 1991.

Craig MacKinlay and Lubos Pastor. Asset pricing models: Implications for expected returns and portfolio selection. *Review of Financial Studies*, 13:883–916, 2000.

Sébastien Maillard, Thierry Roncalli, and Jérôme Teïletche. The properties of equally-weighted risk contributions portfolios. *Journal of Portfolio Management*, 36(4):60—70, 2010.

Bertrand Maillet, Sessi Tokpavi, and Benoit Vaucher. Global minimum variance portfolio optimisation under some model risk: A robust regression-based approach. *European Journal of Operational Research*, 244(1):289–299, 2015.

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

Harry Markowitz. The optimization of a quadratic function subject to linear constraints. *Naval Research Logistics Quarterly*, 3:111–33, 1956.

Lionel Martellini, Philippe Priaulet, and Stéphane Priaulet. *Fixed-income securities: valuation, risk management and portfolio strategies*, volume 237. John Wiley & Sons, 2003.

Vladimir Alexandrovich Marčenko and Leonid Andreevich Pastur. Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1:507–36, 1967.

Madan Mehta. *Random matrices*, volume 142. Elsevier, 2004.

Robert Merton. On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, 8(4):323–361, 1980.

Richard Michaud. The markowitz optimization enigma: Is optimized optimal? *Financial Analysts Journal*, 45:31–42, 1989.

Richard Michaud and Robert Michaud. *Efficient asset management: a practical guide to stock portfolio optimization and asset allocation.* Oxford University Press, 2008.

Gyu-Hyen Moon, Wei-Choun Yu, and Chung-Hyo Hong. Dynamic hedging performance with the evaluation of multivariate GARCH models: evidence from KOSTAR index futures. *Applied Economics Letters*, 16(9):913–919, 2009.

Giancarlo Moschini and Robert Myers. Testing for constant hedge ratios in commodity markets: a multivariate GARCH approach. *Journal of Empirical Finance*, 9(5):589–603, 2002.

Diaa Noureldin, Neil Shephard, and Kevin Sheppard. Multivariate rotated arch models. *Journal of Econometrics*, 179(1):16–30, 2014.

Szilárd Pafka, Marc Potters, and Imre Kondor. Exponential weighting and random matrix theory based filtering of financial covariance matrices for portfolio optimization. *arXiv preprint cond-mat/0402573*, 2004.

Ester Pantaleo, Michele Tumminello, Fabrizio Lillo, and Rosario Mantegna. When do improved covariance matrix estimators enhance portfolio optimization? an empirical comparative study of nine estimators. *Quantitative Finance*, 11:1067–80, 2011.

Tae Park and Lorne Switzer. Bivariate GARCH estimation of the optimal hedge ratios for stock index futures: A note. *Journal of Futures Markets*, 15(1):61–67, 1995.

Cédric Perret-Gentil and Maria-Pia Victoria-Feser. Robust mean-variance portfolio selection. FAME Research Paper Series rp140, International Center for Financial Asset Management and Engineering, 2005.

Greet Pison, Peter Rousseeuw, Peter Filzmoser, and Christophe Croux. Robust factor analysis. *Journal of Multivariate Analysis*, 84(1):145–172, 2003.

Edward Qian. On the financial interpretation of risk contribution: Risk budgets do add up. *Journal of Investment Management*, 4:41–51, 2006.

RiskMetrics. Technical document. *JP Morgan Inc., New York, NY*, 1996.

Daren Roberts, Paul Bosse, Scott Donaldson, and Matthew Tufano. The portfolio currency–hedging decision, by objective and block by block. Technical report, Vanguard Research, 2018.

Stephen Ross. The arbitrage theory of capital asset pricing, 1976.

Peter Rousseeuw. Least median of squares regression. *Journal of the American Statistical Association*, 79(388):871–880, 1984.

Peter Rousseeuw and Katrien Van Driessen. A fast algorithm for the minimum covariance determinant estimator. *Technometrics*, 41(3):212–223, 1999.

Peter Rousseeuw and Victor Yohai. Robust regression by means of s-estimators. In *Robust and nonlinear time series analysis*, pages 256–272. Springer, 1984.

Peter Rousseeuw, Stefan Van Aelst, Katrien Van Driessen, and Jose A Gulló. Robust multivariate regression. *Technometrics*, 46(3):293–305, 2004.

Michael Schyns. Robust portfolio selection. *JSM Proceedings, Statistical Computing Section*, 2008.

Juliane Schäfer and Korbinian Strimmer. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Applications in Genetics and Molecular Biology*, 4:1175–89, 2005.

Marc Senneret, Yannick Malevergne, Patrice Abry, Gerald Perrin, Laurent Jaffres, and Vivienne Investissement. Covariance versus precision matrix estimation for efficient asset allocation. *IEEE Journal on Selected Topics in Signal Processing*, 10(6):982–993, 2016.

Haim Shalit. Mean-Gini hedging in futures markets. *Journal of Futures Markets*, 15(6): 617–635, 1995.

William Sharpe. A simplified model for portfolio analysis. *Management Science*, 9:277–93, 1963.

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

Annastiina Silvennoinen and Timo Teräsvirta. Multivariate GARCH models. In *Handbook of financial time series*, pages 201–229. Springer, 2009.

Charles Stein. Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. Technical report, Stanford University, 1956.

Guy Stevens. On the inverse of the covariance matrix in portfolio analysis. *The Journal of Finance*, 53(5):1821–1827, 1998.

Ruili Sun, Tiefeng Ma, and Shuangzhe Liu. A stein-type shrinkage estimator of the covariance matrix for portfolio selections. *Metrika*, pages 1–22, 2018.

Vincenzo Tola, Fabrizio Lillo, Mauro Gallegati, and Rosario Mantegna. Cluster analysis for portfolio optimization. *Journal of Economic Dynamics and Control*, 32(1):235–258, 2008.

Sandra Torri, Gabriele Paterlini and Rosella Giacometti. Sparse precision matrices for minimum variance portfolios. In *CFE 2017*, pages 160–160. GB, 2017.

Calum Turvey and Govindaray Nayak. The semivariance-minimizing hedge ratio. *Journal of Agricultural and Resource Economics*, pages 100–115, 2003.

Michael Wolf. Resampling vs. shrinkage for benchmarked managers. *Wilmott Magazine*, 1, 2007.

David Woodruff and David Rocke. Computable robust estimation of multivariate location and shape in high dimension using compound estimators. *Journal of the American Statistical Association*, 89(427):888–896, 1994.

Wenling Yang and David Allen. Multivariate GARCH hedge ratios and hedging effectiveness in Australian futures markets. *Accounting & Finance*, 45(2):301–321, 2005.