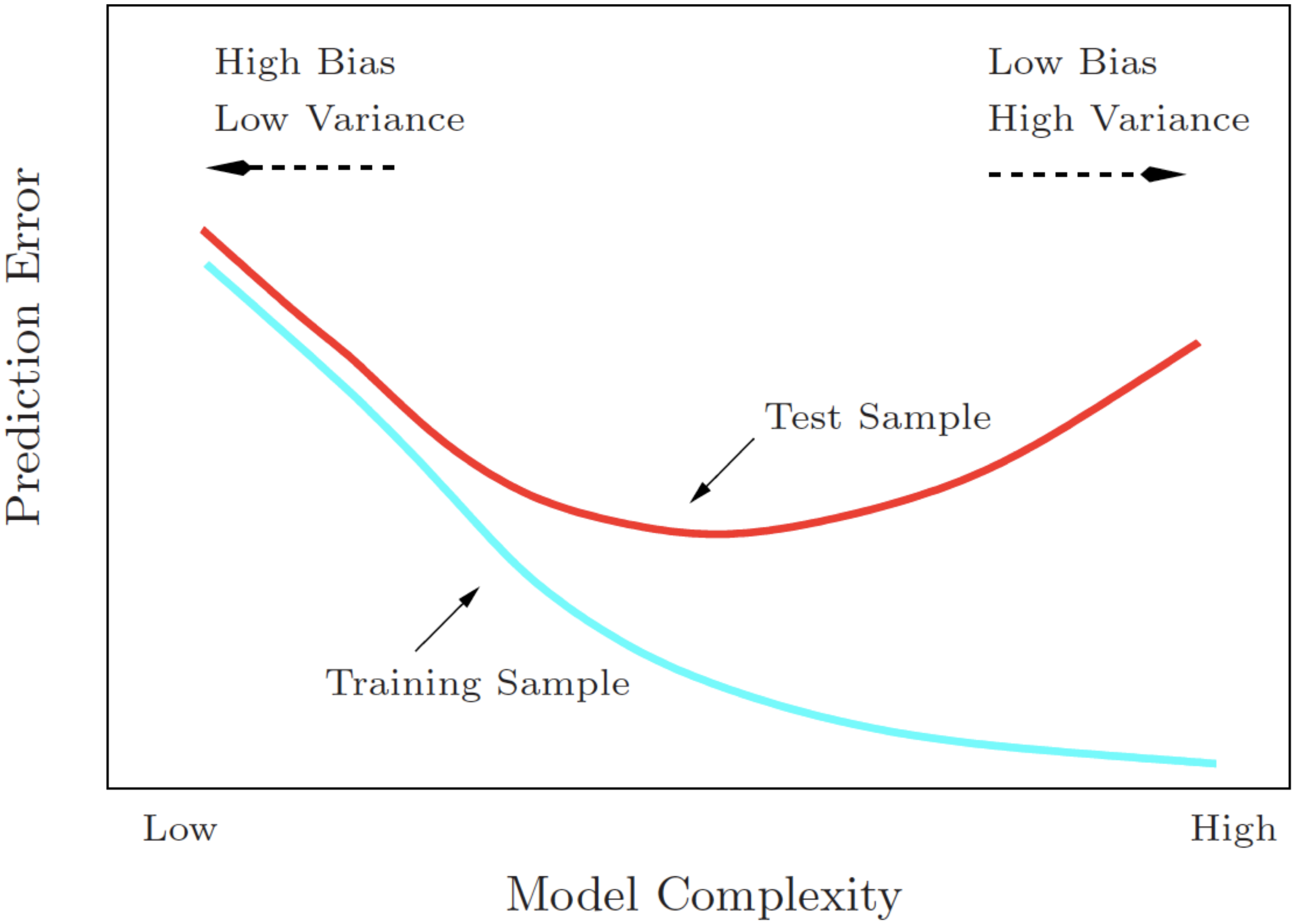


Source: Hastie, Tibshirani, Friedman, "Elements of Statistical Learning", 2011.

Rui Pedro Gonçalves de Brito

New Ways of Measuring and Dealing with Risk and Return in Portfolio Optimization

Tese de doutoramento em Economia, orientada pelo
Professor Doutor Helder Miguel Correia Virtuoso Sebastião e pelo Professor Doutor Pedro Manuel Cortesão Godinho,
apresentada à Faculdade de Economia da Universidade de Coimbra

julho de 2017

· U C ·

UNIVERSIDADE DE COIMBRA

• U C •

FEUC FACULDADE DE ECONOMIA
UNIVERSIDADE DE COIMBRA

Rui Pedro Gonçalves de Brito

New Ways of Measuring and Dealing with Risk and Return in Portfolio Optimization

Tese de doutoramento em Economia, apresentada à Faculdade de Economia da Universidade de Coimbra para a obtenção do grau de Doutor

Orientadores:
Prof. Doutor Helder Miguel Correia Virtuoso Sebastião
e
Prof. Doutor Pedro Manuel Cortesão Godinho

Coimbra, 2017

Dedicated to my father's memory and to my mother.

Acknowledgments

It is difficult to find the right words to express my deep gratitude to my supervisors Professor Helder Sebastião and Professor Pedro Godinho. Both were always available and had a great patience to help me to overcome all the difficulties encountered in the course of this work. I will be forever in debt with them for guiding me in this extraordinary experience. I will always remember our long energizing conversations where we had discussed not only this work but also intertemporal questions about science and life. Their commitment to excellence, their continuous support and motivation, were to me the key insights to always try to go a little further in all phases of the research work. An infinite thank you to both to contribute to the training of my mind to think.

I would like to express my eternal gratitude to Professor Nunes Vicente. During the Masters in Quantitative Methods in Finance, was under his guidance that I was introduced in the exciting research world of Financial Mathematics. I will never forget his scientific and personal teachings. As a researcher assistant at the Department of Mathematics of the University of Coimbra, it was a remarkable experience to work with him and had written my first scientific paper with his collaboration.

I am grateful to every teacher I met during all my studies, who have guided me and showed me the right tools to develop my ability to learn and to think. Particularly, I want to thank to all the professores that I met during this PhD Program and that have undoubtedly contributed for the development of my scientific knowledge.

I want to thank to all members of the Faculty of Economics, University of Coimbra and to the School of Economics and Management, University of Minho, for all the help and support in the accomplishment of this work.

I am grateful for the funding provided by the Portuguese National Funding Agency for Science, Research and Technology (FCT) under the scholarship SFRH/BD/94778/2013. Without this financial support would not be possible accomplish this work. I am also thankful to the Group of Monetary and Financial Studies (GEMF) and to the Faculty of Economics, University of Coimbra, as the hosting institutions, for provided me the right conditions for the completion of this work. In particular, I am very grateful to Professor Paulino Teixeira, the PhD Program Coordinator, for all his guidance, support and motivation.

A big thank you to all my PhD colleagues. The sharing of all the joys and anxieties with each one of you, have made this whole experience unforgettable.

For last and definitely not the least, I would not have produced this work without the unconditional support and motivation of my family. To my wife for her love that completes my days. To my sisters, my brothers-in-law, my sister-in-law, my parents-in-law and my wonderful nieces for enriching my life. To my parents, for all the amazing support and encouragement to follow my dreams. In particular, to the memory of my father who so many times heard about this work; thank you dad for everything you gave me and continue to give me.

Funding

Support was provided by FCT under the scholarship SFRH/BD/94778/2013.

"Whenever a theory appears to you as the only possible one, take this as a sign that you have neither understood the theory nor the problem which it was intended to solve."

—Karl Popper

Resumo

Motivados pelas limitações do modelo de otimização de média-variância, nesta tese propomos a construção de carteiras de investimento em cenários diferentes. Olhando de modo diferente para como lidar com o risco e o retorno na construção de carteiras de investimento, tentamos sugerir cenários mais realistas do que o cenário clássico.

Começamos por propor uma metodologia flexível para a construção de carteiras de investimento, utilizando um modelo de otimização biobjetivo de assimetria/semivariância. As soluções deste problema de otimização biobjetivo permitem ao investidor analisar o compromisso eficiente entre a assimetria e a semivariância. Esta metodologia é utilizada empiricamente em quatro conjuntos de dados, obtidos na coleção de *Fama/French*. A performance fora-da-amostra do modelo de assimetria/semivariância foi aferida escolhendo três carteiras de investimento pertencentes a cada fronteira de *Pareto* dentro-da-amostra e medindo a sua performance em termos do rácio de assimetria por semivariância, rácio de *Sharpe* e rácio de *Sortino*. Ambas as análises de performance dentro-da-amostra e fora-da-amostra foram realizadas utilizando três retornos alvo diferentes para os cálculos da semivariância. Os resultados mostram que as carteiras de investimento eficientes de assimetria/semivariância são consistentemente competitivas quando comparadas com diferentes carteiras de investimento de referência.

Posteriormente, estendemos o estudo do impacto da cardinalidade na performance das carteiras de investimento, do cenário tradicional de média-variância a cenários mais gerais que incluem momentos de ordem superior. Para cada cenário, nós propomos um modelo biobjetivo que permite ao investidor analisar explicitamente o compromisso eficiente entre a utilidade esperada e a cardinalidade. Aplicamos a metodologia proposta a dados relativos a títulos pertencentes ao Índice do Mercado de Ações Português (Índice *PSI* 20). Os resultados empíricos mostram que, dentro-da-amostra, em todos os cenários o equivalente certo e o rácio de *Sharpe* aumentam com o nível de cardinalidade. Os resultados também sugerem que não existem ganhos de performance, dentro-da-amostra, em termos de equivalente certo, quando se consideram momentos de ordem superior. Fora-da-amostra, a rotação da carteira de investimento aumenta até um certo nível de cardinalidade, decrescendo posteriormente. Para certos níveis de cardinalidade, existem ganhos em termos de equivalente certo e rácio de *Sharpe* fora-da-amostra, quando são consideradas a assimetria e a curtose. Confirmamos a robustez destes resultados num conjunto maior de dados relativos a títulos pertencentes ao Índice do Mercado de Ações da Zona Euro (Índice *EURO STOXX* 50).

Finalmente, dentro de um cenário de maximização da utilidade com aversão relativa ao risco constante (*CRRA*), sugerimos a construção de duas carteiras de investimento diferentes: uma carteira de investimento de baixa frequência e uma carteira de investimento de alta frequência. Para dez níveis de aversão ao risco diferentes, comparamos a performance de ambas a carteiras de investimento em termos de várias medidas fora-da-amostra. Utilizando dados de catorze ações do Índice do Mercado de Ações Francês (Índice *CAC* 40), concluímos que para todas as medidas de avaliação de performance consideradas o "combate" é sempre "ganho" pela carteira de investimento de alta frequência. Posteriormente, consideramos um cenário onde o investidor, com preferências *CRRA*, tem dois objetivos: a maximização da utilidade esperada e a minimização da iliquidez esperada da carteira de investimento. A utilidade-*CRRA* é medida utilizando a volatilidade realizável, a assimetria realizável e a curtose realizável da carteira de investimento, enquanto que a iliquidez da carteira de investimento é medida utilizando o bem conhecido rácio de iliquidez de *Amihud*. Assim, o investidor é capaz de realizar as suas escolhas diretamente no espaço bidimensional de utilidade esperada/liquidez (*EU/L*). Conduzimos uma análise empírica no mesmo conjunto de ações do Índice *CAC* 40. A robustez do modelo proposto é averiguada de acordo com a performance fora-da-amostra de diferentes carteiras de investimento *EU/L* em relação à carteira de investimento de variância mínima e à carteira de investimento em que todos os títulos têm ponderações idênticas. Para diferentes níveis de aversão ao risco, a carteiras de investimento *EU/L* são bastante competitivas e em vários casos consistentemente superam aquelas carteiras de referência, em termos de utilidade, liquidez e equivalente certo.

Classificação JEL: C44; C55; C58; C61; C63; C88; G11; G17

Palavras-chave: otimização de carteiras de investimento, maximização da utilidade, semivariância, momentos de ordem superior, cardinalidade, dados de alta frequência, momentos de ordem superior realizáveis, liquidez, performance fora-da-amostra, otimização multiobjetivo, otimização sem derivadas

Abstract

Motivated by the limitations of the mean-variance optimization model, in this thesis we propose to approach the portfolio selection problem with different frameworks. Looking differently at how to deal with risk and return in portfolio construction, we try to suggest more realistic frameworks than the classical one.

We begin by proposing a flexible methodology for portfolio choice, using a skewness/semivariance biobjective optimization model. The solutions of this biojective optimization problem allow the investor to analyze the efficient tradeoff between skewness and semivariance. This methodology is used empirically on four datasets, collected from the Fama/French data library. The out-of-sample performance of the skewness/semivariance model was assessed by choosing three portfolios belonging to each in-sample Pareto frontier and measuring their performance in terms of skewness per semivariance ratio, Sharpe ratio and Sortino ratio. Both the in-sample and the out-of-sample performance analyses were conducted using three different target returns for the semivariance computations. The results show that the efficient skewness/semivariance portfolios are consistently competitive when compared with several benchmark portfolios.

Then we extend the study of the cardinality impact on the portfolio performance, from the traditional mean-variance framework to more general frameworks that include higher moments. For each framework, we propose a biobjective model that allows the investor to explicitly analyze the efficient tradeoff between expected utility and cardinality. We applied the proposed methodology to data from the Portuguese Stock Market Index (PSI 20 Index). The empirical results show that, in-sample, the certainty equivalent and the Sharpe ratio increase with the cardinality level in all frameworks. The results also suggest that there are no performance gains, in-sample, in terms of certainty equivalent, when higher moments are considered. Out-of-sample, the turnover increases up to a certain cardinality level, then decreases. For certain cardinality levels, there are gains in terms of out-of-sample certainty equivalent and Sharpe ratio, when skewness and kurtosis are considered. We check the robustness of these results in a large dataset from the Eurozone Stock Market Index (EURO STOXX 50 Index).

Finally, within a CRRA-utility maximization framework, we suggest the construction of two different portfolios: a low and a high frequency portfolio. For ten different risk aversion levels, we compare the performance of both portfolios in terms of several out-of-sample measures. Using data on fourteen stocks of the French Stock Market Index (CAC 40 Index), we conclude that the "fight" is always "won" by the high frequency portfolio for all the considered performance evaluation measures. Then, we consider a framework where the investor, with CRRA preferences, has two objectives: the maximization of the expected utility and the minimization of the portfolio expected illiquidity. The CRRA-utility is measured using the portfolio realized volatility, realized skewness and realized kurtosis, while the portfolio illiquidity is measured using the well-known Amihud illiquidity ratio. Therefore, the investor is able to make her choices directly in the expected utility/liquidity (EU/L) bidimensional space. We conduct an empirical analysis in the same set of stocks of the CAC 40 Index. The robustness of the proposed model is analyzed taking into account the out-of-sample performance of different EU/L portfolios relative to the minimum variance and equally weighted portfolios. For different risk aversion levels, the EU/L portfolios are quite competitive and in several cases consistently outperform those benchmarks, in terms of utility, liquidity and certainty equivalent.

JEL Classification: C44; C55; C58; C61; C63; C88; G11; G17

Keywords: portfolio optimization, utility maximization, semivariance, higher moments, cardinality, high frequency data, realized higher moments, liquidity, out-of-sample performance, multiobjective optimization, derivative-free optimization

Acronyms

BSM - **B**lack-**S**holes **M**odel

CAPM - **C**apital **A**sset **P**ricing **M**odel

CML - **C**apital **M**arket **L**ine

CRRA - **C**onstant **R**elative **R**isk **A**version

CAC - **C**otation **A**ssistée en **C**ontinu

DMS - **D**irect **M**ulti**S**earch

EUROFIDAI - **Euro**pean **Fi**nancial **Da**ta **I**nstitute

HARA - **H**yperbolic **A**bsolute **R**isk **A**version

IS - **I**n-**S**ample

MV - **M**ean-**V**ariance

MVS - **M**ean-**V**ariance-**S**kewness

MVSK - **M**ean-**V**ariance-**S**kewness-**K**urtosis

MPT - **M**odern **P**ortfolio **T**heory

MOO - **M**ulti**O**bjective **O**ptimisation

OOS - **O**ut-**O**f-**S**ample

PhD - **Ph**ilosophiæ **D**octor

PSI - **P**ortuguese **S**tock **I**ndex

QP - **Q**uadratic **P**rogramming

List of Figures

List of Tables

Contents

Chapter 1

Introduction

The importance of randomness in our lives is often underestimated. One day, at the University of Chicago, while waiting for a meeting with his advisor, to discuss his thesis topic, a PhD student had a casual conversation with a stockbroker. During the conversation, the stockbroker suggested to the PhD student that he should consider to study the application of mathematical methods to the stock market in his thesis. As a result from following this specific advice, the Modern Portfolio Theory (MPT) emerged, associated to the famous paper of Markowitz (1952).

In that seminal work, Markowitz has pointed out that a rule for portfolio choice in which the investor only seeks to maximize the expected return is a poor one, since no diversification is carried out. Following the prevailing theory at the time (systematized by Williams, 1938), the investor could compute the expected return of each security as a discounted return (making use of the discounted cash flow theory). Nevertheless, if the investor was only concerned with the maximization of the portfolio's expected return (regardless of how it is computed), the portfolio would always be formed by only one security, or by a combination of securities with equal expected returns. Motivated by this fact, Markowitz proposed that, in addition to the maximization of the portfolio's expected return, the investor should seek to minimize the risk of the portfolio (measured by the portfolio's variance), thus promoting security diversification. This mean-variance (MV) rule is mathematically given in the form of a MV optimization model. Typically, the MV optimization model is formulated as a single objective optimization problem, where the investor minimizes the portfolio's variance for a given level of expected return, over the set of feasible portfolios. By varying the level of expected return, the set of nondominated portfolios can be identified in the MV bidimensional space. This set of nondominated portfolios defines the efficient MV frontier.

According to the MPT, a rational investor should choose a portfolio on the efficient MV frontier; therefore, this is a normative theory. According to Fabozzi et al. (2002, p.7), a normative theory "is one that describes a standard or norm of behavior that investors should pursue in constructing a portfolio, in contrast to a theory that is actually followed." Besides that, the theoretical consistency of the MPT with the rational axioms of choice, proposed by Von Neumann and Morgenstern (1953), is dependent on a quadratic utility function for the investor's preferences or on the returns following a Gaussian distribution.

As recently explained by Markowitz (2014), a quadratic utility function for the investor's preferences or the returns following a Gaussian distribution, are sufficient conditions for the MV analysis. Still, these are not necessary conditions, since an appropriate choice on the efficient MV frontier can correspond to an approximation of the investor's utility maximization for a variety of utility functions. However, if the investor follows non-quadratic preferences and the distribution of the returns does not correspond to a Gaussian distribution, a more realistic approach is required to the portfolio choice.

MPT has been criticized both in theoretical and practical grounds. Over the years, several studies have showed that the returns' distribution does not follow a Gaussian distribution (see, e.g., Mandelbrot, 1963; Fama, 1965; Beedles, 1979; Campbell et al., 1997). Furthermore, the portfolios constructed according to the MV optimization model tend to be very unstable, due to the ill-conditioning of the covariance matrix, hence exhibiting a poor out-of-sample (OOS) performance. The estimation error present in the estimates of the expected returns, the variances and the covariances is too big to be overlooked (Chopra and Ziemba, 1993; Broadie, 1993). In fact, a provocative article by DeMiguel et al. (2009b) presents compelling evidence on this estimation error problem. DeMiguel et al. (2009b) applied, on seven datasets, fourteen portfolio choice models and showed that none was able to consistently beat (in terms of different OOS performance measures) the equally weighted portfolio. Moreover, DeMiguel et al. (2009b) showed that for a portfolio of 50 securities, based on the U.S. stock market, it is necessary an estimation window of approximately 6000 months (500 years!) for the MV optimization model to outperform the equally weighted strategy. Even models designed to deal with estimation error need unreasonable large estimation windows to outperform the equally weighted strategy (DeMiguel et al., 2009b).

Motivated by the limitations of the MPT, several improvements to the MV optimization model have arisen in the literature. Such improvements have been achieved through a better estimation of the covariance matrix, a better estimation of the vector of expected returns and using more real constraints. Several studies have contributed to a better estimation of the covariance matrix by using high-frequency data (see, e.g., Merton, 1980; Schwert, 1989; Hsieh, 1991; Andersen et al., 2001b), or, by applying factor models (see, e.g., Chang et al., 1999; Ikeda and Kubokawa, 2016), shrinkage estimators (see, e.g., Ledoit and Wolf, 2004a;b), robust optimization (see, e.g., Goldfarb and Iyengar, 2003; Tütüncü and Koenig, 2004; Zhu and Fukushima, 2009; Gotoh et al., 2013; Fernandes et al., 2016), and robust estimation (see, e.g., DeMiguel and Nogales, 2009; Huo et al., 2012). Other studies contributed to a better estimation of the expected return by applying Bayesian estimation (see, e.g., Klein and Bawa, 1976; Jorion, 1986; Pastor, 2000; Pastor and Stambaugh, 2000), factor models (see, e.g., Fama and French, 1992; 1993; 1996; Carhart, 1997; Fama and French, 2015; DeMiguel et al., 2017), robust optimization (see, e.g., Goldfarb and Iyengar, 2003), using option-implied information (see, e.g., DeMiguel et al., 2013), and exploiting stock return serial dependence (see, e.g., DeMiguel et al., 2014). Finally, other studies have addressed the use of better constraints, such as the inclusion of non-short selling constraints (see, e.g., Jagannathan and Ma, 2003), the introduction of constraints to explore information about the cross-sectional characteristics of securities (see, e.g., Brandt et al., 2009), the use of performance-based regularization approaches (see, e.g., Ban et al., 2016), and the introduction of norm constraints (see, e.g., Brodie et al., 2009; DeMiguel et al., 2009a; Brito and Vicente, 2014).

In this thesis we follow a holistic approach suggesting and analyzing the construction of data-driven portfolios beyond the MV optimization framework. Our main focus is to contribute to the existing literature with more comprehensive and realistic models, while using a minimum number of assumptions. Although we try to suggest models that intend to be closer to reality, we are well aware of the limitations of modelling complex systems, particularly financial systems. In general modelling and, particularly, in financial modelling, it is extremely important to always bear in mind the eloquent description of what a model is, given in Derman (2011, p.112): "There is a gap between the model and the object of its focus. The model is not the object, though we may wish it were. A model is a metaphor of limited applicability, not the thing itself." After the financial crisis of 2008, many criticisms have arisen to the type of models used in the financial field. One of the best summary reflections about this issue was written by Emanuel Derman and Paul Willmott, in 2009. In the document coined as "The Financial Modelers' Manifesto" (see Appendix A), the authors describe the precautions to be undertaken in the construction and application of financial models. We think it is truly important to put in practice the principles set out in "The Modelers' Hippocratic Oath" (see Appendix A), presented in the referred manifesto.

This thesis results from the compilation of four articles (one published, two accepted for publication and one submitted). Since the articles were written independently, here we articulate them in order to provide the reader with a coherent sequence of the theoretical aspects of each proposed model, as well as the presentation of the corresponding empirical findings. We seek to follow a logical sequence, eliminating any redundancies between those articles. Accordingly, the organization of this thesis is as follows.

Chapter 2 presents the MPT and discusses it limitations. Then, in order to overcome some of the limitations of the MPT, we take one step forward and go beyond the MV optimization framework, introducing the higher moments' analysis in portfolio choice. This more realistic approach is essential for the theoretical understanding of the models proposed in the subsequent chapters. In Chapter 2 we define three important benchmark portfolios, documented in the related literature as hard to beat strategies in terms of OOS performance. These benchmark portfolios are used for performance comparisons throughout the thesis.

Observing the investor's aversion for losses (regarding some reference point), i.e., aversion for semivariance, and that the investor has preference to face higher probability for extreme gains and limited losses, i.e., preference for positive skewness, in Chapter 3 we propose a flexible skewness/semivariance biobjective model. To the best of our knowledge, for the first time in the literature, we overcome the endogeneity problem of the cosemivariance matrix through a derivative-free algorithm. Independently of the target return used in the semivariance computation, with the proposed model the investor is able to directly analyze the efficient tradeoff between the portfolio skewness and the portfolio semivariance. As an application example, an empirical exercise is conducted using four datasets from the Fama/French data library and three different target returns for the semivariance computation. [This chapter partially corresponds to the published article Brito et al. (2016)]

Chapter 4 extends the analysis of the cardinality impact on the portfolio performance from the classical MV framework to frameworks where higher moments are considered (namely skewness and kurtosis). Furthermore, we analyze the performance gains obtained with the inclusion of skewness and kurtosis, at different cardinality levels. For the in-sample (IS) analysis we use, as performance measures, the certainty equivalent and the Sharpe ratio. The OOS performance is assessed through the certainty equivalent, the Sharpe ratio, the turnover and the Sharpe ratio of returns net of transaction costs. We present empirical results for the PSI 20 Index and for the EURO STOXX 50 Index. [This chapter partially corresponds to the accepted for publication article Brito et al. (2017c)]

In Chapter 5, motivated by the strong growth of the available high frequency data, we assess the benefits of using such data in a portfolio choice framework and we propose a new approach for portfolio choice by means of an expected constant relative risk aversion (CRRA)-utility/liquidity model. All the empirical analysis is done on a set of fourteen stocks from the French Stock Market Index (CAC 40 Index). This data was provided by the European Financial Data Institute (EUROFIDAI). [This chapter partially corresponds to the working paper Brito et al. (2017a) (submitted for publication) and to the article Brito et al. (2017b), accepted for publication]

Finally, Chapter 6 presents the main conclusions and limitations of this thesis.

Chapter 2

MPT, limitations and going further

In this chapter we begin by presenting the MPT (Markowitz, 1952; 1959), operationally defined by a MV optimization model (Section 2.1). Then we discuss the two main limitations of the MV optimization model (Section 2.2): the non-Gaussian distribution of returns and the poor OOS performance of the MV portfolios. Finally, this chapter ends by introducing the higher moments' analysis in portfolio choice (Section 2.3), which allows us to go beyond the MPT in the subsequent chapters.

2.1 The MV optimization model

This presentation of the MV optimization model follows the notation used in Brito and Vicente (2014) and Brito et al. (2016). Here we designate a portfolio as a basket of financial assets or securities. Portfolios can include, for example, stocks, bonds, currencies or even derivatives[1]. Suppose that an investor has a universe of N securities where she can allocate her wealth, W. The returns of each security i (with $i = 1, \ldots, N$), at time t (with $t = 1, \ldots, T$), are denoted by $r_{i,t}$. Based on historical data, a portfolio return matrix, RP, is defined as

$$RP = \begin{bmatrix} r_{1,1} & r_{2,1} & \cdots & r_{N,1} \\ r_{1,2} & r_{2,2} & \cdots & r_{N,2} \\ \cdots & \cdots & \cdots & \cdots \\ r_{1,T-1} & r_{2,T-1} & \cdots & r_{N,T-1} \\ r_{1,T} & r_{2,T} & \cdots & r_{N,T} \end{bmatrix}, \tag{2.1}$$

where $T + 1$ corresponds to the total number of price observations. The expected

[1]This thesis only deals empirically with stocks and portfolios of stocks; however the suggested methodologies can be applied to any kind of financial assets.

return of each security i, with $i = 1, \ldots, N$, at time t is denoted by $\mu_{i,t} = E_t(r_{i,t+1})$. And $\mu_t = [\mu_{1,t}\ \mu_{2,t}\ \ldots\ \mu_{N,t}]^\top$, of dimension $N \times 1$, is the vector of the expected returns. Given the universe of N securities, at each time t there is a set of weights, $w_{i,t}$, $i = 1, \ldots, N$, representing the proportional amounts of the total wealth, W, invested in the corresponding securities. Thus, a portfolio can be defined by an $N \times 1$ vector $w_t = [w_{1,t}\ w_{2,t}\ \ldots\ w_{N,t}]^\top$ of weights, which have to satisfy the constraint

$$1_N^\top w_t = \sum_{i=1}^{N} w_{i,t} = 1, \tag{2.2}$$

where 1_N is the $N \times 1$ vector of ones.

Denoting the portfolio return at time t as $r_{p,t}$, the portfolio expected return, $m_t(r_{p,t+1})$, can be written as

$$\begin{aligned} m_t(r_{p,t+1}) &= E_t(r_{p,t+1}) = E(w_{1,t}r_{1,t+1} + \cdots + w_{N,t}r_{N,t+1}) \\ &= w_{1,t}\mu_{1,t} + \cdots + w_{N,t}\mu_{N,t} = \mu_t^\top w_t. \end{aligned} \tag{2.3}$$

In turn, the variance of the portfolio return, $v_t(r_{p,t+1})$, is given by

$$\begin{aligned} v_t(r_{p,t+1}) &= E_t\left[r_{p,t+1} - E_t(r_{p,t+1})\right]^2 \\ &= E_t\left[\sum_{i=1}^{N} w_{i,t}r_{i,t+1} - E_t\left(\sum_{i=1}^{N} w_{i,t}r_{i,t+1}\right)\right]^2. \end{aligned} \tag{2.4}$$

Thereby,

$$v_t(r_{p,t+1}) = \sum_{i=1}^{N}\sum_{j=1}^{N} E_t[(r_{i,t+1} - \mu_{i,t})(r_{j,t+1} - \mu_{j,t})]w_{i,t}w_{j,t}. \tag{2.5}$$

Representing each entry i, j (with $i, j = 1, \ldots, N$) of the covariance matrix Σ_t by

$$\sigma_{ij,t} = E_t[(r_{i,t+1} - \mu_{i,t})(r_{j,t+1} - \mu_{j,t})], \tag{2.6}$$

we have

$$v_t(r_{p,t+1}) = w_t^\top \Sigma_t w_t, \tag{2.7}$$

where $\forall_{w_t \in \mathbb{R}^N}\ w_t^\top \Sigma_t w_t \geq 0$, since the variance is always nonnegative. Thus, Σ_t is a symmetric and positive semi-definite matrix. Typically Σ_t is assumed to be positive definite. Otherwise, 0 would be an eigenvalue of Σ_t by definition. Thereby

$$\exists_{w_t \in \mathbb{R}^N (w_t \neq 0_N)} : \Sigma_t w_t \;=\; 0_N, \tag{2.8}$$

with 0_N representing the vector of dimension $N \times 1$ where each entry is equal to 0. This would lead to the existence of redundant securities (Cornnuejols and Tütüncü, 2007).

Accordingly, the portfolio standard deviation, $\sigma_t(r_{p,t+1})$, is given by

$$\sigma_t(r_{p,t+1}) \;=\; \sqrt{w_t^\top \Sigma_t w_t}. \tag{2.9}$$

The MPT (Markowitz, 1952; 1959) is based on the formulation of a MV optimization model. The solution of that model is the portfolio of minimum variance for an expected return not below a certain target value r. Therefore, the aim is to minimize the risk for a given level of return. A standard formulation of the MV optimization model is given as a convex quadratic programming (QP) problem

$$\begin{aligned} \min_{w_t \in \mathbb{R}^N} \quad & w_t^\top \Sigma_t w_t \\ \text{subject to} \quad & \mu_t^\top w_t \geq r, \\ & 1_N^\top w_t = 1, \\ & w_t \geq 0_N. \end{aligned} \tag{2.10}$$

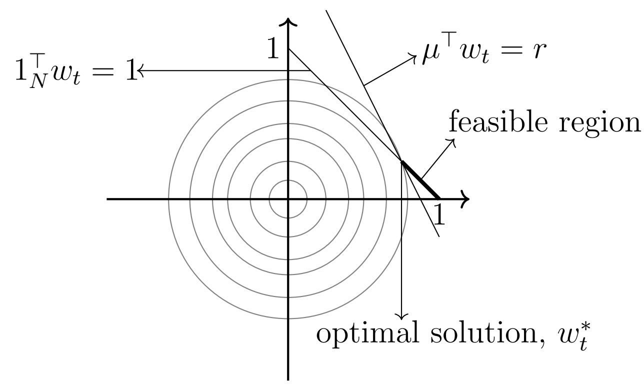
Problem (2.10) includes the usual additional constraint $w_t \geq 0_N$, which excludes from the feasible region the possibility of short selling.

Let w_t^* be an optimal solution of Problem (2.10). The Karush-Kuhn-Tucker (KKT) conditions for Problem (2.10) can be written as

$$
\begin{aligned}
2\Sigma_t w_t^* - \lambda_1 \mu_t - \lambda_2 1_N - \nu &= 0_N, \\
\nu \geq 0_N, \quad \nu^\top w_t^* &= 0, \\
\lambda_1 \geq 0, \quad \lambda_1(\mu^\top w_t^* - r) &= 0, \\
\mu_t^\top w^* \geq r, \quad 1_N^\top w_t^* = 1, \quad w_t^* &\geq 0_N,
\end{aligned} \tag{2.11}
$$

for some $\lambda_1,\ \lambda_2 \in \mathbb{R}$ and $\nu \in \mathbb{R}^N$. Since Problem (2.10) is a strictly convex problem (the objective function is strictly convex and the feasible region is convex), if there is a solution it is unique. The feasible region is bounded and closed, thus by the Weierstrass theorem, if the feasible region is not empty, there is always a solution. Hence, if the feasible region of Problem (2.10) is not empty then the problem has a unique solution, i.e., the KKT 4-tuple $(w_t^*, \lambda_1, \lambda_2, \nu)$ can always be found (see Figure 2.1, for a graphical illustration of how to find the optimal solution in a $N = 2$ case).

Figure 2.1: A graphical solution of the MV optimization model (an example when $N = 2$)



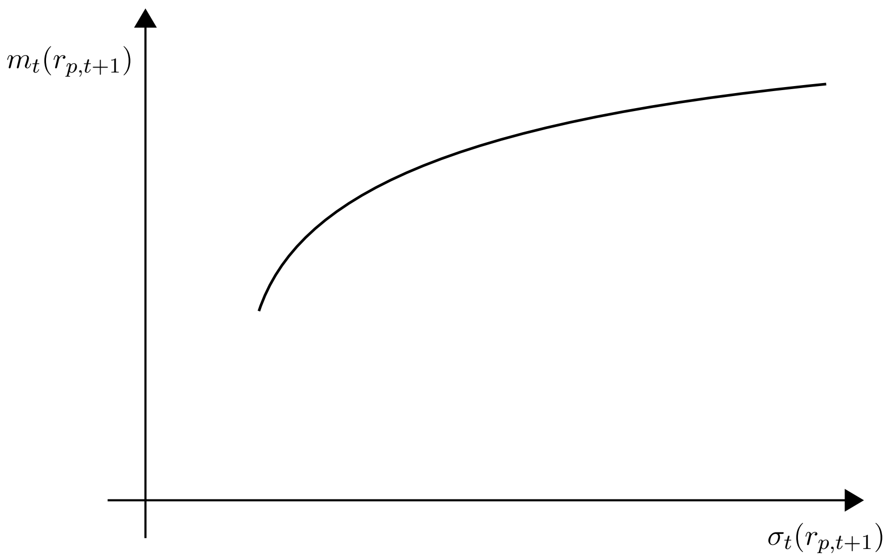
This figure illustrates the process of finding the optimal solution for Problem (2.10), when $N = 2$. The figure represents the contour lines of the objective function ($w_t^\top \Sigma_t w_t$), the feasible region defined by the three constraints ($\mu_t^\top w_t \geq r \ \wedge \ 1_N^\top w_t = 1 \ \wedge \ w_t \geq 0_N$), and the optimal solution (w_t^*) of Problem (2.10).

The MV optimization model can be reformulated as a biobjective problem, which consists of simultaneously minimizing the portfolio variance and maximizing the portfolio expected return

$$
\begin{aligned}
\min_{w_t \in \mathbb{R}^N} \quad & v_t(r_{p,t+1}) \;=\; w_t^\top \Sigma_t w_t \\
\max_{w_t \in \mathbb{R}^N} \quad & m_t(r_{p,t+1}) \;=\; \mu_t^\top w_t \\
\text{subject to} \quad & w_t \in P_t,
\end{aligned}
\tag{2.12}
$$

where $P_t = \left\{ w_t \in \mathbb{R}^N : 1_N^\top w_t = 1 \ \wedge \ w_t \geq 0_N \right\}$ denotes the feasible region. It is easy to prove that a solution of Problem (2.10) is nondominated, efficient or Pareto optimal for Problem (2.12) (see Appendix B). Efficient portfolios are thus the ones that have the minimum variance for at least a certain expected return, or, alternatively, those that have the maximal expected return up to a certain variance. The efficient Pareto frontier is typically represented as a bidimensional curve in the expected return-standard deviation space (see Figure 2.2).

Figure 2.2: Efficient Pareto frontier for the MV optimization model



This figure illustrates the solution of the biobjective Problem (2.12). The vertical axis corresponds to the expected return ($m_t(r_{p,t+1})$) and the horizontal axis corresponds to the standard deviation ($\sigma_t(r_{p,t+1})$).

Following Cornnuejols and Tütüncü (2007), one easy way to obtain the efficient Pareto frontier (Figure 2.2) is to proceed as follows:

- Let r_{max} be the maximum return for an admissible portfolio;
- Find the minimum variance portfolio, w_t^{min}, solution of the strictly convex problem

$$
\begin{aligned}
\min_{w_t \in \mathbb{R}^N} \quad & w_t^\top \Sigma_t w_t \\
\text{subject to} \quad & w_t \in P_t;
\end{aligned}
$$

- Define $r_{min} = \mu_t^\top w_t^{min}$;
- Define the function

$$\sigma : \; [r_{min}, \; r_{max}] \to \mathbb{R} \quad | \quad \sigma(r) = \sqrt{(w_t^r)^\top \Sigma w_t^r},$$

  where w_t^r is the optimal solution of Problem (2.10);
- The efficient Pareto frontier corresponds to the plot of each pair $(\sigma(r), r)$, with $r \in [r_{min}, \; r_{max}]$.

Once the efficient Pareto frontier is determined, the investor can choose, along the frontier, the portfolio that best fits her subjective preferences. The particular choice will depend on the weight that the investor gives to the expected return and to the portfolio variance (MV tradeoff).

In the efficient Pareto frontier there is a portfolio that deserves a special attention. This portfolio (the *ms* portfolio) results from the maximization of the risk premium per unit of risk; that is, the portfolio that maximizes the ratio between the reward and the variability of the investment (reward-to-variability ratio). The *ms* portfolio is obtained by maximizing the so-called Sharpe ratio

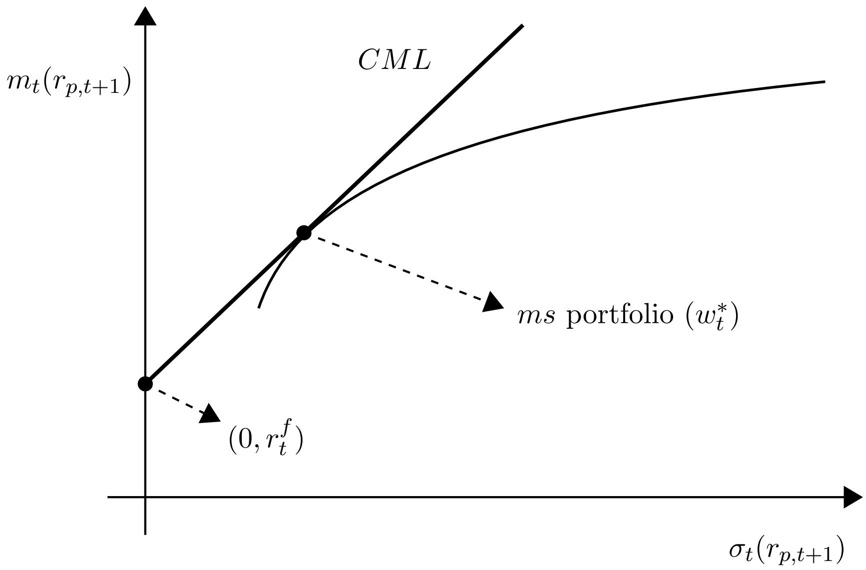
$$
\begin{aligned}
\max_{w_t \in \mathbb{R}^N} \quad SR \; & = \; \frac{\mu_t^\top w_t - r_t^f}{\sqrt{w_t^\top \Sigma_t w_t}} \\
\text{subject to} \quad & w_t \in P_t,
\end{aligned}
\tag{2.13}
$$

where r_t^f is the risk-free rate. Figure 2.3, illustrates how to find the *ms* portfolio geometrically. This portfolio will be used as a benchmark portfolio in Chapter 3.

2.2 Limitations of the MV optimization model

The MV optimization model (Markowitz, 1952; 1959), the Capital Asset Pricing model (CAPM) (Sharpe, 1964) and the Black-Sholes model (BSM) (Black and Scholes, 1973), are among the most prominent models of Modern Finance. The BSM, inspired in the works of Bachelier (1900) and Osborne (1959), assumes that the securities' returns

Figure 2.3: Maximum Sharpe ratio portfolio



This figure illustrates how to find the ms portfolio (also known as the market portfolio), solution of the nonlinear Problem (2.13), along the efficient Pareto frontier. The ms portfolio corresponds to the tangency point between the capital market line (CML) - the line that is tangent to the Pareto frontier and includes the point $(0, r_t^f)$ *- and the efficient Pareto frontier.*

follow a Brownian motion. As explained in Mandelbrot and Hudson (2004), this assumption implies that: 1) each price change is independent from the last price change (more formally, the returns follow a random walk); 2) the generating process of the price changes is always the same over time (in other words, the returns are stationary); and 3) returns follow a Gaussian distribution. In the next five paragraphs we will discuss how these three assumptions are not observed in practice. Our main focus will be on the evidence against the third assumption (the assumption of a Gaussian distribution for the securities returns). If we do not assume that the investor has quadratic preferences, the theoretical consistency of the MV optimization model requires the assumption that the returns' distribution is Gaussian. However, as we will discuss, this is not a realist assumption, thus becoming a major limitation of the MV optimization model. This weakness is also shared by the CAPM, since it is an equilibrium model based on the MV framework.

Several studies have revealed the nonconformity of the first assumption (independence of the price changes) with reality (see, e.g., Jegadeesh, 1990; Ding et al., 1993; Sadique and Silvapulle, 2001; Lewellen, 2002; DeMiguel et al., 2014; Huang et al., 2015). For instance, some authors provide empirical evidence on the existence of a short-term momentum pattern in different markets (see, e.g., Jegadeesh and Titman, 1993; Chang et al., 1996; 2000a).

Economic reasoning cannot support that the nature (the generator) of the securities returns is always the same independently of the period (at least in the long run). It is quite a strong assumption to think that the behavior (nature) of the stocks returns of the Standard & Poor's 500 Stock Index is the same for different periods, say 1926-1940 and 1960-1980, for example. As explained in Hsu (1984, p.13), "It is clear to many economists that the stock market during the late 1920s and 1930s was quite different from that after 1940 in regard to market institutions, regulations, margin requirements, and the efficiency of the market allocating funds for productive purposes (other than pure speculations)." This reasoning can be generalized to other periods and other markets. In fact, financial returns have an evolutionary (nonstationary) nature in the long run, which is probably related to regime switching (Hsu, 1984). Locally, the stationary assumption holds with good results (see, e.g., Stărică and Granger, 2005, for further details), but in the long run such assumption is shaky.

Now let us look at the third assumption, which states that the returns follow a Gaussian distribution. A random variable X follows a Gaussian distribution, with mean μ and variance σ^2, if the probability of X taking values in the interval $[a, b] \in \mathbb{R} \times \mathbb{R}$ is given by

$$Prob(a < X \leq b) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_a^b \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}. \tag{2.14}$$

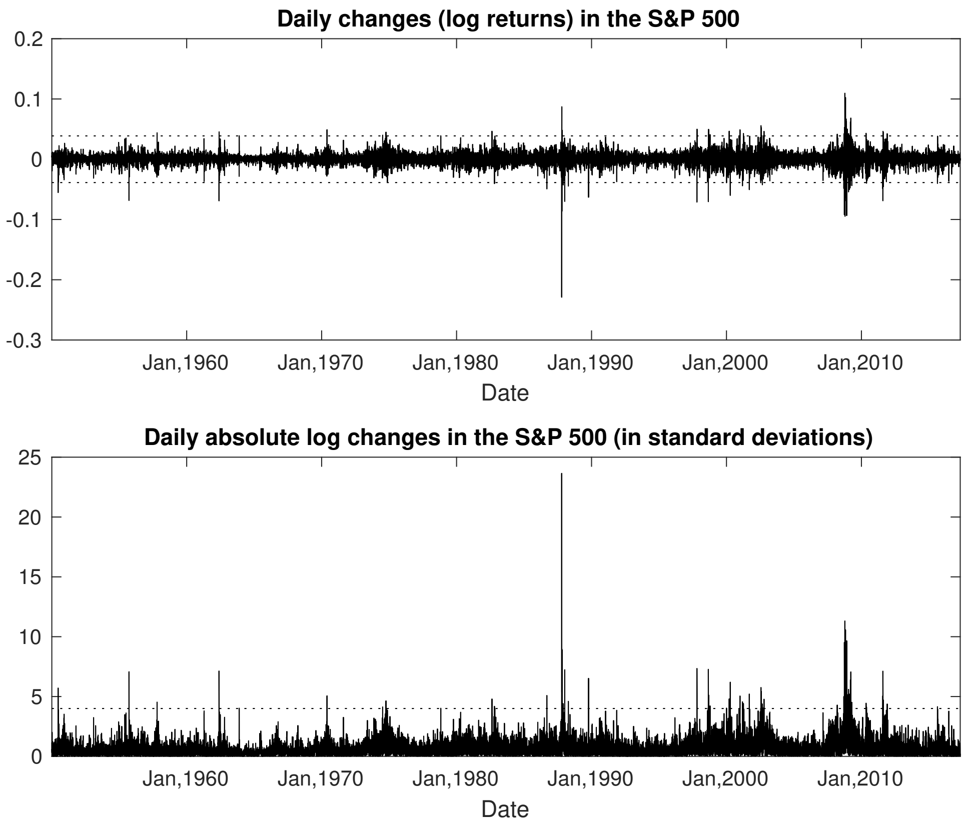
An important feature of the Gaussian distribution is that the probability of a deviation declines exponentially as one moves away from the mean. This translates into the so-called 68-95-99.7 rule (computed according Equation (2.14)): 68.27%, 95.45% and 99.73% of the observations are within the interval $[m - z\sigma,\ m + z\sigma]$, with $z = 1, 2$ and 3, respectively. Thus, events corresponding to 4σ deviations from the mean should be extremely rare (the probability of such occurrences is equal to 6.33×10^{-5}).

A simple exercise provides evidence against this Gaussian assumption. We collect historical daily data on the S&P 500 Stock Market Index from January 1950 to May 2017[2]. Figure 2.4 plots the daily changes (log returns) and the daily absolute changes in standard deviations. Clearly we do not observe the 68-95-99.7 rule. Two events stand out: the crash of October 1987 (an event that corresponds to 23σ!) and the subprime crisis in October 2008 (an event that corresponds to 11σ!). Notice that under the Gaussian assumption, the odds of an event such the crash of October 1987 would be 1 in 10^{117}. Just compare the magnitude of this number with the age of the

[2]At the current date, July 2017, this data is publicly available, for example, at Yahoo Finance (`https://finance.yahoo.com/`).

universe ($\approx 13.7 \times 10^9$) or even with the number of particles in the universe ($\approx 10^{80}$)! Furthermore, during the 16941 days present in Figure 2.4 there are 97 days where price changes exceed the 4σ threshold. Thus, large deviations occur far more frequently in financial markets than the Gaussian distribution would predict.

Figure 2.4: Daily changes in the S&P 500



This figure shows the daily returns and the absolute daily returns in standard deviations of the S&P 500. The sample covers the period since January 3, 1950 to May 2, 2017. The dotted lines, in the first plot, defines the interval $[-4\sigma, 4\sigma]$. *In the second plot, the dotted line, represents a standard deviation of* 4σ.

In fact, there is ample empirical evidence showing that the securities returns' distribution is fat-tailed and skewed, therefore being more appropriately described by a Paretian or Lévy stable (with the stable parameter $\alpha \in (1, 2)$) distribution (Mandelbrot, 1963; 1967; Fama, 1965; Rachev and Mittnik, 2000). A stable distribution has the suitable (for modeling the securities returns' distribution) property of being invariant under addition, i.e., the normed sum of random variables that follow a stable distribution will tend towards a stable distribution as the number of variables increases (this corresponds to the central limit theorem without the assumption of finite variance). The Gaussian distribution is a particular case of a stable distribution, corresponding to the case when the stable parameter, α, is equal to 2. When $1 < \alpha < 2$, the distribution has undefined variance. We need to point out, that although there is a broad consensus that a Paretian distribution fits better to the financial data than a Gaussian distribution, the estimation of α is itself subject to estimation error, thus conferring model instability.

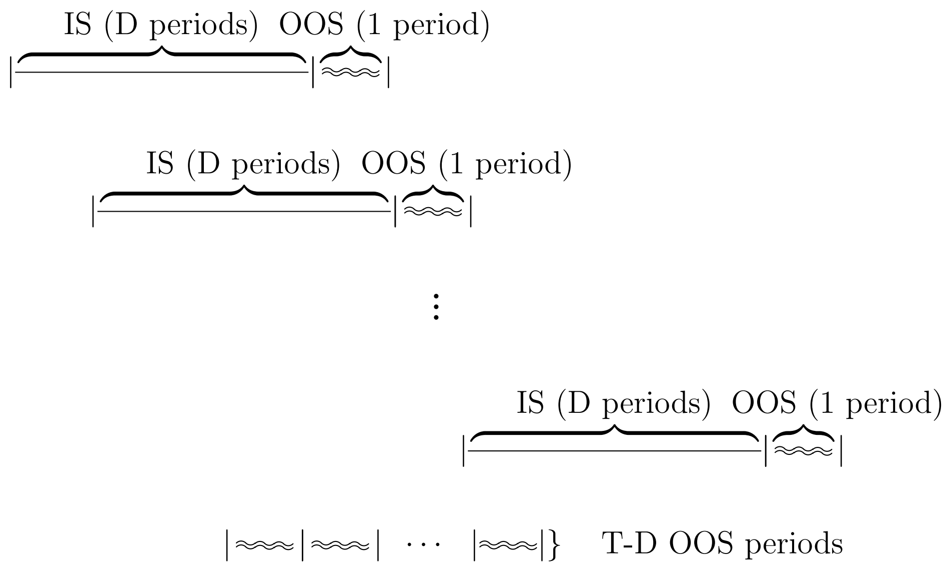
This important fat-tailed stylized fact of financial time series was also analyzed/explored in important papers associated to ARCH type (see, e.g., Engle, 1982; Bollerslev, 1986; Nelson, 1991; Zakoian, 1994; Tavares et al., 2008) and stochastic volatility models (see, e.g., Taylor, 1986; Bai et al., 2003). Even in the specific case of the Portuguese Stock Market Index (PSI 20 Index), we can find studies that analyze this issue (see, e.g., Curto et al., 2003; Rege et al., 2013).

Although the empirical evidence is quite compelling, models are still being built and practitioners are still operating, under the Gaussian distribution assumption. A relevant example was the edge fund Long Term Capital Management (LTCM), founded in 1994, which had among its principals Robert C. Merton and Myron S. Scholes (both Nobel laureates in 1997). The fund took its investment decisions upon risk models assuming a Gaussian framework. Then, the Asian financial crisis (in 1997) and the Russian financial crisis (in 1998) simply crashed LTCM. By the end of August 1998, the LTCM had lost $\$1,710$ millions in a month, which corresponds to a 8.3σ event (Jorion, 2000). Under the Gaussian assumption this event would occur, on average, once in $40,000$ times the age of the universe! As shown by Jorion (2000), if the risk models used by LTCM, instead of assuming a Gaussian distribution, assumed a fat-tailed t-distribution, the odds of such event would be a realistic number of 1 in 8 years (Jorion, 2000).

As we discussed previously, the assumption that returns follow a Gaussian distribution is violated in reality and it constitutes a major limitation of the MV optimization model. In response to this limitation, Markowitz (2014) argues that normal distributions or quadratic utility functions are sufficient but not necessary conditions for the optimality of MV portfolios. A careful choice on the MV efficient Pareto frontier can approximately maximize the expected utility for a variety of utility functions. But, clearly, if these hypotheses do not hold, another type of analysis is required.

Furthermore, the MV optimization model has shown to have a poor OOS performance. The most important test to the validity of a portfolio choice model rests on the portfolios' OOS performance, commonly evaluated through a rolling window approach (see Figure 2.5). The MV optimization model peremptorily fails this test. There is a vast literature showing that the portfolios constructed according the MV optimization model are unstable and exhibits a poor OOS performance (see, e.g., Michaud, 1989; Chopra and Ziemba, 1993; Broadie, 1993; DeMiguel et al., 2009b). As explained by Michaud (1989), the MV optimization model can be seen as a problem in two stages: Firstly, the input parameters (expected returns, variances and covariances) are estimated, typically, using historical data, and then the optimization routine is performed. Thus, the estimation error present in the first stage is "optimized" in the second stage, leading to the overestimation of some securities' weights and to the underestimation of others.

Figure 2.5: Rolling window approach



This figure illustrates a rolling window approach. Following DeMiguel et al. (2009b), the idea is to fix a certain estimation window length, say with D observations. For each one-period ahead, starting from $t = D+1$, the previous D IS observations are used to estimate the input parameters of a certain model and the portfolio is determined. Then the portfolio is held fixed and its OOS return is observed in t. The estimation window is then moved forward one period and the portfolio return is observed in the next OOS period. The process is repeated until exhausting the T periods of time. In the end, there is a time series of $T - D$ portfolio returns that can be used to compute OOS performance measures (e.g., the OOS certainty equivalent, the OOS Sharpe ratio...)

When using the MV optimization model in a rolling window approach, the weights tend to vary a lot from period-to-period, especially if short selling is allowed, due to the presence of estimation error. In fact, it is not unusual for the weight of a particular security to vary, in just one period, from one extreme to the opposite extreme in the domain.

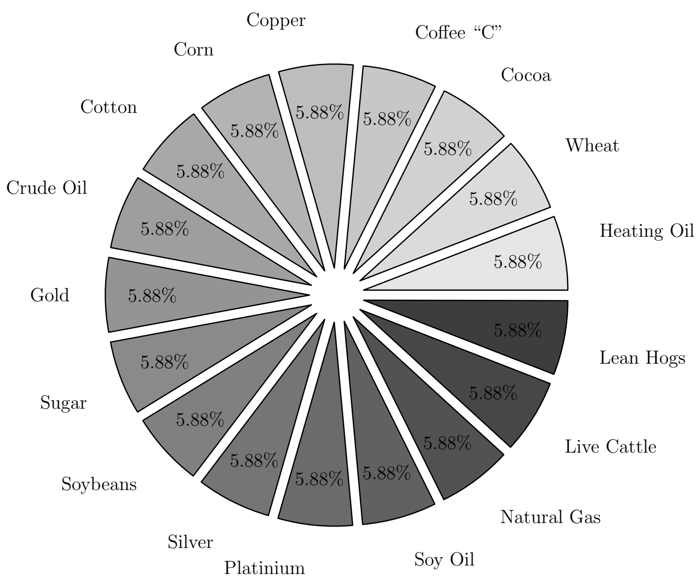
The poor OOS performance of the MV optimization model was exemplary evidenced in DeMiguel et al. (2009b). The authors evaluated the MV optimization model (while allowing for short selling) and some of its extensions, across seven empirical datasets, and showed that none is consistently better (in terms of OOS certainty equivalent, OOS Sharpe ratio and turnover) than the equally weighted portfolio.

The equally weighted portfolio (*ew* portfolio) is the one in which the available investor's wealth is divided equally among the available securities, that is:

$$w_{i,t} = \frac{1}{N}, \quad i = 1, \ldots, N. \tag{2.15}$$

This portfolio corresponds to the adage "do not put all the eggs in the same basket". As described in Benartzi and Thaler (2001), many investors, pursuing diversification, use the equally weighted portfolio to allocate their wealth. A well-known real example of an equally weighted portfolio used as a benchmark index is the Thomson Reuters Equal Weight Commodity Index (see Figure 2.6).

Figure 2.6: Thomson Reuters Equal Weight Commodity Index



This figure describes the Thomson Reuters Equal Weight Commodity Index. At the current date, July 2017, a detailed description of this index can be found at `https://financial.thomsonreuters.com/`.

DeMiguel et al. (2009b) simulate with the US equity market, and conclude that for a portfolio with 25 securities an estimation window of approximately 3000 months (250 years!) is needed for the MV optimization model to outperform the equally weighted portfolio. Moreover, as the number of available securities increases, the length of the necessary estimation window increases accordingly. Arguably, this puzzling result is due to the absence of estimation error and to the intrinsic high level of diversification of the *ew* portfolio. Furthermore, the results presented in DeMiguel et al. (2009b) indicate that there is still much to be done aiming to reduce the estimation error present in several portfolio choice models. The equally weighted portfolio is therefore an excellent benchmarking strategy, that will be used in Chapter 3 and Chapter 5 of this thesis.

In order to mitigate the error present in the estimation of the various parameters (expected returns, variances and covariances), several different approaches have been proposed in the literature. For instance, Chang et al. (1999); Andersen et al. (2001a); Goldfarb and Iyengar (2003); Ledoit and Wolf (2004a;b); DeMiguel and Nogales (2009) propose procedures to improve the estimation of the covariance matrix; while Klein and Bawa (1976); Jorion (1986); Fama and French (1992); Pastor and Stambaugh (2000); DeMiguel et al. (2017) deal with the estimation of the expected return. In fact, special attention has been devoted to the estimation of the expected returns, since its estimation error is much bigger than the one in the covariance matrix (Merton, 1980; Chopra and Ziemba, 1993). Under the Gaussian distribution assumption, Merton (1980) has shown, for the first time, that the accuracy of the estimation of the expected return increases only with T, whereas the accuracy of the estimation of the variance increases with the frequency Q of the observations for a fixed T. Additionally, Chopra and Ziemba (1993) found that the estimation errors present in the estimation of the expected returns are over ten times as costly (in terms of certainty equivalent loss) as the estimation error present in the variances. Accordingly, Kan and Zhou (2007) have observed that the estimation error of the expected returns have a larger influence on the OOS performance.

Motivated by the difficulty in computing accurate estimates of the expected returns, Jagannathan and Ma (2003) showed the superior OOS performance of the minimum variance portfolio (*mv* portfolio) with the constraint on non-short selling.

The *mv* portfolio corresponds to the solution of the following convex QP problem

$$\begin{aligned} \min_{w_t \in \mathbb{R}^N} \quad & w_t^\top \Sigma_t w_t \\ \text{subject to} \quad & 1_N^\top w_t = 1, \\ & w_t \geq 0_N. \end{aligned} \tag{2.16}$$

Jagannathan and Ma (2003) observed that the estimates of the mean returns are so noisy that it is preferable to ignore these estimates at all and use only the covariance matrix. The authors also have shown that not allowing for short selling on the minimum variance portfolio has a regularizing effect (Jagannathan and Ma, 2003). Moreover, and contrary to MPT, Baker et al. (2011) have documented and analyzed the so-called low-volatility anomaly, i.e., stocks with higher variance have historically underperformed low variance stocks. Thus, the *mv* portfolio tends to exhibit a good OOS performance.

For this reason, the *mv* portfolio is used as a benchmark portfolio in Chapter 3 and Chapter 5.

2.3 Portfolio choice with higher moments

In order to overcome some limitations of the MPT, in this thesis we explore the potential of new models for portfolio choice. The line of research is to built models where return and risk are treated differently in comparison to the MV optimization model. Driven by skeptical empiricism, we put aside the Gaussian framework and, without forcing any assumption on the returns' distribution, we look at higher statistical moments. In fact, there are several studies suggesting gains if, instead of considering just the first two moments of the returns' distribution, higher moments are also included in the portfolio choice problem (see, e.g., Chunhachinda et al., 1997; Athayde and Flôres, 2004; Maringer and Parpas, 2009; Mencia and Sentana, 2009; Harvey et al., 2010).

We could include the higher moments directly into moment-based objective functions, however, moment-based objective functions do not allow the study of the dependence of the portfolio weights on the initial wealth, while the utility-based framework allows us to investigate the dependence of the portfolio structure on different levels of investment (Bamberg and Dorfleitner, 2013). In other words, in general, moment-based objective functions do not necessarily fit with expected utility, as utility depends on the initial wealth, while moments of the return distribution do not (see Bamberg and Dorfleitner, 2013, for further details). Nevertheless, we highlight that a utility-based framework is fully compatible with a framework based on the moments of returns, if the investor considers a CRRA-utility (Bamberg and Dorfleitner, 2013). We have thus decided to formulate the suggested models according to the expected utility maximization criteria, where the dependence on the inicial wealth, W, can be studied. But we would like to make it clear that the suggested models are not restricted to the normative expected utility theory.

According to the expected utility theory, the standard utility maximization investor's problem can be formulated as

$$\begin{aligned} \max_{w_t \in \mathbb{R}^N} \quad & E_t\left[u\left(r_{p,t+1}\right)\right] = E_t\left[u\left(\sum_{i=1}^{N} w_{i,t} r_{i,t+1}\right)\right] \\ \text{subject to} \quad & w_t \in P_t, \end{aligned} \tag{2.17}$$

where $u(\cdot)$ is a Von Neumann and Morgenstern (1953) utility function. For example, assuming that the investor's preferences are described by a quadratic utility, the MV optimization model (see Problem (2.10) and Problem (2.12)) can be stated as

$$\begin{aligned} \max_{w_t \in \mathbb{R}^N} \quad & \mu_t^\top w_t - \tfrac{\xi}{2} w_t^\top \Sigma_t w_t \\ \text{subject to} \quad & w_t \in P_t, \end{aligned} \tag{2.18}$$

where $\xi \geq 0$ represents the investor's absolute risk aversion parameter.

Instead of considering MV preferences, we follow Brandt et al. (2009) and assume that the investor maximizes her expected utility, characterized by CRRA preferences. The main reason why we adopt this particular utility function is the same one that was stated by Brandt et al. (2009, p. 3421): "the advantage of CRRA utility is that it incorporates preferences toward higher moments in a parsimonious manner. In addition, the utility function is twice continuously differentiable, which allows us to use more efficient numerical optimization algorithms that make use of the analytic gradient and Hessian of the objective function." Also, as referred in Aït-Sahalia and Brandt (2001, p. 1312): "CRRA preferences are by far the most popular objective function in the portfolio choice literature. This is largely because the investor's portfolio (and consumption) policy is proportional to wealth and the value function is homothetic in wealth."

Assuming that the investor has CRRA preferences her utility can be given by

$$u\left(r_{p,t+1}\right) = \begin{cases} \dfrac{\left(1 + r_{p,t+1}\right)^{1-\gamma}}{1-\gamma} & \text{if} \quad \gamma > 1, \\ \\ \log(1 + r_{p,t+1}) & \text{if} \quad \gamma = 1, \end{cases} \tag{2.19}$$

where γ represents the relative risk aversion coefficient (the higher the value of γ the more risk averse is the investor). Note that the mathematical definition of the CRRA-utility presented in Equation (2.19) is exactly the same used in Brandt et al. (2009). However, the standard definition of a CRRA-utility is

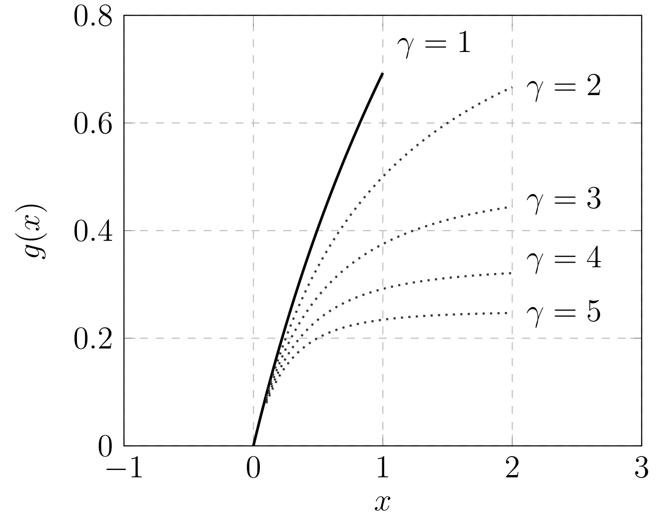
$$g\left(r_{p,t+1}\right) = \frac{\left(1 + r_{p,t+1}\right)^{1-\gamma} - 1}{1-\gamma}, \quad \gamma > 1. \tag{2.20}$$

This standard definition ensures that the CRRA-utility, $g(\cdot)$, when $\gamma \to 1$ converges to the log-utility[3], $\log(\cdot)$ (see Figure 2.7). However, notice that $g(\cdot)$ and $u(\cdot)$ represent exactly the same preferences, since for $\gamma > 1$, $g(\cdot) = u(\cdot) - \frac{1}{1-\gamma}$, i.e., $g(\cdot)$ corresponds to

[3]Using the L'Hôpital's rule:

a positive affine transformation of $u(\cdot)$ (the expected utility form is preserved, see Mas-Colell et al., 1995, for further details). In this thesis, for the sake of simplicity, we use $u(\cdot)$, except in the first part of Chapter 5, where we use $g(\cdot)$, just to evidence a particular pattern.

Figure 2.7: The CRRA utility function for different values of γ



This figure plots the CRRA utility function, $g(\cdot)$, for different risk aversion parameters ($\gamma = 1, 2, 3, 4$ and 5). When $\gamma \to 1$ the CRRA utility function collapses into $\log(1+x)$.

Now, let us denote

$$
\begin{aligned}
v_t\left(r_{p,t+1}\right) &= E_t\left[r_{p,t+1} - E_t\left(r_{p,t+1}\right)\right]^2, \\
s_t\left(r_{p,t+1}\right) &= E_t\left[r_{p,t+1} - E_t\left(r_{p,t+1}\right)\right]^3, \\
k_t\left(r_{p,t+1}\right) &= E_t\left[r_{p,t+1} - E_t\left(r_{p,t+1}\right)\right]^4,
\end{aligned}
\tag{2.21}
$$

as the portfolio variance, skewness and kurtosis at day $t+1$, respectively. Then, the fourth[4] order Taylor expansion of the expected utility, $E_t\left[u\left(r_{p,t+1}\right)\right]$, around the expected return of the portfolio, $E_t\left(r_{p,t+1}\right)$, is given by

$$\lim_{\gamma\to 1}\frac{\left(1+r_{p,t+1}\right)^{1-\gamma}-1}{1-\gamma} = \lim_{\gamma\to 1}\frac{-1\left(1+r_{p,t+1}\right)^{1-\gamma}}{-1}\log(1+r_{p,t+1}) = \log(1+r_{p,t+1}).$$

[4]The Taylor expansion of the expected utility is commonly truncated at order four, since there are no theoretical indications, in terms of the investor's preferences, for the inclusion of higher polynomial terms (see, e.g., Kimball, 1993; Dittmar, 2002; Martellini and Ziemann, 2010, for further details).

$$
\begin{aligned}
E_t\left[u\left(r_{p,t+1}\right)\right] &\approx u\left[E_t\left(r_{p,t+1}\right)\right] + \frac{1}{2!}u''\left[E_t\left(r_{p,t+1}\right)\right]v_t\left(r_{p,t+1}\right) \\
&+ \frac{1}{3!}u'''\left[E_t\left(r_{p,t+1}\right)\right]s_t\left(r_{p,t+1}\right) \\
&+ \frac{1}{4!}u''''\left[E_t\left(r_{p,t+1}\right)\right]k_t\left(r_{p,t+1}\right).
\end{aligned} \tag{2.22}
$$

Defining

$$
\begin{aligned}
\theta_1\left[E_t\left(r_{p,t+1}\right)\right] &= u\left[E_t\left(r_{p,t+1}\right)\right], \quad \theta_2\left[E_t\left(r_{p,t+1}\right)\right] = -\frac{u''\left[E_t\left(r_{p,t+1}\right)\right]}{2}, \\
\theta_3\left[E_t\left(r_{p,t+1}\right)\right] &= \frac{u'''\left[E_t\left(r_{p,t+1}\right)\right]}{6}, \quad \theta_4\left[E_t\left(r_{p,t+1}\right)\right] = -\frac{u''''\left[E_t\left(r_{p,t+1}\right)\right]}{24},
\end{aligned} \tag{2.23}
$$

where

$$u\left[E_t\left(r_{p,t+1}\right)\right] = \begin{cases} \dfrac{\left[1+E_t\left(r_{p,t+1}\right)\right]^{1-\gamma}}{1-\gamma} & \text{if} \quad \gamma > 1, \\ \\ \log\left[1+E_t\left(r_{p,t+1}\right)\right] & \text{if} \quad \gamma = 1, \end{cases}$$

$$u''\left[E_t\left(r_{p,t+1}\right)\right] = \begin{cases} -\gamma\left[1+E_t\left(r_{p,t+1}\right)\right]^{-(\gamma+1)} & \text{if} \quad \gamma > 1, \\ \\ -\dfrac{1}{\left[1+E_t\left(r_{p,t+1}\right)\right]^2} & \text{if} \quad \gamma = 1, \end{cases}$$

$$u'''\left[E_t\left(r_{p,t+1}\right)\right] = \begin{cases} \gamma(\gamma+1)\left[1+E_t\left(r_{p,t+1}\right)\right]^{-(\gamma+2)} & \text{if} \quad \gamma > 1, \\ \\ \dfrac{2}{\left[1+E_t\left(r_{p,t+1}\right)\right]^3} & \text{if} \quad \gamma = 1, \end{cases} \tag{2.24}$$

$$u''''\left[E_t\left(r_{p,t+1}\right)\right] = \begin{cases} -\gamma(\gamma+1)(\gamma+2)\left[1+E_t\left(r_{p,t+1}\right)\right]^{-(\gamma+3)} & \text{if} \quad \gamma > 1, \\ \\ -\dfrac{6}{\left[1+E_t\left(r_{p,t+1}\right)\right]^4} & \text{if} \quad \gamma = 1, \end{cases}$$

we can rewrite Equation (2.22) as

$$\begin{aligned} E_t\left[u\left(r_{p,t+1}\right)\right] &\approx \theta_1\left[E_t\left(r_{p,t+1}\right)\right] - \theta_2\left[E_t\left(r_{p,t+1}\right)\right] v_t\left(r_{p,t+1}\right) \\ &+ \theta_3\left[E_t\left(r_{p,t+1}\right)\right] s_t\left(r_{p,t+1}\right) - \theta_4\left[E_t\left(r_{p,t+1}\right)\right] k_t\left(r_{p,t+1}\right). \end{aligned} \tag{2.25}$$

In Equation (2.25), $v_t\left(r_{p,t+1}\right)$ can be computed according Equation (2.7). In turn, $s_t\left(r_{p,t+1}\right)$ can be computed as a third moment tensor and can be visualized as a $N \times N \times N$ cube in the three-dimensional space. It is possible to transform the skewness tensor into a $N \times N^2$ matrix (see Athayde and Flôres, 2004, for further details). Following this idea $s_t\left(r_{p,t+1}\right)$ can be computed as

$$s_t\left(r_{p,t+1}\right) = E_t\left[r_{p,t+1} - E_t\left(r_{p,t+1}\right)\right]^3 = w_t^\top \Phi_t(w_t \otimes w_t), \tag{2.26}$$

where $\otimes$ denotes the Kronecker product and Φ_t is the coskewness matrix. The coskewness matrix of dimension $N \times N^2$ can be represented by N $A_{i,t}$ matrixes of dimensions $N \times N$ such that

$$\Phi_t = [\ A_{1,t} \mid A_{2,t} \mid \ \cdots \ \mid A_{N,t}\ ], \tag{2.27}$$

where

$$A_{i,t} = \begin{bmatrix} a_{i11,t} & a_{i12,t} & \cdots & a_{i1N,t} \\ a_{i21,t} & a_{i22,t} & \cdots & a_{i2N,t} \\ \vdots & \vdots & \ddots & \vdots \\ a_{iN1,t} & a_{iN2,t} & \cdots & a_{iNN,t} \end{bmatrix}, \tag{2.28}$$

and each element, $a_{ijk,t}$, is given by

$$a_{ijk,t} = \frac{1}{t}\sum_{\tau=1}^{t}(r_{i,\tau} - \mu_{i,\tau})(r_{j,\tau} - \mu_{j,\tau})(r_{k,\tau} - \mu_{k,\tau}), \tag{2.29}$$

with $i, j, k = 1, \ldots, N$.

Analogously, $k_t\,(r_{p,t+1})$, can be computed as

$$k_t\,(r_{p,t+1}) \ = \ E_t\left[r_{p,t+1} - E_t\,(r_{p,t+1})\right]^4 \ = \ w_t^\top \Psi_t(w_t \otimes w_t \otimes w_t), \tag{2.30}$$

where Ψ_t is the cokurtosis matrix. The coskurtosis matrix corresponds to N^2 matrixes $B_{ij,t}$ of dimension $N \times N$ such that

$$\Psi_t = [B_{11,t} \mid B_{12,t} \mid \ \cdots \ \mid B_{1N,t} \mid B_{21,t} \mid B_{22,t} \mid \ \cdots \ \mid B_{2N,t} \mid \ \cdots \ \mid B_{N1,t} \mid B_{N2,t} \mid \ \cdots \ \mid B_{NN,t}], \tag{2.31}$$

with

$$B_{ij,t} = \begin{bmatrix} b_{ij11,t} & b_{ij12,t} & \cdots & b_{ij1N,t} \\ b_{ij21,t} & b_{ij22,t} & \cdots & b_{ij2N,t} \\ \vdots & \vdots & \ddots & \vdots \\ b_{ijN1,t} & b_{ijN2,t} & \cdots & b_{ijNN,t} \end{bmatrix}, \tag{2.32}$$

and where each element, $b_{ijkl,t}$, is given by

$$b_{ijkl,t} = \frac{1}{t} \sum_{\tau=1}^{t} (r_{i,\tau} - \mu_{i,\tau})(r_{j,\tau} - \mu_{j,\tau})(r_{k,\tau} - \mu_{k,\tau})(r_{l,\tau} - \mu_{l,\tau}), \tag{2.33}$$

with $i, j, k, l = 1, \ldots, N$.

According to the objective function given by Equation (2.25), we can formulate the investor's problem with higher moments as

$$\begin{aligned} \max_{w_t \in \mathbb{R}^N} \quad & \theta_1 \left[E_t \left(r_{p,t+1} \right) \right] - \theta_2 \left[E_t \left(r_{p,t+1} \right) \right] v_t \left(r_{p,t+1} \right) \\ & + \theta_3 \left[E_t \left(r_{p,t+1} \right) \right] s_t \left(r_{p,t+1} \right) - \theta_4 \left[E_t \left(r_{p,t+1} \right) \right] k_t \left(r_{p,t+1} \right) \\ \text{subject to} \quad & w_t \in P_t. \end{aligned} \tag{2.34}$$

Chapter 3

Efficient skewness/semivariance portfolios

3.1 Introduction

With non-normal return distributions, the use of a downside risk measure is more suitable than the traditional use of the variance (Nawrocki, 1999). Roy (1952) was the first author to use a downside risk measure in portfolio selection, in the form of a "safety first" rule, that measures the probability of outcomes falling below a predetermined target return (Sing and Ong, 2000). Markowitz (1959) recognised the importance of Roy's work, arguing that there are more plausible measures of risk than the variance, and proposed the use of a below-mean semivariance or a target return semivariance. These metrics belong to a more general family of downside risk measures known as lower partial moments (Bawa, 1975; Fishburn, 1977). Quirk and Saposnik (1962) confirmed the theoretical superiority of the semivariance versus the variance, while Ang and Chua (1979) showed the superiority of the target return semivariance relative to the below-mean semivariance. However, the semivariance is seldom used in portfolio selection problems due to the endogeneity of the cosemivariance.

On the grounds of the semivariance being a more plausible measure of risk than the variance and the investor's preferences and skewness being positively related (see, e.g., Arditti, 1975; Kraus and Litzenberger, 1976; Harvey and Siddique, 2000), we suggest a direct analysis of the efficient tradeoff between skewness and semivariance by means of a biobjective optimization problem. This methodology is flexible, in the sense that the investor is free to choose the target return, required for the semivariance computation. Another strong point is that skewness is interpreted as a third moment tensor and the endogeneity issue of the cosemivariance matrix is addressed explicitly. Given the endogeneity of the cosemivariance matrix, we use a derivative-free algorithm (based on direct multisearch (DMS)) to obtain the solution of the biobjective optimization problem. DMS (see Appendix B) is a class of methods used in multiobjective optimization (MOO) problems (see Appendix B) that does not use derivatives and does not aggregate or scalarize any of the objective function components. It essentially generalizes all direct-search methods of directional type from single to MOO. For a complete description of DMS see the algorithmic framework in Custódio et al. (2011).

The contribution of this chapter is therefore twofold. First, we suggest a skewness/semivariance biojective model that allows the investor to directly analyze the efficient tradeoff between skewness and semivariance (regardless of the target return used in the semivariance calculation). Second, through a derivative-free algorithm we overcome the endogeneity problem of the cosemivariance matrix.

The empirical work is conducted on four datasets collected from the Fama/French online data library. First, the Pareto frontiers on the skewness/semivariance space are computed using three different target returns, corresponding to the returns of the ms, the mv and the ew portfolios, respectively. Then, an extensive OOS performance analysis is implemented on three efficient skewness/semivariance portfolios from the IS Pareto frontiers: the portfolios with the maximum skewness per semivariance ratio, with the maximum Sharpe ratio and with the maximum Sortino ratio. The OOS performance is measured in terms of skewness per semivariance ratio, Sharpe ratio and Sortino ratio. We conclude that the efficient portfolios exhibit a competitive OOS performance compared with the three benchmark portfolios. One interesting result is that in the four datasets, at least one of the three chosen efficient skewness/semivariance portfolios consistently outperform the ew portfolio in terms of Sharpe ratio, which is known to be difficult to achieve (DeMiguel et al., 2009b).

The rest of the chapter is organized as follows. In Section 3.2 it is shown that the expected utility of a risk averse investor is an increasing function of the skewness and a decreasing function of the semivariance. Section 3.3 presents the proposed model. Section 3.4 shows the empirical results and, finally, Section 3.5 summarizes the main findings and discusses future research.

3.2 Investor expected utility

3.2.1 Expected utility based on skewness

Let $u(\cdot)$ be the utility function of a typical investor. If instead of a fourth order Taylor expansion (as considered in Equation (2.22)) of the expected utility, $E_t\left[u\left(r_{p,t+1}\right)\right]$, around the expected return of the portfolio, $E_t\left(r_{p,t+1}\right)$, we consider the third order approximation of the expected utility, as in Joro and Na (2006), then

$$
\begin{aligned}
E_t\left[u\left(r_{p,t+1}\right)\right] \;&\approx\; u\left[E_t\left(r_{p,t+1}\right)\right] \;+\; \frac{1}{2!}u''\left[E_t\left(r_{p,t+1}\right)\right] v_t\left(r_{p,t+1}\right) \\
&+\; \frac{1}{3!}u'''\left[E_t\left(r_{p,t+1}\right)\right] s_t\left(r_{p,t+1}\right).
\end{aligned} \tag{3.1}
$$

Since $u'''\left[E_t\left(r_{p,t+1}\right)\right] \geq 0$ (see Arditti (1967) and Kraus and Litzenberger (1976)), the expected utility, $E_t\left[u\left(r_{p,t+1}\right)\right]$, of a risk averse investor is an increasing function of skewness, $s_t\left(r_{p,t+1}\right)$, which is consistent with the desirable properties for an investor's utility function ($u'(\cdot) > 0$, $u''(\cdot) < 0$ and $u'''(\cdot) > 0$) suggested by Arrow (1971).

3.2.2 Expected utility based on semivariance

The semivariance can be defined as

$$
sv_{R,t}(r_{p,t+1}) \;=\; E_t\left\{\left[\min(r_{p,t+1} - R, 0)\right]^2\right\}, \tag{3.2}
$$

where R represents a target return and should be independent of the probability distribution being ranked (Ang and Chua, 1979). If instead of the variance, we consider the semivariance, then the utility function should have a kink at the reference point R. On the basis of Koekebakker and Zakamouline (2009), the utility function has the form

$$
u(r_{p,t+1}) = \begin{cases} u_+(r_{p,t+1}) & \text{if } r_{p,t+1} \geq R, \\ \\ u_-(r_{p,t+1}) & \text{if } r_{p,t+1} < R, \end{cases} \tag{3.3}
$$

where $u_+(\cdot)$ is the utility function for gains and $u_-(\cdot)$ is the utility function for losses. This is in accordance with the descriptive framework proposed by Kahneman and Tversky (1979). Considering the second order Taylor expansion approximation of $E_t\left[u\left(r_{p,t+1}\right)\right]$, around the target return, R, we have

$$
E_t\left[u\left(r_{p,t+1}\right)\right] \approx \begin{cases} u_+(R) + u'_+(R)E_t\left[(r_{p,t+1} - R)\right] + \dfrac{1}{2}u''_+(R)E_t\left[(r_{p,t+1} - R)^2\right] & \text{if } r_{p,t+1} \geq R, \\ \\ u_-(R) + u'_-(R)E_t\left[(r_{p,t+1} - R)\right] + \dfrac{1}{2}u''_-(R)E_t\left[(r_{p,t+1} - R)^2\right] & \text{if } r_{p,t+1} < R. \end{cases} \tag{3.4}
$$

If the investor is risk averse in the domain of losses (the utility function for losses is concave, $u''_{-}(\cdot) < 0$), then the expected utility, $E_t\left[u\left(r_{p,t+1}\right)\right]$, is a decreasing function of semivariance, $sv_{R,t}(r_{p,t+1})$. Again, this is consistent with the desirable properties for an investor's utility function ($u'(\cdot) > 0$, $u''(\cdot) < 0$ and $u'''(\cdot) > 0$).

3.3 The skewness/semivariance biobjective model

To overcome the limitation of the MV models, some researchers used downside risk measures, which only gauge the negative deviations from a reference return level. One famous downside risk measure was introduced in the "safety first" criterion (Roy, 1952). Other downside risk measures were proposed, for example, in Bawa (1975); Fishburn (1977); Harlow and Rao (1989); Nawrocki (1999). See Nawrocki (1999) for a survey on downside risk measures.

Markowitz (1959) favored one of the best-known downside risk measures: the semivariance of returns. The semivariance can be handled by considering an asymmetric cosemivariance matrix (Hogan and Warren, 1974) or considering a symmetric and exogenous cosemivariance matrix (Estrada, 2008). Another way of handling the semivariance is outside the stochastic environment, considering the fuzzy set environment as in Huang (2008).

Following Markowitz (1959), the endogenous cosemivariance matrix is the approach adopted here. Therefore the exact estimation of the semivariance of a portfolio, $sv_t^R(r_{p,t+1})$, is obtained as

$$sv_t^R(r_{p,t+1}) = \sum_{i=1}^{N}\sum_{j=1}^{N} w_{i,t}w_{j,t}cs_{ij,t} = w_t^\top \Sigma_{R,t}(w_t)w_t, \tag{3.5}$$

where $\Sigma_{R,t}$ is the cosemivariance matrix in which each entry $cs_{ij,t}$ is given by

$$cs_{ij,t} = \frac{1}{t-1}\sum_{k\in U}(r_{i,k} - R)(r_{j,k} - R), \tag{3.6}$$

with $i, j = 1, \ldots, N$, and

$$U = \{\tau \mid r_{p,\tau} < R\} \subseteq \{1, \ldots, t-1\}\,. \tag{3.7}$$

The cosemivariance matrix is endogenous in the sense that a change in the portfolio's weights affects the periods when the portfolio underperforms the benchmark, which in turn affects the elements of the cosemivariance matrix (Estrada, 2008). Notice that the target return, R, is a parameter that can be freely chosen by the investor, according to her own preferences.

There is an intuitive explanation for why the skewness is important for the investor. Clearly, the investor has a preference for positive skewness in order to have higher probability for extreme profit values and limited losses. Alderfer and Bierman (1970) showed empirically that investors prefer positive skewness, even if this positive skewness is associated with a lower expected return. Arditti (1975), Kraus and Litzenberger (1976) and Harvey and Siddique (2000) have shown theoretically that investors should prefer positive skewness. Moreno and Rodríguez (2009) showed that the coskewness is taken into account by funds managers, representing an important factor in the selection of securities. However, skewness is often neglected in the performance evaluation literature, possibly due to computational difficulties (Joro and Na, 2006).

In this study, we propose the simultaneous consideration of the two investor's objectives

- Maximizing the portfolio skewness: $s_t(r_{p,t+1}) \;=\; w_t^\top \Phi_t(w_t \otimes w_t)$,
- Minimizing the portfolio semivariance: $sv_t^R(r_{p,t+1}) \;=\; w_t^\top \Sigma_{R,t}(w_t) w_t$,

over the set of feasible portfolios. In this case, the skewness/semivariance biobjective optimization model can be written as

$$\begin{aligned} \max_{w_t \in \mathbb{R}^N} \quad & w_t^\top \Phi_t(w_t \otimes w_t) \\ \min_{w_t \in \mathbb{R}^N} \quad & w_t^\top \Sigma_{R,t}(w_t) w_t \\ \text{subject to} \quad & w_t \in P_t. \end{aligned} \tag{3.8}$$

By solving Problem (3.8), we identify a skewness/semivariance Pareto frontier. A portfolio in this frontier is such that there exists no other feasible one which simultaneously presents a higher skewness and a lower semivariance. Given such an efficient frontier and a specific semivariance level, an investor may directly find the answers to the questions of what is the optimal (higher) skewness level that can be chosen and what are the portfolios leading to such a skewness level. Problem (3.8) has two objective functions, a linear constraint and N inequality constraints. The first objective, $w_t^\top \Phi_t(w_t \otimes w_t)$, is nonlinear but smooth. However, the second objective, $w_t^\top \Sigma_{R,t}(w_t) w_t$, is nonlinear and nonsmooth due to the endogeneity problem on the cosemivariance matrix, $\Sigma_{R,t}(w_t)$. We have thus decided to solve the problem through the DMS algorithm (see Appendix B, for a description of DMS). This derivative-free algorithm was previously and for the first time used in the portfolio selection framework for solving a cardinality constrained problem (Brito and Vicente, 2014).

The skewness/semivariance biobjective model can be extensively explored by the investor. It is clear that the gains of this approach in relation to the MV one are, in a great extent, a function of the degree of asymmetry in the distribution of the portfolio returns. Given this asymmetric property, a careful choice on a skewness/semivariance efficient frontier can approximately maximize the expected utility for a variety of utility functions, as we demonstrate in Section 3.2. Moreover, the proposed model is quite flexible since it makes use of a general definition of the semivariance, by not restricting the target return (in the related literature, the target return is often set to the mean of the distribution). In fact, giving her specific preferences, the investor can choose this parameter freely.

3.4 Empirical analysis

The empirical analysis is conducted on four datasets collected from the Fama/French data library, which is publicly available (at the current date, July 2017) from the site: `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`. The datasets are constructed according to different criteria, and each one is composed by portfolios, rebalanced annually at the end of June.

The SBM6 dataset corresponds to six portfolios based on size and book to market ratio. These portfolios are the intersections of 2 portfolios based on size (market equity, ME) and 3 portfolios based on the ratio of book equity to market equity (BE/ME).

The FF10 dataset corresponds to 10 industry portfolios.

The SOP25 dataset corresponds to 25 portfolios based on size and operating profitability. These portfolios, are the intersections of 5 portfolios based on size (market equity, ME) and 5 portfolios based on operating profitability (OP).

The BMOP25 dataset corresponds to 25 portfolios formed on book-to-market and operating profitability. These portfolios, are the intersections of 5 portfolios formed on the ratio of book equity to market equity (BE/ME) and 5 portfolios formed on profitability (OP).

The overall sample for all four datasets is formed by monthly data from 07/1964 to 06/2014 (600 months). Table 3.1 reports some descriptive statistics for each dataset. The monthly continuous returns showed, on average, negative skewness and a kurtosis well above that of normal distribution. The application of the Jarque-Bera test for normality to all the portfolios of each dataset (SBM6, FF10, SOP25 and BMOP25) showed that the null hypothesis of normality was rejected with p-values lower than 1%.

Table 3.1: Descriptive statistics

Dataset	Mean	Variance	Skewness	Kurtosis
SBM6	0.0105	0.0035	-0.5401	6.3148
FF10	0.0103	0.0043	-0.4678	6.5228
SOP25	0.0098	0.0034	-0.6280	6.3244
BMOP25	0.0114	0.0039	-0.5324	6.6087

This table reports some descriptive statistics for the four datasets collected from the Fama/French data library. The reported values of skewness and kurtosis concern to the third and fourth standardized moments, respectively. These statistics are averaged cross-sectionally.

3.4.1 IS analysis

We applied the `dms`[5] solver (see Appendix B) to determine the Pareto frontier of the skewness/semivariance biobjective optimization model (solution of Problem (3.8)) for three different values of R: R_{ms}, the return of the maximum Sharpe ratio portfolio[6] (solution of Problem (2.13)); R_{mv}, the return of the minimum variance portfolio (solution of Problem (2.16)); and R_{ew}, the return of the *ew* portfolio (given by Equation (2.15)). In this empirical exercise, we choose these three different values for the target return, R, in order to show the robustness of the model. However, this parameter can be freely chosen according to the specific preferences of the investor.

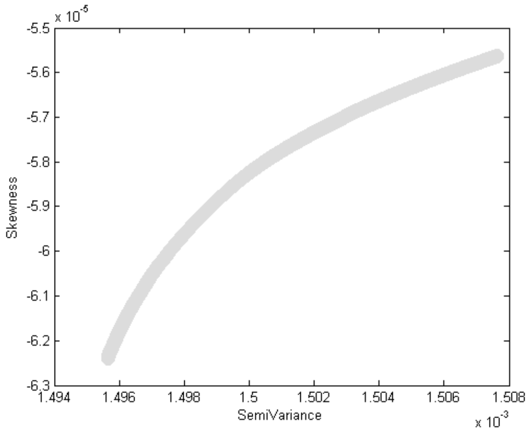
Figures 3.1, 3.2 and 3.3, contain the plots of the Pareto frontiers, computed using the overall sample period, for the SBM6 dataset, with target returns given by R_{ms}, R_{mv} and R_{ew}, respectively. Similarly, we obtained the Pareto frontiers for the remaining datasets (we do not report all the plots since similar patterns were found). For each dataset, the efficient skewness/semivariance frontier differs according to the choice of the target return, R, emphasizing the importance of this parameter in the proposed approach.

As it is clear from the visualization of the plots, DMS was able to determine the Pareto frontiers for the biobjective skewness/semivariance optimization problems (solutions of Problem (3.8)) for all the instances considered. Thus, this methodology offers a direct way for analyzing the efficient tradeoff between skewness and semivariance.

[5] This solver is public and available (at the current date, July 2017) by request at `http://www.mat.uc.pt/dms/`.

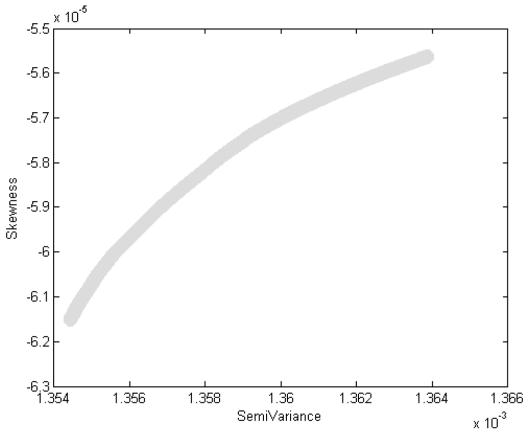
[6] We considered as a risk-free asset the 90-day Treasury-Bills US. Such data is public and made available (at the current date, July 2017) by the Federal Reserve, at the site `www.federalreserve.gov`.

Figure 3.1: SBM6 skewness/semivariance efficient frontier, with R_{ms} as the target return for semivariance computation



This figure shows the plot of the efficient skewness/semivariance frontier for the SBM6 dataset and considering as target return R_{ms}. The horizontal axis corresponds to the portfolio semivariance. The vertical axis corresponds to the portfolio skewness.

Figure 3.2: SBM6 skewness/semivariance efficient frontier, with R_{mv} as the target return for semivariance computation

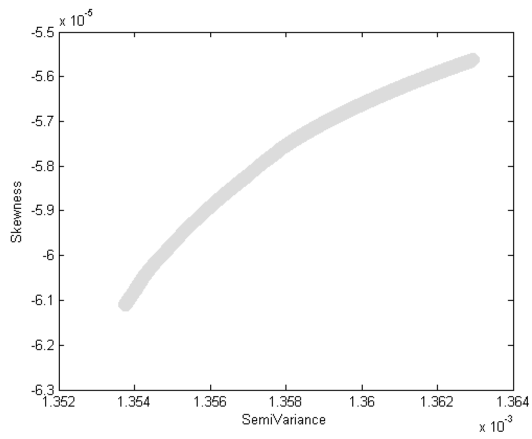


This figure shows the plot of the efficient skewness/semivariance frontier for the SBM6 dataset and considering as target return R_{mv}. The horizontal axis corresponds to the portfolio semivariance. The vertical axis corresponds to the portfolio skewness.

3.4.2 OOS analysis

The validation of a new methodology for portfolio selection must be based on an OOS performance analysis. This section deals with an extensive OOS analysis of the efficient skewness/semivariance portfolios, constructed according to the proposed model in Section 3.3, and compared with each of the benchmark portfolios: the *ms* portfolio (see Problem (2.13)), the *mv* portfolio (see Problem (2.16)) and the *ew* portfolio (see

Figure 3.3: SBM6 skewness/semivariance efficient frontier, with R_{ew} as the target return for semivariance computation



This figure shows the plot of the efficient skewness/semivariance frontier for the SBM6 dataset and considering as target return R_{ew}. The horizontal axis corresponds to the portfolio semivariance. The vertical axis corresponds to the portfolio skewness.

Equation (2.15)).

The OOS analysis relies on a rolling window approach (see Figure 2.5). We considered an estimation window of 120 months, with an initial estimation period from 07/1964 to 06/1974. The evaluation period comprised 480 months, from 07/1974 to 06/2014. For each estimation window, the benchmark portfolios (the ms portfolio, the mv portfolio and the ew portfolio) were computed. Then, the Pareto frontiers of the skewness/semivariance biobjective optimization Problem (3.8) were determined, considering R_{ms}, R_{mv} and R_{ew} as target returns for the semivariance computation. Finally, three efficient skewness/semivariance portfolios in each of the IS Pareto frontiers were selected. The first one, w_{SSR}, was the portfolio that maximizes a skewness per semivariance ratio (SSR)

$$SSR = \frac{w_{SSR}^{\top}\Phi(w_{SSR} \otimes w_{SSR})}{w_{SSR}^{\top}\Sigma_R(w_{SSR})w_{SSR}}. \tag{3.9}$$

The second one, w_{SR}, was the portfolio that maximizes the Sharpe ratio (SR)

$$SR = \frac{\mu^{\top}w_{SR} - r^f}{\sqrt{w_{SR}^{\top}\Sigma w_{SR}}}. \tag{3.10}$$

The third one, w_{SOR}, was the portfolio that maximizes the Sortino ratio (SOR)

$$SOR = \frac{\mu^\top w_{SOR} - R}{\sqrt{w_{SOR}^\top \Sigma_R(w_{SOR}) w_{SOR}}}. \tag{3.11}$$

The numerator of these ratios for the chosen portfolios may be negative. Thus, in order to have correct rankings, the denominators are modified as proposed by Israelsen (2005). Finally, each portfolio is held fixed and its returns were observed over the next month. The estimation window was then moved forward 1 month, and the returns were calculated for the next month of the evaluation period. The process was thus repeated until the end of the evaluation period was reached.

Performance measured by a skewness per semivariance ratio

We computed an OOS skewness per semivariance ratio, defined as the sample skewness, $\widehat{s}$, divided by the sample semivariance, $\widehat{sv}_R$:

$$\widehat{\text{SSR}} = \frac{\widehat{s}}{\widehat{sv}_R}. \tag{3.12}$$

Then, we computed the bootstrap p-values of the difference between the skewness per semivariance ratio of each efficient skewness/semivariance portfolio and those of the benchmarks. Since none of the differences were statistically significant, we decided do not report these results here. Once the computed skewness per semivariance ratios are negative (we are in the presence of negative skewness), in order to achieve a correct rank of the portfolios, it is necessary to refine the ratios. We modified the denominator according the methodology proposed by Israelsen (2005). Thus we computed the refined skewness per semivariance ratio as

$$\widehat{\text{SSR}}_{\text{ref}} = \frac{\widehat{s}}{\widehat{sv}_R^{\,\widehat{s}/abs(\widehat{s})}}, \tag{3.13}$$

where $abs(\cdot)$ is the absolute value function.

Table 3.2 reports the refined skewness per semivariance ratios, when we choose as a target return the maximum Sharpe ratio portfolio return (R_{ms}). We can see that the efficient skewness/semivariance portfolios w_{SR} and w_{SOR} have a higher refined skewness per semivariance ratio than two (the ms and the ew portfolios) of the three benchmarks portfolios, for all the datasets. For the datasets with the highest number of securities (SOP25 and BMOP25) all the efficient skewness/semivariance portfolios (w_{SSR}, w_{SR}

and w_{SOR}) have a higher refined skewness per semivariance ratio than two (the *ms* and the *ew* portfolios) of the three benchmarks portfolios. The same pattern is found when we choose as a target return the minimum variance portfolio return (R_{mv}) (see Table 3.3) and the *ew* portfolio return (R_{ew}) (see Table 3.4).

These results suggests the robustness of the efficiency provided by the skewness/ semivariance model.

Table 3.2: Portfolio refined skewness per semivariance ratios for the target return R_{ms}

Strategy	SBM6	#Rank	FF10	#Rank	SOP25	#Rank	BMOP25	#Rank
ms portfolio	-179.1711	4	-63.4597	4	-193.5598	5	-204.7025	5
mv portfolio	-134.5284	1	-8.8977	1	-76.8224	1	-135.2051	1
ew portfolio	-246.1647	6	-272.9040	6	-278.8570	6	-298.5436	6
Efficient skewness/semivariance portfolios								
w_{SSR}	-216.9560	5	-90.9043	5	-151.9139	4	-182.6608	4
w_{SR}	-151.7857	2	-28.1328	2	-116.4111	2	-166.0968	2
w_{SOR}	-171.6702	3	-57.4604	3	-128.0863	3	-175.9555	3

This table reports, for each dataset, the monthly refined skewness per semivariance ratios using as benchmark portfolios the maximum Sharpe ratio portfolio (ms portfolio), the minimum variance portfolio (mv portfolio) and the equally weighted portfolio (ew portfolio). All the monthly refined skewness per semivariance ratios values are multiplied by a factor of 10^9. *The target return used in the computation of the semivariance is the maximum Sharpe ratio portfolio return* (R_{ms}). *This table also reports the monthly refined skewness per semivariance ratios for the efficient skewness/semivariance portfolios referred in Section 3.4.2: the maximum skewness per semivariance ratio portfolio* (w_{SSR}), *the maximum Sharpe ratio portfolio* (w_{SR}) *and the maximum Sortino ratio portfolio* (w_{SOR}). *The correct rank of each portfolio according to the refined skewness per semivariance ratios is also reported.*

Performance measured by the Sharpe ratio

Given the time series of monthly OOS returns, for each portfolio we computed the OOS Sharpe ratio, defined as the sample mean of excess returns (over the risk-free asset), $\widehat{m}$, divided by its sample standard deviation, $\widehat{\sigma}$:

$$\widehat{\text{SR}} = \frac{\widehat{m}}{\widehat{\sigma}}. \tag{3.14}$$

The results are presented in Table 3.5. This table also reports the bootstrap *p*-values (Ledoit and Wolf, 2008) for the statistical significance of the difference between the Sharpe ratios of the benchmarks and the efficient skewness/semivariance portfolios.

For the SBM6 dataset, independently of the target return used in the computation of the semivariance (R_{ms}, R_{mv} or R_{ew}), the efficient skewness/semivariance portfolio w_{SOR}

Table 3.3: Portfolio refined skewness per semivariance ratios for the target return R_{mv}

Strategy	SBM6	#Rank	FF10	#Rank	SOP25	#Rank	BMOP25	#Rank
ms portfolio	-152.8876	4	-62.8689	4	-174.6185	5	-182.4181	4
mv portfolio	-112.5562	1	-8.7818	1	-68.1002	1	-118.4064	1
ew portfolio	-209.8894	6	-270.8331	6	-252.9930	6	-265.9876	6
Efficient skewness/semivariance portfolios								
w_{SSR}	-185.5340	5	-92.3814	5	-141.7455	4	-184.6669	5
w_{SR}	-129.6049	2	-28.5004	2	-106.5681	2	-156.2354	2
w_{SOR}	-150.6649	3	-51.6755	3	-115.8668	3	-166.2062	3

This table reports, for each dataset, the monthly refined skewness per semivariance ratios using as benchmark portfolios the maximum Sharpe ratio portfolio (ms portfolio), the minimum variance portfolio (mv portfolio) and the equally weighted portfolio (ew portfolio). All the monthly refined skewness per semivariance ratios values are multiplied by a factor of 10^9*. The target return used in the computation of the semivariance is the minimum variance portfolio return (*R_{mv}*). This table also reports the monthly refined skewness per semivariance ratios for the efficient skewness/semivariance portfolios referred in Section 3.4.2: the maximum skewness per semivariance ratio portfolio (*w_{SSR}*), the maximum Sharpe ratio portfolio (*w_{SR}*) and the maximum Sortino ratio portfolio (*w_{SOR}*). The correct rank of each portfolio according to the refined skewness per semivariance ratios is also reported.*

Table 3.4: Portfolio refined skewness per semivariance ratios for the target return R_{ew}

Strategy	SBM6	#Rank	FF10	#Rank	SOP25	#Rank	BMOP25	#Rank
ms portfolio	-161.6179	4	-65.4694	4	-180.9697	5	-185.5512	4
mv portfolio	-119.8545	1	-9.2942	1	-71.0221	1	-120.5573	1
ew portfolio	-221.9983	6	-279.9076	6	-261.7112	6	-270.2061	6
Efficient skewness/semivariance portfolios								
w_{SSR}	-195.3805	5	-96.0909	5	-146.9785	4	-202.0638	5
w_{SR}	-135.8992	2	-29.6964	2	-109.3740	2	-155.2521	2
w_{SOR}	-153.7604	3	-60.1979	3	-123.1379	3	-176.1691	3

This table reports, for each dataset, the monthly refined skewness per semivariance ratios using as benchmark portfolios the maximum Sharpe ratio portfolio (ms portfolio), the minimum variance portfolio (mv portfolio) and the equally weighted portfolio (ew portfolio). All the monthly refined skewness per semivariance ratios values are multiplied by a factor of 10^9*. The target return used in the computation of the semivariance is the equally weighted portfolio return (*R_{ew}*). This table also reports the monthly refined skewness per semivariance ratios for the efficient skewness/semivariance portfolios referred in Section 3.4.2: the maximum skewness per semivariance ratio portfolio (*w_{SSR}*), the maximum Sharpe ratio portfolio (*w_{SR}*) and the maximum Sortino ratio portfolio (*w_{SOR}*). The correct rank of each portfolio according to the refined skewness per semivariance ratios is also reported.*

has a higher Sharpe ratio than all the three benchmark portfolios (the *ms* portfolio, the *mv* portfolio and the *ew* portfolio). For all the target returns (R_{ms}, R_{mv} and R_{ew}), the difference between the Sharpe ratios of the efficient skewness/semivariance portfolio w_{SOR} and the *ew* portfolio is statistically significant (at the 5% level).

Table 3.5: OOS Sharpe ratios

Benchmark	Strategy	SBM6	FF10	SOP25	BMOP25
	ms portfolio	0.2572	0.2736	0.2113	0.2993
	mv portfolio	0.2480	0.3188	0.2130	0.2935
	ew portfolio	0.2235	0.2115	0.1976	0.2357
		Efficient skewness/semivariance portfolios			
R_{ms}	w_{SSR}	0.2541	0.2689	0.2041	0.2624
		$(0.8941)^{ms}(0.7353)^{mv}(0.1459)^{ew}$	$(0.9201)^{ms}(0.1139)^{mv}(0.1628)^{ew}$	$(0.7463)^{ms}(0.6315)^{mv}(0.7431)^{ew}$	$(0.1586)^{ms}(0.2474)^{mv}(0.2559)^{ew}$
	w_{SR}	0.2472	0.3078	0.2071	0.3151
		$(0.5155)^{ms}(0.9471)^{mv}(0.1858)^{ew}$	$(0.3477)^{ms}(0.5874)^{mv}(0.0230)^{ew}$	$(0.8352)^{ms}(0.7041)^{mv}(0.6054)^{ew}$	$(0.2279)^{ms}(0.1418)^{mv}(0.0000)^{ew}$
	w_{SOR}	0.2640	0.2829	0.2125	0.3019
		$(0.7343)^{ms}(0.3846)^{mv}(0.0320)^{ew}$	$(0.8065)^{ms}(0.2142)^{mv}(0.0732)^{ew}$	$(0.9481)^{ms}(0.9840)^{mv}(0.4366)^{ew}$	$(0.8876)^{ms}(0.6927)^{mv}(0.0030)^{ew}$
R_{mv}	w_{SSR}	0.2539	0.2676	0.2018	0.2596
		$(0.8761)^{ms}(0.7430)^{mv}(0.1429)^{ew}$	$(0.8921)^{ms}(0.1074)^{mv}(0.1698)^{ew}$	$(0.6813)^{ms}(0.5395)^{mv}(0.8391)^{ew}$	$(0.1339)^{ms}(0.1974)^{mv}(0.3071)^{ew}$
	w_{SR}	0.2485	0.3057	0.2077	0.3109
		$(0.5845)^{ms}(0.9640)^{mv}(0.1748)^{ew}$	$(0.3786)^{ms}(0.5071)^{mv}(0.0212)^{ew}$	$(0.8571)^{ms}(0.7353)^{mv}(0.5659)^{ew}$	$(0.3676)^{ms}(0.2174)^{mv}(0.0000)^{ew}$
	w_{SOR}	0.2629	0.2842	0.2097	0.3051
		$(0.7702)^{ms}(0.3766)^{mv}(0.0320)^{ew}$	$(0.8002)^{ms}(0.2169)^{mv}(0.0610)^{ew}$	$(0.9386)^{ms}(0.8596)^{mv}(0.5052)^{ew}$	$(0.7113)^{ms}(0.5451)^{mv}(0.0000)^{ew}$
$R_{1/N}$	w_{SSR}	0.2540	0.2658	0.2004	0.2566
		$(0.8872)^{ms}(0.7522)^{mv}(0.1494)^{ew}$	$(0.8656)^{ms}(0.1051)^{mv}(0.1850)^{ew}$	$(0.6324)^{ms}(0.5073)^{mv}(0.8941)^{ew}$	$(0.0979)^{ms}(0.1694)^{mv}(0.3677)^{ew}$
	w_{SR}	0.2488	0.3050	0.2057	0.3127
		$(0.6099)^{ms}(0.9365)^{mv}(0.1688)^{ew}$	$(0.3921)^{ms}(0.4835)^{mv}(0.0240)^{ew}$	$(0.7802)^{ms}(0.6611)^{mv}(0.6547)^{ew}$	$(0.2814)^{ms}(0.1948)^{mv}(0.0000)^{ew}$
	w_{SOR}	0.2634	0.2806	0.2099	0.3073
		$(0.7642)^{ms}(0.3876)^{mv}(0.0280)^{ew}$	$(0.8661)^{ms}(0.1942)^{mv}(0.0856)^{ew}$	$(0.9540)^{ms}(0.8601)^{mv}(0.5097)^{ew}$	$(0.6142)^{ms}(0.4953)^{mv}(0.0000)^{ew}$

This table reports, the monthly Sharpe ratios using as benchmark portfolios the maximum Sharpe ratio portfolio (ms portfolio), the minimum variance portfolio (mv portfolio) and the equally weighted portfolio (ew portfolio). This table also reports the monthly Sharpe ratios for the efficient skewness/semivariance portfolios referred in Section 3.4.2: the maximum skewness per semivariance ratio portfolio (w_{SSR}), the maximum Sharpe ratio portfolio (w_{SR}) and the maximum Sortino ratio portfolio (w_{SOR}). The computation of the semivariance is carried out using three different values for the target return: the maximum Sharpe ratio portfolio return (R_{ms}), the minimum variance portfolio return (R_{mv}) and the equally weighted portfolio return (R_{ew}). In parenthesis are the bootstrap p-values of the difference between the Sharpe ratio of each efficient skewness/semivariance portfolio from those of the benchmarks: from the ms portfolio, from the mv portfolio and from the ew portfolio; respectively in the first, second and third parenthesis. These p-values are computed according the Ledoit and Wolf (Ledoit and Wolf, 2008) methodology.

In the case of the FF10 dataset, for all the target returns (R_{ms}, R_{mv} and R_{ew}), the efficient skewness/semivariance portfolios w_{SR} and w_{SOR} have a higher Sharpe ratio than two (the *ms* portfolio and the *ew* portfolio) of the three benchmark portfolios. The difference between the Sharpe ratio of the efficient skewness/semivariance portfolio w_{SR} and the benchmark *ew* portfolio is always statistically significant (at the 5% level).

We do not observe statistically significant differences, between the Sharpe ratios of the efficient skewness/semivariance portfolios and the benchmark portfolios, for the SOP25 dataset. However, we can see that when the semivariance is computed using as a target return R_{ms}, the efficient skewness/semivariance portfolio w_{SOR} has a higher Sharpe ratio than two (the *ms* portfolio and the *ew* portfolio) of the three benchmark portfolios.

Finally, for the BMOP25 dataset, independently of the target return used in the computation of the semivariance, the efficient skewness/semivariance portfolios w_{SR} and w_{SOR} have a higher Sharpe ratio than all the benchmark portfolios. The differences between the Sharpe ratios of these two efficient skewness/semivariance portfolios (w_{SR} and w_{SOR}) and the benchmark *ew* portfolio are always statistically significant (at the 1% level).

These results show that the efficient skewness/semivariance portfolios are consistently competitive, and often superior, comparatively to the benchmark portfolios. The efficient skewness/semivariance portfolios consistently outperform the *ew* benchmark portfolio. To achieve a higher Sharpe ratio, this analysis suggests that the investor should choose a target return (for the computation of the semivariance) according to the specific nature of the data.

Performance measured by the Sortino ratio

We computed the OOS Sortino ratio, defined as the sample mean of OOS excess returns (over the target return R), $\widehat{m}_R$, divided by their sample standard semideviation, $\widehat{\sigma}_R$:

$$\widehat{\text{SOR}} = \frac{\widehat{m}_R}{\widehat{\sigma}_R}. \tag{3.15}$$

Then, we computed the bootstrap p-values of the difference between the Sortino ratio of each efficient skewness/semivariance portfolio from those of the benchmarks. Since none of the differences were statistically significant, we decided not to report these results here. This data present negative excess returns and in order to achieve a correct rank of the considered portfolios, we modified the denominator according with the methodology proposed by Israelsen (2005). The refined Sortino ratio is computed as

$$\widehat{\text{SOR}}_{\text{ref}} = \frac{\widehat{m}_R}{\widehat{\sigma}_R^{\widehat{m}_R/abs(\widehat{m}_R)}}, \tag{3.16}$$

where $abs(\cdot)$ is the absolute value function.

Table 3.6 reports the refined Sortino ratios when we choose as a target return for the computation of the semivariance, the maximum Sharpe ratio portfolio return (R_{ms}). We can see that for the SBM6 dataset, the efficient skewness/semivariance portfolios w_{SSR} and w_{SOR} have a higher refined Sortino ratio than all the benchmark portfolios. In the case of the FF10 dataset, the efficient skewness/semivariance portfolios have a higher refined Sortino ratio than one (the *mv* portfolio) of the benchmark portfolios. The efficient skewness/semivariance portfolio w_{SOR}, in the SOP25 dataset, has a higher refined Sortino ratio than two (the *mv* portfolio and the *ew* portfolio) of the three benchmark portfolios. For the BMOP25 dataset, the efficient skewness/semivariance portfolio w_{SR} has the highest refined Sortino ratio among all the portfolios.

Table 3.6: Portfolio refined Sortino ratios for the target return R_{ms}

Strategy	SBM6	#Rank	FF10	#Rank	SOP25	#Rank	BMOP25	#Rank
ms portfolio	-1,850.8734	3	26,295.0701	1	-1,797.3958	1	-1,385.5948	2
mv portfolio	-2,345.3327	5	-386.2736	6	-1,985.5274	6	-2,121.3854	5
ew portfolio	-2,586.1178	6	-202.7054	2	-1,883.3336	3	-2,762.0948	6
Efficient skewness/semivariance portfolios								
w_{SSR}	-1,681.2084	1	-269.2545	4	-1,916.3469	4	-1,792.7265	4
w_{SR}	-2,276.6401	4	-203.2169	3	-1,960.5951	5	-1,261.9732	1
w_{SOR}	-1,748.2885	2	-281.5840	5	-1,831.0087	2	-1,428.1312	3

This table reports, for each dataset, the monthly refined Sortino ratios using as benchmark portfolios the maximum Sharpe ratio portfolio (ms portfolio), the minimum variance portfolio (mv portfolio) and the equally weighted portfolio (ew portfolio). All the refined Sortino ratios values are multiplied by a factor of 10^7. *The target return used in the computation of the semivariance is the maximum Sharpe ratio portfolio return* (R_{ms}). *This table also reports the monthly refined Sortino ratios for the efficient skewness/semivariance portfolios referred in Section 3.4.2: the maximum skewness per semivariance ratio portfolio* (w_{SSR}), *the maximum Sharpe ratio portfolio* (w_{SR}) *and the maximum Sortino ratio portfolio* (w_{SOR}). *The correct rank of each portfolio according to the refined Sortino ratios is also reported.*

In Table 3.7 we can find the refined Sortino ratios for the case in which the target return for the computation of the semivariance is the minimum variance portfolio return (R_{mv}). For the SBM6 dataset, the efficient skewness/semivariance portfolio w_{SOR} has the highest refined Sortino ratio among all the portfolios. In the cases of the FF10 and SOP25 datasets, the efficient skewness/semivariance portfolios have a higher refined Sortino ratio than one (the *mv* portfolio) of the benchmark portfolios. For the BMOP25 dataset, the efficient skewness/semivariance portfolio w_{SR} has the highest refined Sortino ratio among all the portfolios.

Finally, Table 3.8 presents the refined Sortino ratios for the case in which the target return is the *ew* portfolio return (R_{ew}). The efficient skewness/semivariance portfolio w_{SOR} has the highest refined Sortino ratio, among all the portfolios, in two cases (for the SBM6 dataset and for the BMOP25 dataset). For the FF10 dataset, the efficient skewness/semivariance portfolio w_{SR} has a higher refined Sortino ratio than two (the *mv* and the *ew* portfolios) of the three benchmark portfolios. In the case of the SOP25 dataset, the efficient skewness/semivariance portfolios have a higher refined Sortino ratio than one (the *mv* portfolio) of the benchmark portfolios.

Once again, these results suggests the robustness of the efficiency provided by the skewness/semivariance model.

Table 3.7: Portfolio refined Sortino ratios for the target return R_{mv}

Strategy	SBM6	#Rank	FF10	#Rank	SOP25	#Rank	BMOP25	#Rank
ms portfolio	624,615.7422	2	128,181.0851	1	-48.4189	2	479,777.6342	2
mv portfolio	256,858.4995	6	-293.0277	6	-375.1476	6	-109.7444	5
ew portfolio	236,771.0943	5	-52.1048	2	-2.1300	1	-366.0381	6
Efficient skewness/semivariance portfolios								
w_{SSR}	607,874.5739	3	-202.6882	5	-202.9211	4	265,776.7251	4
w_{SR}	355,161.3260	4	-133.4656	3	-252.2228	5	494,904.0660	1
w_{SOR}	665,238.0489	1	-182.6976	4	-181.3330	3	479,179.7536	3

This table reports, for each dataset, the monthly refined Sortino ratios using as benchmark portfolios the maximum Sharpe ratio portfolio (ms portfolio), the minimum variance portfolio (mv portfolio) and the equally weighted portfolio (ew portfolio). All the refined Sortino ratios values are multiplied by a factor of 10^7. The target return used in the computation of the semivariance is the minimum variance portfolio return (R_{mv}). This table also reports the monthly refined Sortino ratios for the efficient skewness/semivariance portfolios referred in Section 3.4.2: the maximum skewness per semivariance ratio portfolio (w_{SSR}), the maximum Sharpe ratio portfolio (w_{SR}) and the maximum Sortino ratio portfolio (w_{SOR}). The correct rank of each portfolio according to the refined Sortino ratios is also reported.

Table 3.8: Portfolio refined Sortino ratios for the target return R_{ew}

Strategy	SBM6	#Rank	FF10	#Rank	SOP25	#Rank	BMOP25	#Rank
ms portfolio	4,624.7267	2	-366.1145	1	-636.8731	2	296,765.0953	3
mv portfolio	-551.1163	5	-705.2135	3	-914.0883	6	-368.7736	5
ew portfolio	-602.2495	6	-711.9219	4	-636.2648	1	-676.9208	6
Efficient skewness/semivariance portfolios								
w_{SSR}	1,629.8212	3	-739.1504	5	-825.3411	4	61,942.1459	4
w_{SR}	-405.5356	4	-589.4037	2	-850.8136	5	335,361.5320	2
w_{SOR}	51,791.2377	1	-759.5383	6	-739.1028	3	338,761.2812	1

This table reports, for each dataset, the monthly refined Sortino ratios using as benchmark portfolios the maximum Sharpe ratio portfolio (ms portfolio), the minimum variance portfolio (mv portfolio) and the equally weighted portfolio (ew portfolio). All the refined Sortino ratios values are multiplied by a factor of 10^7. The target return used in the computation of the semivariance is the ew portfolio return (R_{ew}). This table also reports the monthly refined Sortino ratios for the efficient skewness/semivariance portfolios referred in Section 3.4.2: the maximum skewness per semivariance ratio portfolio (w_{SSR}), the maximum Sharpe ratio portfolio (w_{SR}) and the maximum Sortino ratio portfolio (w_{SOR}). The correct rank of each portfolio according to the refined Sortino ratios is also reported.

3.5 Conclusions

In this chapter we have proposed a direct analysis of the efficient tradeoff between skewness and semivariance through a skewness/semivariance biobjective optimization problem. We computed skewness as a third moment tensor and overcame the endogeneity problem of the cosemivariance matrix using a derivative-free algorithm. To the best of our knowledge, this is the first time that such an algorithm is used in this context. The solver chosen for solving the skewness/semivariance biobjective optimization problem is based on DMS. Direct-search methods based on polling are known to be extremely robust due to their directional properties (Conn et al., 2009). In fact, we have observed its robustness in four empirical datasets collected from the Fama/French data library, since DMS was capable of determining IS the Pareto frontier for the biobjective skewness/semivariance problem, using three different target returns for the computation of the semivariance.

In addition, we performed an extensive OOS analysis. The results showed that the efficient skewness/semivariance portfolios are consistently competitive when compared with the benchmark portfolios, in terms of OOS Sharpe ratio. A surprising fact was that at least one of the three chosen efficient skewness/semivariance portfolios, consistently outperforms the *ew* portfolio in terms of OOS Sharpe ratio. The efficient skewness/semivariance portfolios, also exhibited a consistently good performance in terms of skewness per semivariance ratio and Sortino ratio, which suggests the robustness of the efficiency provided by the skewness/semivariance model.

In the empirical analysis performed here we have just used three different target returns for the computation of the semivariance. Other choices could have been made (for example, the risk-free return or an index return). It is clear that the right choice of this parameter depends on the nature of the data and has a profound impact on the results. Within the skewness/semivariance model, the investor has the freedom to choose this parameter according to any criteria that may suit her preferences. Besides, we only evaluate the performance of three efficient skewness/semivariance portfolios (chosen according to three different criteria). Other efficient portfolios could have been evaluated.

The only constraint that we considered in the proposed model was the absence of short selling, but the proposed skewness/semivariance model could readily incorporate other constraints aiming to improve portfolio stability and/or performance. For instance, we could consider turnover constraints in order to control the transaction costs or introduce constraints to explore information about the cross-sectional characteristics of securities, as in Brandt et al. (2009).

Chapter 4

Portfolio management with higher moments: The cardinality impact

4.1 Introduction

With the aim of improving the performance of the classical MV model, several modifications or alternative strategies have been proposed in the literature. One well-known modification consists in imposing a cardinality constraint (a constraint that limits the number of stocks in the portfolio) to the classical MV optimization model, leading to the cardinality constrained MV model. This model can be seen as a mixed-integer quadratic problem that is no longer solved in polynomial time, therefore most studies have focused on providing efficient algorithms to solve this problem. These algorithms vary from exact algorithms (Bienstock, 1996; Vielma et al., 2008; Bertsimas and Shioda, 2009) to heuristics (Chang et al., 2000b; Fieldsend et al., 2004; Cesarone et al., 2009; Anagnostopoulos and Mamanis, 2011; Woodside-Oriakhi et al., 2011; Brito and Vicente, 2014). With the aim of promoting regularization of ill conditioning, DeMiguel et al. (2009a) constrained the Markowitz classical model by imposing a bound on the l_1-norm of the vector of portfolio weights. Since the l_1-norm is the exact convex relaxation of the l_0-norm (cardinality), DeMiguel et al. (2009a) indirectly studied the effect of cardinality in the classical MV framework. A similar idea was explored in Brodie et al. (2009).

Cardinality is a well-known measure of portfolio diversification. When cardinality increases, the total risk of a portfolio decreases gradually, until a certain level of risk is achieved. After that level, increases in cardinality only produce negligible reductions in the risk level (see, e.g., Evans and Archer, 1968; Statman, 1987; Benartzi and Thaler, 2001). Therefore, the cardinality impact on portfolio performance is a very important issue for both individual and institutional investors.

This chapter focus on the analysis of the cardinality impact on the portfolio performance, in the mean-variance (MV), mean-variance-skewness (MVS) and mean-variance-skewness-kurtosis (MVSK) frameworks. Therefore, the analysis is conducted not only on the impact of the cardinality constraint in each framework but also on the benefits of considering higher moments in portfolio management. For the reasons explained in Section 2.3, and following Brandt et al. (2009), in this study we assume that the investor maximizes her expected utility, characterized by CRRA preferences. This is a common assumption in the related literature (see, e.g., Aït-Sahalia and Brandt, 2001). Nevertheless, we point out that the presented procedures are quite flexible, and may be applied to any other type of utility function. For each framework, we propose a biobjective model that allows the investor to directly analyze the efficient tradeoff between expected utility and cardinality. Given the nonsmoothness of the cardinality function and inspired in Brito and Vicente (2014), we have decided to solve each biobjective model using a derivative-free solver based on DMS (see Appendix B).

The contribution of this chapter is twofold. First, this study extends the analysis of the cardinality impact on the portfolio performance from the standard MV framework to the MVS and MVSK frameworks. Second, we analyze the performance gains (in terms of certainty equivalent and Sharpe ratio) when considering higher moments (skewness and kurtosis) at different cardinality levels. The certainty equivalent, being a measure that takes into account the expected utility, is therefore a more suitable performance measure than the Sharpe ratio when comparing the three frameworks. Although one needs to be careful when using the Sharpe ratio as a performance measure, especially in the MVS and MVSK frameworks (the ratio does not account for skewness or kurtosis), this is one of the most widely used performance measures in the literature, even in non MV settings (see, e.g., DeMiguel et al., 2009b). Although the Sharpe ratio is intrinsically related to the MV framework, it is interesting to see how the MVS and MVSK frameworks perform in this dimension in comparison with the MV framework (even if the existence of performance gains, in terms of Sharpe ratio, of the MVS and MVSK frameworks against the MV framework should be hard to get).

The empirical analysis is conducted on a dataset from the Portuguese Stock Market Index (PSI 20 Index). The daily data is collected from Thomson Reuters Datastream®. First, we compute the expected utility/cardinality efficient frontiers for each one of the three frameworks and analyze the IS cardinality impact using as performance measures the certainty equivalent and the Sharpe ratio. Second, we perform an OOS analysis using the certainty equivalent, the Sharpe ratio, the turnover and the Sharpe ratio of returns net of transaction costs. The results for each efficient expected utility/cardinality portfolio in each of the three frameworks are reported.

The results suggest that there are no performance gains IS (in terms of certainty equivalent) when higher moments are considered, while in some cases the inclusion of higher moments in the portfolio management framework leads to some significant gains in terms of OOS certainty equivalent and Sharpe ratio.

As a kind of robustness check, we present some results for the OOS certainty equivalent, Sharpe ratio, turnover and the net Sharpe ratio for a larger dataset from the EURO STOXX 50 Index. The observed results are consistent with those for the PSI 20 Index.

The remainder of the chapter proceeds as follows. In Section 4.2 we describe our approach and the proposed models. The empirical application is presented in Section 4.3. We conclude in Section 4.4.

4.2 Methodology

4.2.1 The expected utility/cardinality biobjective model

We assume that the investor wants to choose a portfolio taking into account two criteria: an expected utility that considers several moments of the portfolio return and the cardinality of the portfolio.

Let us consider the standard investor's problem (see Problem (2.17)), with general lower and upper bounds on the weights, that is,

$$\begin{aligned} \max_{w_t \in \mathbb{R}^N} \quad & E_t\left[u\left(r_{p,t+1}\right)\right] = E_t\left[u\left(\sum_{i=1}^{N} w_{i,t} r_{i,t+1}\right)\right] \\ \text{subject to} \quad & 1_N^\top w_t = 1, \\ & LO_{i,t} \leq w_{i,t} \leq UP_{i,t}, \; i = 1, \ldots, N, \end{aligned} \tag{4.1}$$

where $LO_{i,t}$ and $UP_{i,t}$ represent, respectively, lower and upper bounds on the weights.

Adding to Problem (4.1) a constraint that limits the number of active positions on the portfolio (a cardinality constraint), we have

$$
\begin{aligned}
\max_{w_t \in \mathbb{R}^N} \quad & E_t\left[u\left(r_{p,t+1}\right)\right] \;=\; E_t\left[u\left(\sum_{i=1}^{N} w_{i,t} r_{i,t+1}\right)\right] \\
\text{subject to} \quad & \text{card}(w_t) \;\leq\; C \\
& 1_N^\top w_t \;=\; 1, \\
& LO_{i,t} \;\leq\; w_{i,t} \;\leq\; UP_{i,t}, \; i = 1, \ldots, N,
\end{aligned}
\tag{4.2}
$$

where $\text{card}(w_t) = |\{i \in \{1, ..., N\} : w_{i,t} \neq 0\}|$ and $C \in \{1, \ldots, N\}$. Despite $\text{card}(\cdot)$ being often called, in the literature, the l_0-norm, it is not a norm due to the fact that it does not satisfy the homogeneity property.

Motivated by Brito and Vicente (2014), here we suggest that cardinality should be included as a second objective in the problem, instead of a constraint. Thus, we reformulate Problem (4.2) as

$$
\begin{aligned}
\max_{w_t \in \mathbb{R}^N} \quad & E_t\left[u\left(r_{p,t+1}\right)\right] \;=\; E_t\left[u\left(\sum_{i=1}^{N} w_{i,t} r_{i,t+1}\right)\right] \\
\min_{w_t \in \mathbb{R}^N} \quad & \text{card}(w_t) \\
\text{subject to} \quad & 1_N^\top w_t \;=\; 1, \\
& LO_{i,t} \;\leq\; w_{i,t} \;\leq\; UP_{i,t}, \; i = 1, \ldots, N.
\end{aligned}
\tag{4.3}
$$

The solution of Problem (4.3) is given by a Pareto frontier, i.e., as an efficient expected utility/cardinality frontier. Efficient expected utility/cardinality portfolios can be seen as those which have the maximum expected utility among all that have at most a certain level of cardinality. The investor can thus directly analyze the efficient tradeoff between expected utility and cardinality. The biobjective Problem (4.3) has two objective functions, a linear constraint and $2N$ inequality constraints. The first objective, $E_t\left[u\left(r_{p,t+1}\right)\right]$, is nonlinear but smooth. However, the second objective, $\text{card}(w_t)$, is a piecewise linear discontinuous function, and consequently it is nonlinear and nonsmooth. Since derivative-free algorithms are applicable to black-box type functions (see Appendix B), they are suitable to deal with the cardinality objective function. According with Custódio et al. (2011), the `dms` solver has the best performance among all derivative-free solvers for MOO optimization. We have thus decided to solve Problem (4.3) using `dms` (see Appendix B).

4.2.2 The MV, MVS and MVSK expected utility/cardinality efficient frontiers

If the investor with CRRA preferences (see Equation (2.19)) considers the second order Taylor expansion of the expected utility, $E_t\left[u\left(r_{p,t+1}\right)\right]$, around the expected return of the portfolio, $E_t\left(r_{p,t+1}\right)$, the second member of Equation (2.22) is truncated at the second term. Hence, in order to investigate the cardinality impact, the following model can be solved when the investor has a MV expected utility

$$\begin{aligned} \max_{w_t \in \mathbb{R}^N} \quad & \theta_1\left[E_t\left(r_{p,t+1}\right)\right] - \theta_2\left[E_t\left(r_{p,t+1}\right)\right] v_t\left(r_{p,t+1}\right) \\ \min_{w_t \in \mathbb{R}^N} \quad & \text{card}(w_t) \\ \text{subject to} \quad & 1_N^\top w_t \ = \ 1, \\ & LO_{i,t} \ \leq \ w_{i,t} \ \leq \ UP_{i,t}, \ i = 1, \ldots, N. \end{aligned} \tag{4.4}$$

By solving Problem (4.4), we identify a MV expected utility/cardinality efficient frontier. A portfolio in this frontier is such that there exists no other feasible one which simultaneously presents a higher expected MV utility and a lower cardinality. Given such an efficient frontier and a cardinality target, an investor may directly find the answers to what is the optimal (highest) expected utility level that can be attained while meeting the cardinality target and what are the portfolios leading to such an expected utility level. The investor can thus directly examine the efficient tradeoff between the MV expected utility and cardinality.

If the investor, besides the mean and variance, also considers the third moment (the second member of Equation (2.22) is truncated at the third term), the following model can be solved

$$\begin{aligned} \max_{w_t \in \mathbb{R}^N} \quad & \theta_1\left[E_t\left(r_{p,t+1}\right)\right] - \theta_2\left[E_t\left(r_{p,t+1}\right)\right] v_t\left(r_{p,t+1}\right) + \theta_3\left[E_t\left(r_{p,t+1}\right)\right] s_t\left(r_{p,t+1}\right) \\ \min_{w_t \in \mathbb{R}^N} \quad & \text{card}(w_t) \\ \text{subject to} \quad & 1_N^\top w_t \ = \ 1, \\ & LO_{i,t} \ \leq \ w_{i,t} \ \leq \ UP_{i,t}, \ i = 1, \ldots, N. \end{aligned} \tag{4.5}$$

By solving Problem (4.5), we identify a MVS expected utility/cardinality efficient frontier. The investor can thus directly analyze the efficient tradeoff between the MVS expected utility and cardinality.

When the investor considers all the moments present in Equation (2.22), the model to be solved is given by

$$\begin{aligned}
\max_{w_t \in \mathbb{R}^N} \quad & \theta_1 \left[E_t\left(r_{p,t+1}\right)\right] - \theta_2 \left[E_t\left(r_{p,t+1}\right)\right] v_t\left(r_{p,t+1}\right) \\
& +\theta_3 \left[E_t\left(r_{p,t+1}\right)\right] s_t\left(r_{p,t+1}\right) - \theta_4 \left[E_t\left(r_{p,t+1}\right)\right] k_t\left(r_{p,t+1}\right) \\
\min_{w_t \in \mathbb{R}^N} \quad & \mathrm{card}(w_t) \\
\text{subject to} \quad & 1_N^\top w_t \;=\; 1, \\
& LO_{i,t} \;\leq\; w_{i,t} \;\leq\; UP_{i,t},\; i = 1, \ldots, N.
\end{aligned} \tag{4.6}$$

By solving Problem (4.6), we identify a MVSK expected utility/cardinality efficient frontier. The investor can thus directly analyze the efficient tradeoff between the MVSK expected utility and cardinality.

4.3 Empirical performance

We tested the proposed methodology in one dataset based on the composition of the Portuguese Stock Market Index (PSI 20 Index). We collected daily data from Thomson Reuters Datastream®, for the time window from July 2007 to June 2014 (seven years). We chose 20 stocks of companies that belonged to the PSI 20 Index at least once and that were traded during all this time (some stocks changed designation, but all stocks were quoted during the entire period under analysis). The composition of this dataset is given in Table 4.1.

Table 4.1: The PSI 20 dataset

List of stocks	
ALTRI SGPS	MARTIFER
BANCO BPI	MEDIA CAPITAL
BANCO COMR.PORTUGUES 'R'	MOTA ENGIL SGPS
BANCO ESPIRITO SANTO	NOS SGPS
COFINA	NOVABASE
EDP ENERGIAS DE PORTUGAL	PHAROL SGPS
GI.GLB.INTEL.TECHS.SGPS	PORTUCEL EMPRESA
IMPRESA SGPS	SEMAPA
INAPA	SONAE INDUSTRIA SGPS
JERONIMO MARTINS	SONAE SGPS

This table lists the composition of the PSI 20 dataset used in the empirical work. Daily closing prices, from July 2007 to June 2014, of these 20 stocks were collected from the Thomson Reuters Datastream®.

Some descriptive statistics of the returns, for the overall sample period, are reported in Table 4.2.

Table 4.2: Descriptive statistics for the PSI 20 dataset

Number of days	1826
Minimum	-0.1573
Median	-0.0001
Maximum	0.2584
Mean	-0.0002
Variance	0.0009
Skewness	0.8187
Kurtosis	15.6667

This table reports some descriptive statistics for the PSI 20 dataset (the composition of the dataset is reported in Table 4.1). The reported values of skewness and kurtosis concern to the third and fourth standardized moments, respectively. These statistics are averaged cross-sectionally, i.e., they are computed for each stock and then the arithmetic mean is taken.

During the period under scrutiny the discrete daily returns of the PSI 20 stocks presented, on average, a negative but near zero mean, positive skewness and about five times above normal kurtosis.

We present the efficient expected utility/cardinality portfolios, with the expected utility based on the MV, MVS and MVSK frameworks, in Section 4.3.1. In Section 4.3.2, based on a rolling window approach (see Figure 2.5), we analyze the OOS performance of these efficient portfolios. Both IS and OOS analysis were conducted assuming that the investor has CRRA preferences and a parameter of relative risk aversion equal to five (this is a common assumption, see, e.g., Brandt et al., 2009).

4.3.1 IS performance

We solved Problem (4.4), Problem (4.5) and Problem (4.6) for the general non-boundary case (i.e., $LO_{i,t} = -\infty$ and $UP_{i,t} = +\infty$) using the `dms` solver[7] (see Appendix B). Other choices for the boundaries could be made; for instance, if short selling is undesirable then $LO_{i,t} \geq 0$ and consequently $UP_{i,t} \leq 1$. We have decided to allow for short selling in order to make the models as general as possible and to avoid the possibility

[7]Following Chang et al. (2000b) and Jobst et al. (2010), in practice the true cardinality is approximated by introducing a tolerance $\varepsilon = 1\%$, such that $\text{card}(w_t) = \sum_{i=1}^{N} \mathbb{1}_{\{|w_{i,t}|>\varepsilon\}}$, where $\mathbb{1}$ represents the indicator function.

of getting Pareto frontiers with few feasible cardinality levels. Note that the greater the restriction on the portfolio weights, the more likely it is to obtain Pareto frontiers without all the feasible cardinality levels[8] (these "missing points" correspond to strictly dominated points). Allowing for short selling in each model does not make them less realistic, since most markets around the world allow short positions and, particularly, the Portuguese Securities Market Commission (CMVM) allows to perform short selling in each one of the constituents of the PSI 20 Index.

In the IS analysis we used the overall sample period, from July 2007 to June 2014 (seven years). The efficient expected utility/cardinality frontiers[9], for the MV, MVS and MVSK frameworks, are plotted in Figure 4.1.

DMS aims to approximate the true Pareto frontier. Theoretically it is only possible to prove that there is a limit point in a stationary form of the Pareto frontier, as no aggregation or scalarization technique is incorporated (see Custódio et al., 2011, for further details). However, some empirical applications have shown that this method has a very good performance, even when applied to problems with discontinuous and non-convex Pareto frontiers (Custódio et al., 2011). In our application, we can observe that DMS was effective, in the sense that it allowed us to obtain frontiers for the three frameworks, and we have all the reasons to believe that they are good approximations to the true Pareto frontiers. From the analysis of these frontiers the investor can thus directly analyse the efficient tradeoff (in each framework) and choose (according to some pre-established criteria), in the expected utility/cardinality space, the portfolio that best fits her preferences.

Table 4.3 reports the certainty equivalent return for the MV, MVS and MVSK frameworks. The certainty equivalent return, $\widehat{CE}$, is defined as

$$u(\widehat{CE}) = E_t\left[u(r_{p,t+1})\right], \tag{4.7}$$

and can be interpreted as the risk-free rate that an investor is willing to accept in order to give up a particular risky investment.

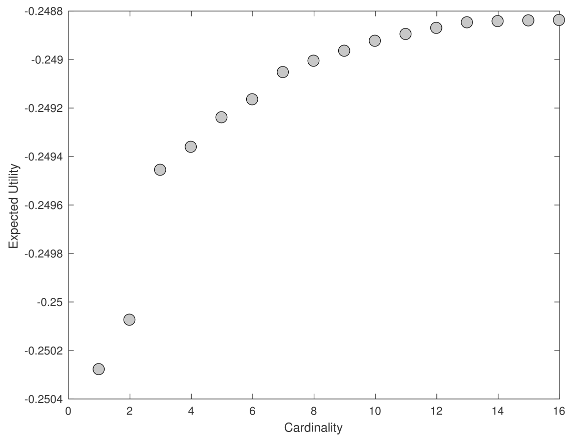
Tables 4.4, 4.5 and 4.6 report some descriptive statistics (the IS mean, $\widehat{m}$, standard deviation, $\widehat{\sigma}$, standardized skewness, $\widehat{s}_{stand}$, standardized kurtosis, $\widehat{k}_{stand}$, expected utility, $\widehat{u}$, and Sharpe ratio[10], $\widehat{SR}$) for the MV, MVS and MVSK frameworks, respectively. These tables also report the ratio between the IS Sharpe ratio of each efficient portfolio

[8] Theoretically the solutions of Problem (4.4), Problem (4.5) and Problem (4.6) can be a single point that strictly dominates all the others.

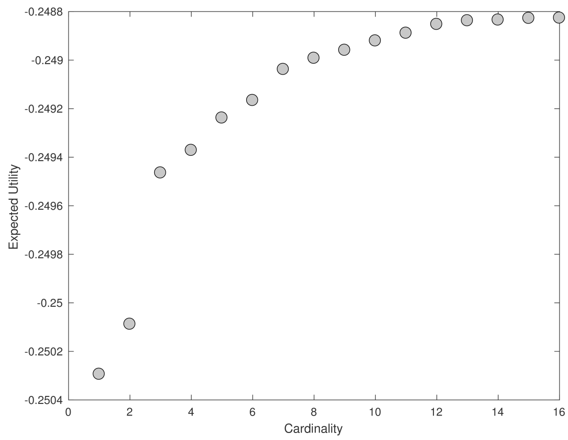
[9] Note that, with the 1% threshold for the cardinality computation, the maximum cardinality level that led to nondominated solutions was 16.

[10] Without loss of generality, in this chapter we set $r_t^f=0$.

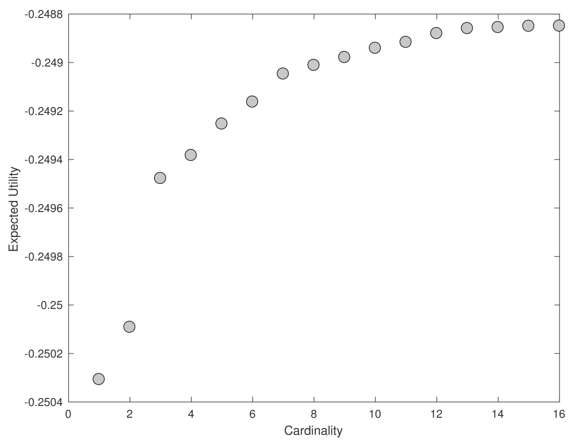
Figure 4.1: Efficient expected utility/cardinality frontiers, for the MV, MVS and MVSK frameworks



Efficient MV expected utility/cardinality portfolios



Efficient MVS expected utility/cardinality portfolios



Efficient MVSK expected utility/cardinality portfolios

This figure displays the efficient expected utility/cardinality frontiers. The vertical axis represents the expected utility (MV, MVS and MVSK expected utility, respectively). The horizontal axis corresponds to the cardinality.

Table 4.3: IS certainty equivalent return

Efficient portfolios			
Number of stocks	MV	MVS	MVSK
1	-0.2499	-0.2637	-0.2762
2	-0.0171	-0.0245	-0.0280
3	0.5494	0.5498	0.5365
4	0.6475	0.6437	0.6340
5	0.7735	0.7814	0.7661
6	0.8469	0.8516	0.8536
7	0.9543	0.9694	0.9638
8	0.9970	1.0127	0.9936
9	1.0596	1.0457	1.0256
10	1.0976	1.1020	1.0795
11	1.1279	1.1319	1.1074
12	1.1626	1.1694	1.1463
13	1.1815	1.1846	1.1623
14	1.1859	1.1889	1.1665
15	1.1905	1.1930	1.1709
16	1.1956	1.1989	1.1760

This table lists the IS certainty equivalent return ($\widehat{CE}$) of the efficient expected utility/cardinality portfolios. MV refers to the efficient MV expected utility/cardinality portfolios, MVS refers to the efficient MVS expected utility/cardinality portfolios and MVSK refers to the efficient MVSK expected utility/cardinality portfolios. All the certainty equivalent values are multiplied by a factor of 10^3.

and the IS Sharpe ratio of the efficient portfolio with cardinality 16, $\widehat{SR}/\widehat{SR}_{16}$ (corresponding to the portfolio with the maximum cardinality level). This ratio can be seen as a measure of the impact of portfolio diversification on the Sharpe ratio.

Table 4.4: IS analysis of the efficient MV expected utility/cardinality portfolios

Efficient portfolios							
Number of stocks	$\widehat{m}$	$\widehat{\sigma}$	$\widehat{s}_{stand}$	$\widehat{k}_{stand}$	$\widehat{u}$	$\widehat{SR}$	$\widehat{SR}/\widehat{SR}_{16}$
1	0.0008	0.0204	-0.3440	8.4078	-0.2503	0.0385	0.3485 (0.0038)
2	0.0007	0.0164	-0.3867	7.5859	-0.2500	0.0398	0.3593 (0.0036)
3	0.0019	0.0233	-0.1037	5.2381	-0.2495	0.0816	0.7373 (0.0486)
4	0.0019	0.0225	-0.1122	5.3278	-0.2494	0.0847	0.7659 (0.0556)
5	0.0023	0.0248	-0.0292	4.8302	-0.2492	0.0929	0.8394 (0.1356)
6	0.0023	0.0238	-0.0045	5.3832	-0.2492	0.0948	0.8566 (0.1480)
7	0.0025	0.0252	0.1553	6.1148	-0.2490	0.1005	0.9081 (0.2685)
8	0.0025	0.0243	0.1966	6.3847	-0.2490	0.1014	0.9164 (0.2531)
9	0.0027	0.0256	0.0374	5.8750	-0.2489	0.1049	0.9484 (0.4015)
10	0.0028	0.0259	0.0250	5.7805	-0.2489	0.1067	0.9644 (0.4727)
11	0.0029	0.0265	0.0159	5.7081	-0.2489	0.1083	0.9792 (0.6155)
12	0.0029	0.0263	0.0460	5.5404	-0.2488	0.1094	0.9893 (0.7229)
13	0.0029	0.0261	0.0054	5.5654	-0.2488	0.1100	0.9946 (0.7778)
14	0.0029	0.0261	0.0037	5.5670	-0.2488	0.1102	0.9963 (0.8238)
15	0.0029	0.0262	0.0105	5.6330	-0.2488	0.1105	0.9986 (0.9426)
16	0.0029	0.0261	0.0144	5.4855	-0.2488	0.1106	1

This table reports the IS mean ($\widehat{m}$), standard deviation ($\widehat{\sigma}$), standardized skewness ($\widehat{s}_{stand}$), standardized kurtosis ($\widehat{k}_{stand}$), expected utility ($\widehat{u}$) and the IS Sharpe ratio ($\widehat{SR}$) of each efficient MV expected utility/cardinality portfolio. It also reports the ratio, $\widehat{SR}/\widehat{SR}_{16}$, between the IS Sharpe ratio of each efficient portfolio and the IS Sharpe ratio of the efficient portfolio with cardinality 16 (the highest Sharpe ratio portfolio). The Ledoit and Wolf (2008) bootstrap p-values for the statistical significance of the difference between the IS Sharpe ratio of the efficient portfolio with cardinality 16 and the respective efficient portfolio are presented in parenthesis.

One interesting result is that, in each of the three frameworks, the IS certainty equivalent and Sharpe ratio increases gradually with cardinality, until the maximum cardinality level of 16 is achieved. Moreover, we computed the bootstrap p-values for the statistical significance of the difference between the IS Sharpe ratio of the efficient portfolio with cardinality 16 and each other efficient portfolio. These p-values were computed according the Ledoit and Wolf (2008) robust methodology. For the MV framework these bootstrap p-values show that only the portfolios with cardinality 1, 2 and 3, consistently offer a Sharpe ratio significantly lower than the Sharpe ratio of the portfolio with cardinality 16. For the MVS and MVSK frameworks, this happens only for the portfolios with cardinality 1 and 2. This seems to suggest that, for this dataset, most of the diversification gains occur at cardinalities of 3 or 4.

Table 4.5: IS analysis of the efficient MVS expected utility/cardinality portfolios

	Efficient portfolios						
Number of stocks	$\widehat{m}$	$\widehat{\sigma}$	$\widehat{s}_{stand}$	$\widehat{k}_{stand}$	$\widehat{u}$	$\widehat{SR}$	$\widehat{SR}/\widehat{SR}_{16}$
1	0.0008	0.0204	-0.3441	8.4074	-0.2503	0.0385	0.3486 (0.0034)
2	0.0006	0.0163	-0.3851	7.5589	-0.2500	0.0396	0.3584 (0.0040)
3	0.0019	0.0232	-0.1021	5.2614	-0.2495	0.0817	0.7395 (0.0620)
4	0.0019	0.0224	-0.1103	5.3212	-0.2494	0.0847	0.7661 (0.0734)
5	0.0023	0.0246	-0.0158	4.8868	-0.2492	0.0929	0.8408 (0.1476)
6	0.0023	0.0239	-0.0035	5.3673	-0.2492	0.0950	0.8596 (0.1672)
7	0.0026	0.0254	0.2250	6.7244	-0.2490	0.1005	0.9095 (0.2589)
8	0.0025	0.0245	0.2437	6.8288	-0.2490	0.1015	0.9180 (0.2599)
9	0.0026	0.0251	0.2110	6.3837	-0.2490	0.1033	0.9349 (0.3155)
10	0.0028	0.0259	0.0857	6.1959	-0.2489	0.1066	0.9642 (0.4797)
11	0.0029	0.0265	0.0792	6.1399	-0.2489	0.1082	0.9790 (0.6107)
12	0.0029	0.0263	0.1165	6.0007	-0.2488	0.1093	0.9893 (0.7099)
13	0.0029	0.0260	0.0742	5.9680	-0.2488	0.1099	0.9942 (0.7754)
14	0.0029	0.0261	0.0722	5.9641	-0.2488	0.1101	0.9960 (0.8268)
15	0.0029	0.0261	0.0727	5.8957	-0.2488	0.1102	0.9974 (0.8518)
16	0.0029	0.0262	0.0832	5.9889	-0.2488	0.1105	1

This table reports the IS mean ($\widehat{m}$), standard deviation ($\widehat{\sigma}$), standardized skewness ($\widehat{s}_{stand}$), standardized kurtosis ($\widehat{k}_{stand}$), expected utility ($\widehat{u}$) and the IS Sharpe ratio ($\widehat{SR}$) of each efficient MVS expected utility/cardinality portfolio. It also reports the ratio, $\widehat{SR}/\widehat{SR}_{16}$, between the IS Sharpe ratio of each efficient portfolio and the IS Sharpe ratio of the efficient portfolio with cardinality 16 (the highest Sharpe ratio portfolio). The Ledoit and Wolf (2008) bootstrap p-values for the statistical significance of the difference between the IS Sharpe ratio of the efficient portfolio with cardinality 16 and the respective efficient portfolio are presented in parenthesis.

Table 4.6: IS analysis of the efficient MVSK expected utility/cardinality portfolios

	Efficient portfolios						
Number of stocks	$\widehat{m}$	$\widehat{\sigma}$	$\widehat{s}_{stand}$	$\widehat{k}_{stand}$	$\widehat{u}$	$\widehat{SR}$	$\widehat{SR}/\widehat{SR}_{16}$
1	0.0008	0.0204	-0.3441	8.4074	-0.2503	0.0385	0.3497 (0.0038)
2	0.0006	0.0162	-0.3841	7.5402	-0.2500	0.0395	0.3589 (0.0044)
3	0.0019	0.0232	-0.1020	5.2611	-0.2495	0.0817	0.7418 (0.0566)
4	0.0019	0.0222	-0.1086	5.3120	-0.2494	0.0846	0.7675 (0.0624)
5	0.0023	0.0245	-0.0176	4.8963	-0.2492	0.0928	0.8426 (0.1488)
6	0.0024	0.0247	0.0047	5.0890	-0.2491	0.0966	0.8763 (0.1948)
7	0.0025	0.0246	0.1855	6.0367	-0.2490	0.1005	0.9124 (0.2460)
8	0.0025	0.0243	0.2066	6.4703	-0.2490	0.1015	0.9209 (0.2759)
9	0.0026	0.0249	0.1721	6.0183	-0.2490	0.1033	0.9375 (0.3251)
10	0.0027	0.0254	0.0649	6.0057	-0.2489	0.1062	0.9638 (0.4711)
11	0.0028	0.0259	0.0559	5.9411	-0.2489	0.1078	0.9783 (0.6153)
12	0.0028	0.0257	0.0920	5.7790	-0.2489	0.1090	0.9891 (0.7141)
13	0.0028	0.0255	0.0524	5.7804	-0.2488	0.1096	0.9943 (0.7774)
14	0.0028	0.0255	0.0517	5.7863	-0.2488	0.1098	0.9960 (0.8246)
15	0.0028	0.0255	0.0523	5.7182	-0.2488	0.1099	0.9975 (0.8472)
16	0.0028	0.0257	0.0599	5.7839	-0.2488	0.1102	1

This table reports the IS mean ($\widehat{m}$), standard deviation ($\widehat{\sigma}$), standardized skewness ($\widehat{s}_{stand}$), standardized kurtosis ($\widehat{k}_{stand}$), expected utility ($\widehat{u}$) and the IS Sharpe ratio ($\widehat{SR}$) of each efficient MVSK expected utility/cardinality portfolio. It also reports the ratio, $\widehat{SR}/\widehat{SR}_{16}$, between the IS Sharpe ratio of each efficient portfolio and the IS Sharpe ratio of the efficient portfolio with cardinality 16 (the highest Sharpe ratio portfolio). The Ledoit and Wolf (2008) bootstrap p-values for the statistical significance of the difference between the IS Sharpe ratio of the efficient portfolio with cardinality 16 and the respective efficient portfolio are presented in parenthesis.

In these tables, it may seem puzzling that there is not a clear decreasing trend of the standard deviation of portfolio returns with the increase of cardinality. However, we must point out that we are, in fact, maximizing the expected utility for each level of cardinality. Therefore, with a marginal increase in the cardinality level, we may choose a completely different set of stocks (and associated weights) for which the benefits resulting from other moments of the return distribution (expected return and, when applicable, skewness and/or kurtosis) more than offset a possible increase in the standard deviation.

Results reported in Tables 4.3, 4.4, 4.5 and 4.6 point out that there are no significant differences between the three frameworks in terms of IS certainty equivalent and Sharpe ratio. For example, in each of the three frameworks the maximum certainty equivalent return is achieved at the maximum cardinality level and the values of $\widehat{CE}$ are quite similar (0.11956%, 0.11989% and 0.11760%, respectively). Also the maximum IS Sharpe ratio is found, in each of the three frameworks, at the maximum feasible cardinality level and with similar values for $\widehat{SR}$ (11.06%, 11.05% and 11.02%, respectively). These results suggest that, IS, the consideration of higher moments does not produce relevant gains (in terms of certainty equivalent) for an investor with CRRA preferences ($\gamma = 5$).

4.3.2 OOS performance

The OOS analysis relies on a rolling window approach (see Figure 2.5), with an estimation window of 1565 days. The initial estimation window begins in July 2007 and ends in June 2013. There are 261 evaluation periods (days) until June 2014.

Table 4.7 reports the OOS certainty equivalent return (computed according Equation (4.7)), $\widehat{CE}$, for each one of the three frameworks.

Tables 4.8, 4.9 and 4.10 report some OOS statistics (the OOS mean, $\widehat{m}$, standard deviation, $\widehat{\sigma}$, standardized skewness, $\widehat{s}_{stand}$, standardized kurtosis, $\widehat{k}_{stand}$, expected utility, $\widehat{u}$, and Sharpe ratio, $\widehat{SR}$) for the MV, MVS and MVSK frameworks, respectively. When the numerator of the Sharpe ratio ($\widehat{SR}$) is negative, it should be refined in order to achieve a correct rank of the portfolios. Once again we use the Israelsen (2005) methodology:

$$\widehat{SR}_{ref} = \frac{\widehat{m}}{\widehat{\sigma}^{\widehat{m}/abs(\widehat{m})}}, \tag{4.8}$$

where $abs(\cdot)$ is the absolute value function. Note that the refined Sharpe ratio ($\widehat{SR}_{ref}$) is equal to the Sharpe ratio ($\widehat{SR}$) when the numerator is nonnegative; in this case,

Table 4.7: OOS certainty equivalent return

Efficient portfolios			
Number of stocks	MV	MVS	MVSK
1	-0.3605	-0.1337	-0.2635
2	-0.2165	-0.2886	-0.5271
3	-1.3200	-1.2945	-1.1657
4	-1.4815	-1.4850	-1.6058
5	-0.5165	-0.7985	-1.1355
6	-1.2108	-1.0303	-1.0729
7	-0.9808	-1.3350	-1.0559
8	-1.5770	-1.2233	-1.2436
9	-1.3603	-1.0098	-1.1505
10	-0.7319	-1.0514	-0.7085
11	-0.8519	-0.9998	-0.9240
12	-0.8622	-0.8066	-0.5946
13	-0.8368	-0.7737	-0.6427
14	-0.7461	-0.7586	-0.5603
15	-0.7741	-0.7319	-0.6841
16	-0.7593	-0.7701	-0.7048

This table lists the OOS certainty equivalent return ($\widehat{CE}$) of the efficient expected utility/cardinality portfolios. MV refers to the efficient MV expected utility/cardinality portfolios, MVS refers to the efficient MVS expected utility/cardinality portfolios and MVSK refers to the efficient MVSK expected utility/cardinality portfolios. All the certainty equivalent values are multiplied by a factor of 10^3.

Equation (4.8) is equivalent to Equation (3.14).

Table 4.8: OOS analysis of the efficient MV expected utility/cardinality portfolios

	Efficient portfolios							
Number of stocks	$\widehat{m}$	$\widehat{\sigma}$	$\widehat{s}_{stand}$	$\widehat{k}_{stand}$	$\widehat{u}$	$\widehat{SR}$	$\widehat{SR}/\widehat{SR}_{16}$	$\widehat{SR}_{ref}$
1	0.0001	0.0137	-0.1952	4.1943	-0.2504	0.0084	0.1472 (0.5781)	0.0084
2	0.0002	0.0137	0.8597	8.7178	-0.2502	0.0177	0.3105 (0.5961)	0.0177
3	-0.0002	0.0209	-0.3217	5.2709	-0.2513	-0.0100	-0.1750 (0.0888)	-0.0000
4	-0.0001	0.0231	0.0797	4.9952	-0.2515	-0.0065	-0.1137 (0.0624)	-0.0000
5	0.0011	0.0249	-0.3433	4.3266	-0.2505	0.0430	0.7555 (0.6273)	0.0430
6	0.0006	0.0266	-0.4120	5.1481	-0.2512	0.0231	0.4060 (0.1562)	0.0231
7	0.0009	0.0272	-0.2345	3.8378	-0.2510	0.0329	0.5784 (0.1998)	0.0329
8	0.0004	0.0278	-0.3922	4.6299	-0.2516	0.0146	0.2569 (0.0282)	0.0146
9	0.0008	0.0287	-0.3905	4.4395	-0.2514	0.0266	0.4675 (0.0750)	0.0266
10	0.0015	0.0294	-0.3697	4.5972	-0.2507	0.0508	0.8928 (0.6575)	0.0508
11	0.0015	0.0301	-0.3385	4.3533	-0.2509	0.0490	0.8606 (0.4619)	0.0490
12	0.0016	0.0309	-0.4063	4.5166	-0.2509	0.0520	0.9129 (0.8708)	0.0520
13	0.0017	0.0312	-0.3999	4.5208	-0.2508	0.0538	0.9456 (0.5527)	0.0538
14	0.0018	0.0314	-0.3855	4.6060	-0.2507	0.0573	1.0064 (0.9100)	0.0573
15	0.0018	0.0314	-0.3909	4.5925	-0.2508	0.0563	0.9885 (0.6667)	0.0563
16	0.0018	0.0314	-0.4056	4.6114	-0.2508	0.0569	1	0.0569

This table reports the OOS mean ($\widehat{m}$), standard deviation ($\widehat{\sigma}$), standardized skewness ($\widehat{s}_{stand}$), standardized kurtosis ($\widehat{k}_{stand}$), expected utility ($\widehat{u}$) and the OOS Sharpe ratio ($\widehat{SR}$) of each efficient MV expected utility/cardinality portfolio. It also reports the ratio, $\widehat{SR}/\widehat{SR}_{16}$, between the OOS Sharpe ratio of each efficient portfolio and the OOS Sharpe ratio of the efficient portfolio with cardinality 16. The Ledoit and Wolf (2008) bootstrap p-values for the statistical significance of the difference between the IS Sharpe ratio of the efficient portfolio with cardinality 16 and the respective efficient portfolio are presented in parenthesis. The OOS refined Sharpe ratio ($\widehat{SR}_{ref}$) of each efficient MV expected utility/cardinality portfolio is also reported. The refined OOS Sharpe ratio was computed according to the Israelsen (2005) methodology.

The OOS certainty equivalent does not exhibit a clear trend with cardinality. In the MV framework, the best OOS certainty equivalent is achieved for the portfolio with cardinality equal to 2 (-0.02165%). The bootstrap p-values (computed according the classical methodology of Efron and Tibshirani, 1994) for the difference between the OOS certainty equivalent of the efficient MV portfolio with cardinality 2 and the OOS certainty equivalent of the MVS and MVSK portfolios with the same cardinality level (equal to 2) are both equal to 0.000. However, in the MVS and MVSK frameworks, the best OOS certainty equivalent return is achieved for the portfolios with cardinality equal to 1 (-0.01337% and -0.02635%, respectively for each framework). The bootstrap p-values for the difference between the OOS certainty equivalent of the efficient MV portfolio with cardinality 1 and the OOS certainty equivalent of the MVS and MVSK portfolios with the same cardinality level (equal to 1) are both equal to 0.000. These results suggest gains in considering skewness and kurtosis, for some cardinality levels.

Table 4.9: OOS analysis of the efficient MVS expected utility/cardinality portfolios

	Efficient portfolios							
Number of stocks	$\widehat{m}$	$\widehat{\sigma}$	$\widehat{s}_{stand}$	$\widehat{k}_{stand}$	$\widehat{u}$	$\widehat{SR}$	$\widehat{SR}/\widehat{SR}_{16}$	$\widehat{SR}_{ref}$
1	0.0003	0.0137	-0.2027	4.2231	-0.2501	0.0249	0.4410 (0.7243)	0.0249
2	0.0002	0.0134	0.8635	9.2178	-0.2503	0.0113	0.2001 (0.5447)	0.0113
3	-0.0002	0.0208	-0.3166	5.3093	-0.2513	-0.0091	-0.1618 (0.1044)	-0.0000
4	-0.0001	0.0231	-0.0339	5.2009	-0.2515	-0.0059	-0.1053 (0.0726)	-0.0000
5	0.0008	0.0254	-0.3214	4.1134	-0.2508	0.0334	0.5918 (0.3915)	0.0334
6	0.0007	0.0264	-0.2889	4.2218	-0.2510	0.0282	0.4994 (0.2456)	0.0282
7	0.0006	0.0274	-0.4334	4.5700	-0.2513	0.0219	0.3888 (0.0690)	0.0219
8	0.0007	0.0278	-0.3456	4.1169	-0.2512	0.0270	0.4788 (0.0852)	0.0270
9	0.0011	0.0287	-0.3050	4.4441	-0.2510	0.0382	0.6770 (0.2444)	0.0382
10	0.0011	0.0290	-0.3803	4.5557	-0.2511	0.0385	0.6829 (0.1652)	0.0385
11	0.0013	0.0302	-0.3448	4.4666	-0.2510	0.0445	0.7889 (0.3589)	0.0445
12	0.0016	0.0306	-0.3728	4.5447	-0.2508	0.0523	0.9269 (0.9046)	0.0523
13	0.0017	0.0311	-0.4084	4.5232	-0.2508	0.0555	0.9840 (0.8580)	0.0555
14	0.0018	0.0313	-0.3878	4.5733	-0.2508	0.0565	1.0026 (0.9690)	0.0565
15	0.0018	0.0313	-0.3999	4.5931	-0.2507	0.0575	1.0196 (0.5287)	0.0575
16	0.0018	0.0313	-0.4143	4.5854	-0.2508	0.0564	1	0.0564

This table reports the OOS mean ($\widehat{m}$), standard deviation ($\widehat{\sigma}$), standardized skewness ($\widehat{s}_{stand}$), standardized kurtosis ($\widehat{k}_{stand}$), expected utility ($\widehat{u}$) and the OOS Sharpe ratio ($\widehat{SR}$) of each efficient MVS expected utility/cardinality portfolio. It also reports the ratio, $\widehat{SR}/\widehat{SR}_{16}$, between the OOS Sharpe ratio of each efficient portfolio and the OOS Sharpe ratio of the efficient portfolio with cardinality 16. The Ledoit and Wolf (2008) bootstrap p-values for the statistical significance of the difference between the IS Sharpe ratio of the efficient portfolio with cardinality 16 and the respective efficient portfolio are presented in parenthesis. The OOS refined Sharpe ratio ($\widehat{SR}_{ref}$) of each efficient MVS expected utility/cardinality portfolio is also reported. The refined OOS Sharpe ratio was computed according to the Israelsen (2005) methodology.

Table 4.10: OOS analysis of the efficient MVSK expected utility/cardinality portfolios

Efficient portfolios								
Number of stocks	$\widehat{m}$	$\widehat{\sigma}$	$\widehat{s}_{stand}$	$\widehat{k}_{stand}$	$\widehat{u}$	$\widehat{SR}$	$\widehat{SR}/\widehat{SR}_{16}$	$\widehat{SR}_{ref}$
1	0.0002	0.0138	-0.1811	4.1731	-0.2503	0.0155	0.2763 (0.6363)	0.0155
2	-0.0001	0.0123	-0.2762	4.3074	-0.2505	-0.0116	-0.2069 (0.3833)	-0.0000
3	-0.0001	0.0206	-0.3015	5.5200	-0.2512	-0.0042	-0.0752 (0.1414)	-0.0000
4	-0.0003	0.0228	0.0790	4.9650	-0.2516	-0.0130	-0.2318 (0.0508)	-0.0000
5	0.0005	0.0251	-0.2708	4.3071	-0.2511	0.0185	0.3304 (0.1702)	0.0185
6	0.0006	0.0257	-0.2989	4.3997	-0.2511	0.0238	0.4246 (0.1846)	0.0238
7	0.0007	0.0265	-0.2967	3.8895	-0.2511	0.0277	0.4939 (0.1394)	0.0277
8	0.0007	0.0273	-0.4212	4.4027	-0.2512	0.0248	0.4410 (0.1260)	0.0248
9	0.0008	0.0274	-0.4249	4.5070	-0.2512	0.0283	0.5044 (0.0764)	0.0283
10	0.0014	0.0283	-0.4169	4.5242	-0.2507	0.0480	0.8545 (0.4993)	0.0480
11	0.0013	0.0295	-0.4497	4.6709	-0.2509	0.0448	0.7974 (0.3523)	0.0448
12	0.0017	0.0298	-0.3885	4.6517	-0.2506	0.0567	1.0100 (0.9812)	0.0567
13	0.0017	0.0304	-0.3968	4.5750	-0.2506	0.0573	1.0199 (0.8254)	0.0573
14	0.0019	0.0307	-0.3844	4.6244	-0.2506	0.0608	1.0838 (0.1760)	0.0608
15	0.0017	0.0307	-0.4052	4.6198	-0.2507	0.0570	1.0148 (0.6133)	0.0570
16	0.0017	0.0307	-0.4072	4.6057	-0.2507	0.0561	1	0.0561

This table reports the OOS mean ($\widehat{m}$), standard deviation ($\widehat{\sigma}$), standardized skewness ($\widehat{s}_{stand}$), standardized kurtosis ($\widehat{k}_{stand}$), expected utility ($\widehat{u}$) and the OOS Sharpe ratio ($\widehat{SR}$) of each efficient MVSK expected utility/cardinality portfolio. It also reports the ratio, $\widehat{SR}/\widehat{SR}_{16}$, between the OOS Sharpe ratio of each efficient portfolio and the OOS Sharpe ratio of the efficient portfolio with cardinality 16. The Ledoit and Wolf (2008) bootstrap p-values for the statistical significance of the difference between the IS Sharpe ratio of the efficient portfolio with cardinality 16 and the respective efficient portfolio are presented in parenthesis. The OOS refined Sharpe ratio ($\widehat{SR}_{ref}$) of each efficient MVSK expected utility/cardinality portfolio is also reported. The refined OOS Sharpe ratio was computed according to the Israelsen (2005) methodology.

As we can see in Table 4.7, both the efficient MVS and the efficient MVSK portfolios outperform the efficient MV portfolios for eight cardinality levels (cardinality equal to 1, 3, 6, 8, 9, 12, 13 and 15).

From the analysis of the refined Sharpe ratios, we conclude that the best OOS performance is achieved by portfolios with cardinality 14, for the MV and MVSK frameworks, and 15, for the MVS framework. For these cardinalities, the best portfolio is either the one obtained within the MVSK framework (for cardinality 14) or the one obtained within the MVS framework (for cardinality 15). However, we cannot reach strong conclusions regarding whether or not there are gains in considering higher moments, since the bootstrap p-values indicate that the performance of these portfolios is never significantly better than the one of the portfolios obtained within the MV framework.

Tables 4.8 to 4.10 also report, for each of the three frameworks, the ratio, $\widehat{SR}/\widehat{SR}_{16}$, between the OOS Sharpe ratio of each efficient portfolio and the OOS Sharpe ratio of the efficient portfolio corresponding to the maximum cardinality of 16. Based on these ratios, we computed (according to the methodology of Ledoit and Wolf 2008) the bootstrap p-values for the statistical significance of the difference between the OOS Sharpe ratio of the efficient portfolio with cardinality 16 and each other efficient portfolio. Except for the difference between the OOS Sharpe ratio of the MV portfolio with cardinality 8 and the MV portfolio with cardinality 16 (where the first is lower than the second), none of these differences were significant. This means that, for the three frameworks and except to the previous referred case, none of the efficient expected utility/cardinality portfolios presents a Sharpe ratio significantly higher or significantly lower than the Sharpe ratio of the efficient portfolio with cardinality 16.

Table 4.11 reports the bootstrap p-values of the difference between the OOS Sharpe ratio of each pair of frameworks (MV vs MVS, MV vs MVSK, and MVS vs MVSK) for the same cardinality level. The efficient MVS portfolios outperform the MV portfolios, in terms of Sharpe ratio, for nine cardinality levels (1, 3, 4, 6, 8, 9, 12, 13 and 15). Also for nine cardinality levels (1, 3, 6, 8, 9, 12, 13, 14, 15), the efficient MVSK portfolios outperform the MV portfolios. In total, for 75% of the cardinality levels (12 in 16), the efficient MV portfolios are outperformed either by the corresponding efficient MVS portfolios or by the corresponding efficient MVSK portfolios (although none of the differences are statistically significant at the 5% significance level). This suggests that, for those cardinality levels, there may be small gains in considering higher moments (skewness and kurtosis).

Given the time series of daily OOS returns, for each efficient portfolio considered, we computed the portfolio turnover, defined as the average, over all time periods, of the

Table 4.11: Bootstrap p-values for the Sharpe ratios

Efficient portfolios			
Number of stocks	MV vs MVS	MV vs MVSK	MVS vs MVSK
1	0.1794 (MVS)	0.4851 (MVSK)	0.1648 (MVS)
2	0.5515 (MV)	0.4803 (MV)	0.4713 (MVS)
3	0.8930 (MVS)	0.3801 (MVSK)	0.3439 (MVSK)
4	0.9540 (MVS)	0.5007 (MV)	0.3787 (MVS)
5	0.4251 (MV)	0.0658 (MV)	0.1792 (MVS)
6	0.6323 (MVS)	0.9478 (MVSK)	0.6743 (MVS)
7	0.1862 (MV)	0.6331 (MV)	0.5469 (MVSK)
8	0.4759 (MVS)	0.5353 (MVSK)	0.8630 (MVS)
9	0.4171 (MVS)	0.8924 (MVSK)	0.2198 (MVS)
10	0.1746 (MV)	0.7409 (MV)	0.2513 (MVSK)
11	0.4935 (MV)	0.5563 (MV)	0.9672 (MVSK)
12	0.9658 (MVS)	0.4541 (MVSK)	0.3415 (MVSK)
13	0.7003 (MVS)	0.3705 (MVSK)	0.2941 (MVSK)
14	0.7960 (MV)	0.1132 (MVSK)	0.0426 (MVSK)
15	0.6179 (MVS)	0.7844 (MVSK)	0.7674 (MVS)
16	0.7816 (MV)	0.6415 (MV)	0.7890 (MVS)

This table lists the bootstrap p-values of the difference between the Sharpe ratio of each pair of frameworks. MV refers to the efficient MV expected utility/cardinality portfolios, MVS refers to the efficient MVS expected utility/cardinality portfolios, and MVSK refers to the efficient MVSK expected utility/cardinality portfolios. The bootstrap p-values were computed according the Ledoit and Wolf (2008) methodology. In parenthesis there is the indication of the framework with the highest Sharpe ratio.

absolute changes in weights (corresponding to trades) across the N available stocks:

$$\text{turnover} = \frac{1}{\#periods} \sum_{t=1}^{\#periods} \sum_{i=1}^{N} \left(\left| w_{i,t+1} - w_{i,t}^{h} \right| \right), \tag{4.9}$$

where $w_{i,t+1}$ is the weight of stock i after rebalancing at time $t+1$ and $w_{i,t}^{h}$ is that weight before rebalancing at time $t+1$. Thus, $w_{i,t}^{h}$, is computed as

$$w_{i,t}^{h} = w_{i,t-1} \frac{1 + r_{i,t}}{1 + r_{p,t}}, \tag{4.10}$$

where $r_{i,t}$ is the return at time t of the stock i and $r_{p,t}$ is the return at time t of the portfolio. The results are reported in Table 4.12. In all the three frameworks, the turnover begins by increasing up to a certain cardinality level (around 8 or 9, which corresponds to about 50% of the maximum cardinality level). This is somehow expected, since the increase in the number of stocks should increase the transaction costs. However, after reaching that cardinality level, we observe, in all the three frameworks, the non-intuitive fact that the turnover decreases when the cardinality increases. The explanation for this result is that the small fluctuations of the portfolio weights compensate the increment of the number of stocks in the portfolio. Thus, in this case, from the cardinality level of about 8 (around 50% of the maximum cardinality), increasing the cardinality of a portfolio allows the investor to reduce transaction costs.

In order to assess the impact of transaction costs in the performance of the different strategies, we computed the OOS Sharpe ratio of returns net of transaction costs, defined as

$$\widehat{SR}_{tc} = \frac{\widehat{m}_{tc}}{\widehat{\sigma}_{tc}}, \tag{4.11}$$

where $\widehat{m}_{tc} = \widehat{m} - tc$ is the OOS time series of the mean of the returns ($\widehat{m}$) deducted by transaction costs (tc) and $\hat{\sigma}_{tc}$ is the standard deviation of the OOS returns after transaction costs. We used proportional transaction costs defined as[11]

[11] Based on the information received from several stock brokers in the Lisbon Stock Exchange (Euronext Lisbon), for the stocks in the PSI 20 Index the proportional transaction cost was set equal to 30 basis points per transaction.

Table 4.12: Portfolio turnover

Efficient portfolios			
Number of stocks	MV	MVS	MVSK
1	5.2542	3.7711	5.8529
2	12.1198	9.7732	12.9851
3	11.2775	11.6299	12.7648
4	45.3348	38.2224	44.9272
5	44.7624	45.2611	60.5814
6	55.2793	50.5924	60.7667
7	53.1998	49.3179	59.9827
8	69.7246	60.9169	78.7103
9	61.6530	62.8226	76.8444
10	47.1448	51.1774	56.6053
11	44.4329	47.8681	53.4236
12	32.5758	39.6748	43.4849
13	33.5141	32.5745	34.4450
14	33.6567	33.2447	36.7984
15	30.0968	29.0990	33.1650
16	28.9799	29.7610	33.4280

This table lists the portfolio turnover of the efficient expected utility/cardinality portfolios. MV refers to the efficient MV expected utility/cardinality portfolios, MVS refers to the efficient MVS expected utility/cardinality portfolios and MVSK refers to the efficient MVSK expected utility/cardinality portfolios.

$$tc = \sum_{t=1}^{\#periods} 0.3\% \sum_{i=1}^{N} \left(\left| w_{i,t+1} - w_{i,t}^{h} \right| \right). \tag{4.12}$$

The refined Sharpe ratios after transaction costs for the efficient portfolios are reported in Table 4.13. The inclusion of transaction costs may lead to completely different results (see, e.g., Brito and Vicente, 2014; DeMiguel et al., 2014; DeMiguel and Olivares-Nadal, 2016). Not taking into account the transaction costs in the OOS analysis may result in the misleading conclusion that diversification is not an important issue in portfolio management. For example, Brodie et al. (2009) have implemented a specific algorithm in order to construct sparse portfolios (portfolios with low cardinality) in the classical MV setting. Their empirical OOS results, for two sets of portfolios constructed by Fama and French (FF48 and FF100), indicate that sparse portfolios outperform (in terms of Sharpe ratio) the equally weighted portfolio, suggesting that diversification does not produce positive results OOS (for the FF48 dataset, the Sharpe ratio even exhibits a decreasing trend with cardinality). However, the authors do not account for transaction costs, limiting the real applicability of their findings.

The results that we obtained for the net Sharpe ratios highlight that in the three frameworks, the pattern of the net Sharpe ratios is inversely related with the turnover pattern, i.e., the net Sharpe ratios decrease to a certain level of cardinality (around 8 or 9) and then increases with the cardinality level. For all frameworks, the portfolio with the minimum cardinality of 1 is the portfolio with the highest refined Sharpe ratio after transaction costs. This is in agreement with previous studies suggesting that sparse portfolios tend to exhibit a good performance (see, e.g., Brodie et al., 2009). However, the results that we obtain go further and suggest that for sufficiently large levels of cardinality, there are gains with diversification.

In Table 4.14 we report the bootstrap p-values of the difference between the Sharpe ratio after transaction costs for each pair of frameworks (MV vs MVS, MV vs MVSK, and MVS vs MVSK) given a specific cardinality level. From the comparison between the MV and the MVS portfolios we found that the MVS portfolios outperform the MV portfolios (the differences being statistically significant) for nine cardinality levels (1, 2, 4, 6, 7, 8, 13, 14 and 15). Looking to the MV and the MVSK portfolios, we found that the MVSK portfolios outperform the MV portfolios only for two cardinality levels (2 and 4), but the differences are not statistically significant. Summarizing, the MV portfolios underperform at least one of the efficient portfolios from the two other frameworks for 9 of 16 cardinality levels. This indicates that when transaction costs are taken into account, there are sometimes gains (in terms of Sharpe ratio) in considering

Table 4.13: Refined Sharpe ratios of returns net of transaction costs

Efficient portfolios			
Number of stocks	MV	MVS	MVSK
1	-0.0566	-0.0406	-0.0631
2	-0.1297	-0.1024	-0.1255
3	-0.1845	-0.1896	-0.2058
4	-0.8184	-0.6921	-0.8038
5	-0.8744	-0.9002	-1.1893
6	-1.1531	-1.0444	-1.2229
7	-1.1312	-1.0588	-1.2446
8	-1.5160	-1.3237	-1.6839
9	-1.3875	-1.4105	-1.6457
10	-1.0867	-1.1639	-1.2563
11	-1.0476	-1.1319	-1.2325
12	-0.7891	-0.9492	-1.0138
13	-0.8198	-0.7939	-0.8202
14	-0.8280	-0.8151	-0.8845
15	-0.7394	-0.7138	-0.7976
16	-0.7126	-0.7303	-0.8026

This table reports the refined Sharpe ratios of returns net of transaction costs for each efficient expected utility/cardinality portfolio. MV refers to the efficient MV expected utility/cardinality portfolios, MVS refers to the efficient MVS expected utility/cardinality portfolios and MVSK refers to the efficient MVSK expected utility/cardinality portfolios. The refined net Sharpe ratios were computed according the Israelsen (2005) methodology.

the skewness, for certain cardinality levels.

Table 4.14: Bootstrap p-values for the net Sharpe ratios

Efficient portfolios			
Number of stocks	MV vs MVS	MV vs MVSK	MVS vs MVSK
1	0.0002 (MVS)	0.0002 (MV)	0.0002 (MVS)
2	0.0002 (MVS)	0.1527 (MVSK)	0.0018 (MVS)
3	0.0002 (MV)	0.0002 (MV)	0.0002 (MVS)
4	0.0002 (MVS)	0.9999 (MVSK)	0.0002 (MVS)
5	0.5037 (MV)	0.0002 (MV)	0.0002 (MVS)
6	0.0006 (MVS)	0.0002 (MV)	0.0002 (MVS)
7	0.0002 (MVS)	0.0002 (MV)	0.0002 (MVS)
8	0.0006 (MVS)	0.0002 (MV)	0.0002 (MVS)
9	0.2841 (MV)	0.0002 (MV)	0.0002 (MVS)
10	0.0002 (MV)	0.0002 (MV)	0.0002 (MVS)
11	0.0002 (MV)	0.0002 (MV)	0.0002 (MVS)
12	0.0002 (MV)	0.0002 (MV)	0.0002 (MVS)
13	0.0002 (MVS)	0.0002 (MV)	0.0002 (MVS)
14	0.0014 (MVS)	0.0002 (MV)	0.0002 (MVS)
15	0.0002 (MVS)	0.0002 (MV)	0.0002 (MVS)
16	0.0002 (MV)	0.0002 (MV)	0.0002 (MVS)

This table lists the bootstrap p-values of the difference between the net Sharpe ratios of two different frameworks. MV refers to the efficient MV expected utility/cardinality portfolios, MVS refers to the efficient MVS expected utility/cardinality portfolios, and MVSK refers to the efficient MVSK expected utility/cardinality portfolios. The bootstrap p-values were computed according the Ledoit and Wolf (2008) methodology. The framework with the highest Sharpe ratio is indicated in parenthesis.

The cardinality becomes an even more important issue in datasets with a larger number of stocks. In order to provide an indication of whether certain patterns found in the results discussed above (for the PSI 20 Index) remain valid in larger datasets, we decided to apply the methodology to a dataset from the EURO STOXX 50 Index (see Table 4.15).

Similarly to the PSI 20 dataset, the daily discrete returns of the EURO STOXX 50 dataset (see Table 4.16), presented, on average, negative but near zero mean, positive skewness and above normal kurtosis.

We applied, to the EURO STOXX 50 dataset, exactly the same rolling window approach (see Figure 2.5) described in the beginning of this Section. Table 4.17 reports the OOS certainty equivalent (solution of Equation (4.7)), $\widehat{CE}$, for the MV, MVS and MVSK frameworks[12].

[12]Similarly to the case of the PSI 20 dataset, the introduction of a threshold of 1% for the cardinality computation reduced the maximum cardinality level that led to nondominated solutions. In the case

Table 4.15: The EURO STOXX 50 dataset

List of stocks	
AIR LIQUIDE	IBERDROLA
AIRBUS GROUP	INDITEX
ALLIANZ (XET)	ING GROEP
ANHEUSER-BUSCH INBEV	INTESA SANPAOLO
ASML HOLDING	L'OREAL
ASSICURAZIONI GENERALI	LVMH
AXA	MUENCHENER RUCK (XET)
BANCO SANTANDER	ORANGE
BASF (XET)	PHILIPS ELTN KONINKLIJKE
BAYER (XET)	REPSOL YPF
BBV.ARGENTARIA	RWE (XET)
BMW (XET)	SAINT GOBAIN
BNP PARIBAS	SANOFI
CARREFOUR	SAP (XET)
CRH (DUB)	SCHNEIDER ELECTRIC SE
DAIMLER (XET)	SIEMENS (XET)
DANONE	SOCIETE GENERALE
DEUTSCHE BANK (XET)	TELEFONICA
DEUTSCHE POST (XET)	TOTAL
DEUTSCHE TELEKOM (XET)	UNIBAIL-RODAMCO
E ON (XET)	UNICREDIT
ENEL	UNILEVER CERTS
ENI	VINCI
ESSILOR INTL	VIVENDI
GDF SUEZ	VOLKSWAGEN PREF (XET)

This table lists the composition of the EURO STOXX 50 dataset. Daily closing prices, from July 2007 to June 2014, of these 50 stocks (from the EURO STOXX 50 Index) were collected from the Thomson Reuters Datastream®.

Table 4.16: Descriptive statistics for the EURO STOXX 50 dataset

Number of days	1826
Minimum	-0.1298
Median	0.0001
Maximum	0.1667
Mean	0.0002
Variance	0.0005
Skewness	0.3399
Kurtosis	10.2218

This table reports some descriptive statistics for the EURO STOXX 50 dataset (the composition of the dataset is reported in Table 4.15). The reported values of skewness and kurtosis, concern to the third and fourth standardized moments, respectively. These statistics are averaged cross-sectionally, i.e., they are computed for each stock and then the arithmetic mean is taken.

From the results in Table 4.17 we can see that, similarly to what occurred with PSI 20 dataset, the certainty equivalent does not show a clear trend with increasing cardinality, and the best certainty equivalent values always occur at a low cardinality level (cardinality 2 for all frameworks).

Table 4.18 reports the bootstrap p-values (computed according the Ledoit and Wolf, 2008 robust methodology) of the difference between the OOS Sharpe ratio of each pair of frameworks (MV vs MVS, MV vs MVSK, and MVS vs MVSK) for the same cardinality level. Table 4.18 also reports, for each cardinality level, which framework shows the highest Sharpe ratio. For most cardinality levels, either the MVS or the MVSK efficient portfolios outperform the efficient MV portfolios, but such differences are never statistically significant.

The results for the OOS certainty equivalent and Sharpe ratio, both suggest gains, for certain cardinality levels, in considering higher moments. These results are in accordance with the results obtained for the PSI 20 dataset.

Similarly to what occurred with the PSI 20 dataset, the turnover exhibits a non-monotonic pattern (see Table 4.19). In all the frameworks we observe that the turnover begins by increasing up to a certain cardinality level (17 in the EURO STOXX 50 dataset case, which corresponds to 42.5% of the maximum cardinality level), then decreases. However, starting from the cardinality level equal to 24 we observe, again, an increase in the turnover. Then, for high cardinality levels, the trend ceases to be clear.

of the EURO STOXX 50 dataset the maximum cardinality is equal to 40.

Table 4.17: Certainty equivalent return for the EURO STOXX 50 dataset

Efficient portfolios			
Number of stocks	MV	MVS	MVSK
1	-1.1228	-1.2180	-0.8401
2	-0.7150	-0.8362	-0.5988
3	-1.6367	-1.5413	-1.5261
4	-1.3593	-1.2041	-1.3267
5	-1.7554	-1.8701	-1.6846
6	-2.1226	-2.3514	-1.9385
7	-2.2929	-2.3568	-2.3579
8	-2.8023	-2.9306	-2.8663
9	-3.0445	-2.9943	-2.9719
10	-3.4889	-3.3824	-2.9750
11	-3.4794	-3.5896	-3.3237
12	-3.4635	-3.7971	-3.5398
13	-3.6888	-4.1857	-3.7955
14	-3.8448	-4.1124	-3.9549
15	-4.0945	-4.3040	-4.0928
16	-4.0015	-4.2972	-4.0927
17	-3.9952	-4.1224	-4.0040
18	-4.0168	-4.1569	-4.1108
19	-4.0383	-4.1908	-4.2306
20	-4.2002	-4.1429	-4.3419
21	-4.1040	-4.3728	-4.3479
22	-4.1218	-4.2493	-4.1209
23	-4.1219	-4.2553	-3.9249
24	-4.3303	-4.2030	-4.1561
25	-4.1947	-4.2743	-4.1720
26	-4.2986	-4.2314	-4.1753
27	-4.0887	-4.1126	-4.1933
28	-4.0956	-3.9658	-4.0411
29	-4.1652	-4.0173	-4.0393
30	-4.1568	-4.0311	-4.0897
31	-4.1389	-4.1134	-4.0677
32	-4.1234	-4.0914	-4.0499
33	-4.1076	-4.0932	-4.0833
34	-3.8679	-4.0807	-4.0476
35	-3.9078	-4.0378	-4.0585
36	-4.2453	-3.9664	-3.9836
37	-3.7742	-3.9499	-3.9635
38	-3.9052	-3.9963	-4.0322
39	-3.7590	-3.9894	-4.0615
40	-3.9687	-3.9451	-3.9964

This table lists the OOS certainty equivalent return ($\widehat{CE}$) of the efficient expected utility/cardinality portfolios. MV refers to the efficient MV expected utility/cardinality portfolios, MVS refers to the efficient MVS expected utility/cardinality portfolios and MVSK refers to the efficient MVSK expected utility/cardinality portfolios. All the certainty equivalent values are multiplied by a factor of 1000.

Table 4.18: Bootstrap p-values for the Sharpe ratios, for the EURO STOXX 50 dataset

Efficient portfolios	Bootstrap p-values		
Number of stocks	MV vs MVS	MV vs MVSK	MVS vs MVSK
1	0.8018 (MVS)	0.3827 (MVSK)	0.3101 (MVSK)
2	0.5583 (MVS)	0.4371 (MVSK)	0.2486 (MVSK)
3	0.5211 (MV)	0.7361 (MVSK)	0.8246 (MVSK)
4	0.4119 (MV)	0.9116 (MVSK)	0.5945 (MVS)
5	0.6313 (MVS)	0.8864 (MVSK)	0.3275 (MVSK)
6	0.4409 (MVS)	0.2242 (MVSK)	0.0058 (MVSK)
7	0.6443 (MVS)	0.5523 (MV)	0.8484 (MVS)
8	0.7281 (MVS)	0.4811 (MV)	0.5559 (MVSK)
9	0.5077 (MV)	0.6737 (MVSK)	0.6775 (MVSK)
10	0.6395 (MV)	0.0772 (MVSK)	0.0940 (MVSK)
11	0.7806 (MVS)	0.6211 (MVSK)	0.4017 (MVSK)
12	0.1296 (MVS)	0.4769 (MV)	0.3759 (MVSK)
13	0.0636 (MVS)	0.4425 (MV)	0.2064 (MVSK)
14	0.2040 (MVS)	0.3207 (MV)	0.7620 (MVSK)
15	0.5417 (MVS)	0.9442 (MVSK)	0.5455 (MVSK)
16	0.1772 (MVS)	0.5427 (MV)	0.3493 (MVSK)
17	0.5688 (MVS)	0.6474 (MVSK)	0.8520 (MVSK)
18	0.5113 (MVS)	0.5085 (MV)	0.9474 (MVSK)
19	0.5303 (MVS)	0.2328 (MV)	0.5893 (MVS)
20	0.6087 (MV)	0.2346 (MV)	0.0310 (MVS)
21	0.2791 (MVS)	0.0963 (MV)	0.5367 (MVSK)
22	0.7271 (MVS)	0.6437 (MVSK)	0.9012 (MVSK)
23	0.9340 (MVS)	0.3761 (MVSK)	0.2244 (MVSK)
24	0.2460 (MV)	0.6237 (MVSK)	0.4115 (MVSK)
25	0.8732 (MVS)	0.8746 (MVSK)	0.9968 (MVSK)
26	0.5815 (MV)	0.6839 (MVSK)	0.8890 (MVSK)
27	0.8546 (MVS)	0.4657 (MV)	0.4809 (MVS)
28	0.3017 (MV)	0.8048 (MVSK)	0.3069 (MVS)
29	0.1588 (MV)	0.3917 (MVSK)	0.5469 (MVS)
30	0.2401 (MV)	0.8660 (MVSK)	0.1978 (MVS)
31	0.4591 (MV)	0.8498 (MVSK)	0.4647 (MVSK)
32	0.3927 (MV)	0.6527 (MVSK)	0.6903 (MVSK)
33	0.4287 (MV)	0.9848 (MVSK)	0.3715 (MVSK)
34	0.3655 (MVS)	0.2334 (MV)	0.6139 (MVSK)
35	0.5713 (MVS)	0.0819 (MV)	0.1556 (MVS)
36	0.1554 (MV)	0.3589 (MVSK)	0.3579 (MVS)
37	0.4453 (MVS)	0.1414 (MV)	0.3659 (MVS)
38	0.9948 (MVS)	0.6247 (MV)	0.3119 (MVS)
39	0.2218 (MVS)	0.0321 (MV)	0.1838 (MVS)
40	0.6309 (MVS)	0.5311 (MV)	0.1858 (MVS)

This table lists the bootstrap p-values of the difference between the Sharpe ratio of each pair of frameworks. MV refers to the efficient MV expected utility/cardinality portfolios, MVS refers to the efficient MVS expected utility/cardinality portfolios, and MVSK refers to the efficient MVSK expected utility/cardinality portfolios. The bootstrap p-values were computed according the Ledoit and Wolf (2008) methodology. In parenthesis there is the indication of the framework with the highest Sharpe ratio.

Table 4.19: Portfolio turnover, for the EURO STOXX 50 dataset

Efficient portfolios			
Number of stocks	MV	MVS	MVSK
1	13.9087	13.8344	15.0348
2	13.1303	14.6790	15.0737
3	40.1251	44.3824	49.4342
4	56.0273	60.3688	64.7256
5	50.1175	60.7830	55.5558
6	80.3495	78.7964	83.3818
7	84.5984	82.0087	87.8145
8	81.9265	97.1036	102.5520
9	72.5228	96.7480	97.5971
10	107.1080	130.0243	130.8012
11	133.0312	132.0839	144.8024
12	138.5046	142.4454	148.6940
13	138.0521	150.0048	146.9185
14	141.3580	138.2097	141.5113
15	137.1693	137.6385	148.3001
16	139.5588	143.8476	149.4672
17	141.5487	141.7356	153.4877
18	126.5044	131.2663	136.0524
19	125.1619	124.8980	129.0187
20	123.9810	126.5021	127.5923
21	125.7462	120.8936	122.2386
22	118.0602	120.9263	122.4530
23	117.5878	119.5078	121.6786
24	118.8219	122.2902	130.7216
25	127.2803	134.0256	135.8087
26	128.9391	135.6442	136.8285
27	129.0486	136.3405	139.5051
28	136.1073	139.2059	140.4296
29	135.3056	137.4664	143.2948
30	139.4502	139.4577	143.5130
31	136.3280	141.3910	141.2399
32	136.1948	140.3978	143.4964
33	133.0652	141.9928	142.5606
34	141.7797	142.7401	141.0815
35	136.6671	140.6893	143.4869
36	166.3403	142.3293	142.4093
37	145.4961	138.7303	144.4675
38	166.2088	160.9871	157.5376
39	145.6927	139.3858	146.6295
40	137.5481	139.7157	141.8584

This table lists the portfolio turnover of the efficient expected utility/cardinality portfolios. MV refers to the efficient MV expected utility/cardinality portfolios, MVS refers to the efficient MVS expected utility/cardinality portfolios and MVSK refers to the efficient MVSK expected utility/cardinality portfolios.

Regarding the net Sharpe ratio (see Table 4.20), as for the PSI 20 dataset, we observe that sparse portfolios tend to exhibit the best performance: in all the three frameworks the efficient portfolios corresponding to the cardinality level equal 2 are the portfolios with the highest refined net Sharpe ratio. Finally, we notice that, for the EURO STOXX 50 dataset, the MV efficient portfolios outperform the MVS and the MVSK efficient portfolios for most cardinalities.

4.4 Conclusions

This chapter extends the analysis of the cardinality impact in the portfolio performance from the standard MV framework to frameworks with higher moments, namely by considering the skewness and kurtosis. We propose a biobjective model that allows the direct analysis of the efficient tradeoff between expected utility and cardinality. This analysis was conducted assuming an investor with CRRA preferences (with a relative risk aversion parameter of five). Although we have assumed a specific form for the investor's utility, the proposed methodology may easily be applied to any utility function.

We have conducted an empirical application of the proposed methodology to a dataset based on the Portuguese Stock Market Index (PSI 20 Index). The IS results show that both the certainty equivalent and the Sharpe ratio increase with the cardinality level and suggest that there are no gains, in terms of certainty equivalent, when considering higher moments. When transaction costs are not considered, the OOS results on the certainty equivalent and Sharpe ratio do not show a clear trend with cardinality. However, the turnover increases up to a certain level of cardinality and then decreases. This result leads to another interesting result that the OOS net Sharpe ratio – that is, the Sharpe ratio net of transaction costs – decreases up to a certain cardinality level, then increases and afterwards behaves more erratically for larger values of the cardinality. So, for an important range of cardinality levels, diversification has a positive effect when transaction costs are taken into account.

We also used a dataset based on the EURO STOXX 50 Index in order to check the robustness of the OOS results. Since in this dataset we have a larger number of stocks, we are able to build portfolios with a larger cardinality. We notice that, for larger levels of cardinality, the turnover and the net Sharpe ratio no longer show a clear trend. After an initial increase in turnover, and the subsequent decrease, the turnover increases again and tends to behave more erratically. Similarly, after an initial decrease in the net Sharpe ratio, and the subsequent increase, the net Sharpe ratio decreases again and tends to behave more erratically. This indicates that the

Table 4.20: Refined Sharpe ratios of returns net of transaction costs, for the EURO STOXX 50 dataset

Efficient portfolios			
Number of stocks	MV	MVS	MVSK
1	-0.1360	-0.1395	-0.1426
2	-0.1073	-0.1237	-0.1213
3	-0.4877	-0.5414	-0.5870
4	-0.6547	-0.7181	-0.7536
5	-0.6448	-0.7818	-0.7024
6	-1.0447	-1.0544	-1.0964
7	-1.1716	-1.1325	-1.2033
8	-1.2258	-1.4757	-1.4913
9	-1.0991	-1.5041	-1.4862
10	-1.6402	-2.0026	-1.9751
11	-2.0884	-2.0941	-2.2504
12	-2.2377	-2.2704	-2.3405
13	-2.2362	-2.4233	-2.3268
14	-2.3225	-2.2756	-2.2848
15	-2.2672	-2.2993	-2.4405
16	-2.3264	-2.3787	-2.4668
17	-2.3837	-2.3738	-2.5310
18	-2.1571	-2.2256	-2.2903
19	-2.1514	-2.1606	-2.2140
20	-2.1534	-2.2148	-2.1834
21	-2.2019	-2.1345	-2.1065
22	-2.0808	-2.1540	-2.1216
23	-2.0638	-2.1430	-2.1198
24	-2.1094	-2.2029	-2.2845
25	-2.2693	-2.4090	-2.3981
26	-2.3133	-2.4471	-2.4359
27	-2.3214	-2.4474	-2.4920
28	-2.4312	-2.5182	-2.5045
29	-2.4173	-2.4838	-2.5619
30	-2.4952	-2.5202	-2.5465
31	-2.4408	-2.5786	-2.5069
32	-2.4298	-2.5431	-2.5554
33	-2.3817	-2.5873	-2.5390
34	-2.5413	-2.5894	-2.5184
35	-2.4547	-2.5517	-2.5375
36	-2.9631	-2.5736	-2.5236
37	-2.6060	-2.5100	-2.5616
38	-2.9530	-2.9100	-2.7977
39	-2.6082	-2.5201	-2.6066
40	-2.4691	-2.5253	-2.5195

This table reports the refined Sharpe ratios of returns net of transaction costs for each efficient expected utility/cardinality portfolio. MV refers to the efficient MV expected utility/cardinality portfolios, MVS refers to the efficient MVS expected utility/cardinality portfolios and MVSK refers to the efficient MVSK expected utility/cardinality portfolios. The refined net Sharpe ratios were computed according the Israelsen (2005) methodology.

best level of cardinality should be a concern for investors wanting to diversify their portfolios, and we cannot assume that a larger cardinality is always better.

In this study we assumed that the investor has CRRA preferences with a relative risk aversion of five. In future research we are also interested in studying the sensitivity of these results to different choices of the relative risk aversion level and also to different utility functions (e.g. hyperbolic absolute risk aversion (HARA)-utility).

Chapter 5

Portfolio choice with high frequency data

5.1 Introduction

The availability of high frequency financial databases has increased in recent years, which has opened new fields of research both in financial economics and financial econometrics. Although Merton (1980) has observed that the variance can be accurately estimated as the sum of the realizations of the squared intraday returns, the researchers attention has been devoted to ARCH type (Engle, 1982; Bollerslev, 1986; Nelson, 1991) and stochastic volatility models (Taylor, 1986), at least until the end of the 90s. Nevertheless, based on the work of Schwert (1989) and Hsieh (1991), authors such as Andersen et al. (2001b) and Barndorff-Nielsen and Shephard (2002) began to use intraday data to estimate the variance as the sum of squared returns, sampled at very short intraday intervals. This new approach, known in the literature as realized volatility, has a straightforward reasoning: since the sample path of the variance is continuous, the accuracy of the variance estimates increases with the sampling frequency (Merton, 1980). Moreover, it is quite appealing as it is a model-free and an error-free measure of volatility (because it is observable) which converges to the quadratic variation (Andersen et al., 2006). Several useful surveys on realized volatility can be found in the literature (see, e.g., Barndorff-Nielsen and Shephard, 2005; Andersen et al., 2006; McAleer and Medeiros, 2008; Meddahi et al., 2011). Papers like Andersen et al. (2001a), Andersen et al. (2001b), Areal and Taylor (2002), and Koopman et al. (2005) use the realized volatility in univariate frameworks. Other papers, like Andersen et al. (2003), Flemming et al. (2003), Barndorff-Nielsen and Shephard (2004), Liu (2009), Fan et al. (2012), and Hautsch et al. (2012) extend the approach to multivariate cases. With the remarkable growth of related literature, special attention has been given to two ubiquitous problems when dealing with high frequency financial data: market microstructure noise and asynchronous price observations. The most common response to the presence of microstructure noise has been to reduce the sampling frequency to some arbitrary level, say 5-minutes or 30-minutes (Andersen et al., 2001a; Hansen and Lunde, 2006). Another possibility is to use all the available high frequency data (seconds, milliseconds) and taking explicitly into account the microstructure noise in the volatility estimation (Aït-Sahalia et al., 2005a;b; 2011). Regarding the nonsynchronous price observations, Barndorff-Nielsen et al. (2011) have proposed the well-known realized kernel estimator for the multivariate case, which has the advantage of being a positive-definite estimator, quite suitable for dealing with asynchronous data.

The approaches originally defined for realized volatility can be extended, with equal interest and potential, to higher moments. Neuberger (2012) introduced the realized skewness as the sum of the 3rd power of returns, while Amaya et al. (2015) defined the realized kurtosis as the sum of the 4th power of returns.

Concerning the variance estimation, there are already some studies suggesting the existence of benefits in using high frequency data (see, e.g., Flemming et al., 2003; Liu, 2009). On the other hand, Amaya et al. (2015) have found a negative effect of skewness and a positive effect of kurtosis on weekly stock returns. However, an important question remains open: are there performance gains, for portfolio choice purposes, in the joint use of the three realized moments (variance, skewness and kurtosis)? In Section 5.3, we try to contribute, empirically, to answering this question.

The majority of studies in portfolio choice assume that securities' returns are the only source of information. However, the subprime crisis, which led to a worldwide market liquidity crisis, has highlighted the importance of liquidity not only for each particular investor but also for the achievement of the allocative rationale inherent to financial markets. Liquidity is the easiness to trade a security. Although quite simple to enunciate, liquidity is an elusive concept, and in fact, one may enumerate three main dimensions of a liquid market: depth (high quantities available for sale or purchase away from the current market price), breadth (large number of market participants) and resiliency (price impacts caused by trading are small and transitory). Some authors have found a positive relationship between stock returns and alternative proxies for liquidity: Amihud and Mendelsen (1986) and Datar et al. (1998) used as a liquidity proxy the bid-ask spread; Brennan and Subrahmanyam (1996) used price impacts; Easley et al. (2002) used the probability of informed trading (PIN). Other papers found the existence of commonality and predictability in liquidity (Chordia et al., 2001; Hasbrouck and Seppi, 2001; Huberman and Halka, 2001; Amihud, 2002; Pastor and Stambaugh, 2003).

It is known that liquidity is a direct function of the implicit and explicit trading costs. However, the quantification of these costs is not a trivial task, not only due to its conceptual vagueness but also, sometimes, due to the lack of information. Thus, there are several proxies to measure liquidity: bid-ask spread, trading volume, turnover, quote size and price impact. Goyenko et al. (2009) compared different liquidity measures and found that the Amihud illiquidity ratio (Amihud, 2002) is one of the best liquidity proxies, having a strong correlation with several other liquidity measures.

Section 5.4 presents a new methodology for portfolio choice in the expected utility-liquidity space. The proposed EU/L model allows the investor to identify the optimal portfolios, which have the maximum expected utility, computed with higher moments, among all that provide at least a certain expected level of liquidity. In this section, we also consider high frequency data by using realized estimators as the inputs of the optimization model.

The remainder of this chapter proceeds as follows. Section 5.2 explains the procedures for estimating higher moments using high frequency data. Section 5.3 conducts an empirical application on fourteen stocks of the French Stock Market Index (CAC 40 Index) comparing between several low and high frequency portfolios. Section 5.4 presents the expected utility-liquidity problem (EU/L problem) and performs an empirical exercise (on the same set of fourteen stocks from the CAC 40 Index). Finally, Section 5.5 presents the main conclusions.

5.2 The investor's problem with higher realized moments

When the available intraday trading data suffer from nonsynchronous trading effects, this induces potentially serious biases in the moments and co-moments of returns (Campbell et al., 1997, p. 84-98). In such case, we should adopt some procedure in order to synchronize the data. One possible way to accomplish this, is using the all refresh-time method (Barndorff-Nielsen et al., 2011). This method will be summarily described in Section 5.3.1.

Suppose that for each stock, $i = 1, \ldots, N$, we have Q synchronized intraday price observations, in day $t+1$ we have $P_{t+(q/Q)}$, with $q = 1, \ldots, Q$, price observations. Note that the closing price of day $t+1$ is given by $P_{t+(Q/Q)} = P_{t+1}$. In this setting, the daily realized variance (Andersen et al., 2001a) at day $t+1$, for each individual stock i, is given by

$$rv^{q}_{i,t+1} = \sum_{q=1}^{Q} r^2_{i,t+(q/Q)}, \tag{5.1}$$

where $r_{i,t+(q/Q)}$ is the return of stock i in the intraday period q. For each pair of stocks, i and j, with $i, j = 1, \ldots, N$, the corresponding daily realized covariance, at day $t+1$, is given by

$$rcov^{Q}_{i,j,t+1} = \sum_{q=1}^{Q} r_{i,t+(q/Q)} r_{j,t+(q/Q)}. \tag{5.2}$$

The daily portfolio realized variance can thus be computed as

$$rv_{t+1} = w^{\top}_{t+1} R\Sigma_{t+1} w_{t+1}, \tag{5.3}$$

where $R\Sigma_{t+1}$ represents the realized covariance matrix. Each entry, $c_{ij,t+1}$, of the $R\Sigma_{t+1}$ matrix is given by

$$c_{ij,t+1} = \frac{1}{t} \sum_{\tau=1}^{t} \sum_{q=1}^{Q} r_{i,\tau+(q/Q)} r_{j,\tau+(q/Q)}. \tag{5.4}$$

Analogously to the realized variance approach, the daily realized skewness (Neuberger, 2012) at day $t+1$, for each individual stock i, can be defined as

$$rs^{Q}_{i,t+1} = \sum_{q=1}^{Q} r^3_{i,t+(q/Q)}. \tag{5.5}$$

The realized coskewness matrix can be computed as a $N \times N^2$ matrix (adapting the procedure, for the computation of the coskewness matrix, described in detail by Athayde and Flôres, 2004). According to this procedure, the daily portfolio realized skewness can be computed as

$$rs_{t+1} = w^{\top}_{t+1} R\Phi_{t+1} (w_{t+1} \otimes w_{t+1}), \tag{5.6}$$

where $R\Phi_{t+1}$ is the realized coskewness matrix and $\otimes$ represents the Kronecker product. The realized coskewness matrix corresponds to N matrixes $RA_{i,t+1}$ of dimension $N \times N$ such that

$$R\Phi_{t+1} = [\ RA_{1,t+1} \mid RA_{2,t+1} \mid \cdots \mid RA_{N,t+1}\ ], \tag{5.7}$$

where

$$RA_{i,t+1} = \begin{bmatrix} ra_{i11,t+1} & ra_{i12,t+1} & \cdots & ra_{i1N,t+1} \\ ra_{i21,t+1} & ra_{i22,t+1} & \cdots & ra_{i2N,t+1} \\ \vdots & \vdots & \ddots & \vdots \\ ra_{iN1,t+1} & ra_{iN2,t+1} & \cdots & ra_{iNN,t+1} \end{bmatrix}, \tag{5.8}$$

where each element, $ra_{ijk,t+1}$, is given by

$$ra_{ijk,t+1} \;=\; \frac{1}{t}\sum_{\tau=1}^{t}\sum_{q=1}^{Q} r_{i,\tau+(q/Q)} r_{j,\tau+(q/Q)} r_{k,\tau+(q/Q)}, \tag{5.9}$$

with $i, j, k = 1, \ldots, N$.

Finally, the daily realized kurtosis (Amaya et al., 2015) at day $t+1$, for each individual stock i, can be defined as

$$rk^{Q}_{i,t+1} \;=\; \sum_{q=1}^{Q} r^{4}_{i,t+(q/Q)}. \tag{5.10}$$

The daily portfolio realized kurtosis, can be obtained by computing the following products

$$rk_{t+1} \;=\; w^{\top}_{t+1} R\Psi_{t+1}(w_{t+1} \otimes w_{t+1} \otimes w_{t+1}), \tag{5.11}$$

where $R\Psi_{t+1}$ represents the realized cokurtosis matrix. The $R\Psi_{t+1}$ matrix corresponds to N^2 matrixes $RB_{ij,t+1}$ of dimension $N \times N$ such that

$$R\Psi_{t+1} = [RB_{11,t+1} \mid RB_{12,t+1} \mid \cdots \mid RB_{1N,t+1} \mid RB_{21,t+1} \mid RB_{22,t+1} \mid \cdots \mid RB_{2N,t+1} \mid \cdots \mid RB_{N1,t+1} \mid RB_{N2,t+1} \mid \cdots \mid RB_{NN,t+1}], \tag{5.12}$$

with

$$RB_{ij,t+1} = \begin{bmatrix} rb_{ij11,t+1} & rb_{ij12,t+1} & \cdots & rb_{ij1N,t+1} \\ rb_{ij21,t+1} & rb_{ij22,t+1} & \cdots & rb_{ij2N,t+1} \\ \vdots & \vdots & \ddots & \vdots \\ rb_{ijN1,t+1} & rb_{ijN2,t+1} & \cdots & rb_{ijNN,t+1} \end{bmatrix}, \tag{5.13}$$

and where each element, $rb_{ijkl,t+1}$, is given by

$$rb_{ijkl,t+1} = \frac{1}{t}\sum_{\tau=1}^{t}\sum_{q=1}^{Q} r_{i,\tau+(q/Q)} r_{j,\tau+(q/Q)} r_{k,\tau+(q/Q)} r_{l,\tau+(q/Q)}, \tag{5.14}$$

with $i, j, k, l = 1, \ldots, N$.

As discussed before, we propose the use of intraday data to compute the realized moments as inputs of Problem (2.34):

$$\begin{aligned} V_t(r_{p,t+1}) &\approx rv_t = w_t^\top R\Sigma_t w_t, \\ S_t(r_{p,t+1}) &\approx rs_t = w_t^\top R\Phi_t (w_t \otimes w_t), \\ K_t(r_{p,t+1}) &\approx rk_t = w_t^\top R\Psi_t (w_t \otimes w_t \otimes w_t). \end{aligned} \tag{5.15}$$

Hence, Problem (2.34) can be reformulated as

$$\begin{array}{ll} \max\limits_{w_t \in \mathbb{R}^N} & \theta_1 [E_t(r_{p,t+1})] - \theta_2 [E_t(r_{p,t+1})] rv_t + \theta_3 [E_t(r_{p,t+1})] rs_t - \theta_4 [E_t(r_{p,t+1})] rk_t \\ \text{subject to} & w_t \in P_t. \end{array} \tag{5.16}$$

Notice that in estimating the daily return, $r_{i,t+1}$, of each stock i (with $i = 1, \ldots, N$), from high-frequency data, only the first and last price observations will matter:

$$
\begin{aligned}
r_{i,t+1} &= \sum_{q=1}^{Q} \left[\ln(P_{i,t+(q/Q)}) - \ln(P_{i,t+((q-1)/Q)})\right] \\
&= \left[\ln(P_{i,t+(1/Q)}) - \ln(P_{i,t})\right] + \left[\ln(P_{i,t+(2/Q)}) - \ln(P_{i,t+(1/Q)})\right] + \ldots \\
&+ \left[\ln(P_{i,t+1}) - \ln(P_{i,t+(Q-1/Q)})\right] \\
&= \left[\ln(P_{i,t+1}) - \ln(P_{i,t})\right].
\end{aligned} \tag{5.17}
$$

Consequently, the estimation of the daily portfolio mean, $E_t\left(r_{p,t+1}\right)$, is given by Equation (2.3).

5.3 On the gains of using high frequency data and higher moments in portfolio selection

In this section, motivated by the work of Brandt et al. (2009), we consider a CRRA-utility framework to incorporate, not only the first two moments of the returns distribution, but also the skewness and kurtosis into the portfolio selection problem. The methodological design is the following: firstly, for a given risk aversion level we build two utility-maximizing portfolios - one based on daily data (which we designate by low frequency portfolio, $w^{(low)}$, corresponding to the solution of Problem (2.34)) and the other based on intraday data (the high frequency portfolio, $w^{(high)}$, corresponding to the solution of Problem (5.16)); then, we compare the OOS performance of the low and high frequency portfolios for ten different risk aversion levels, using several measures (the OOS utility, mean, variance, skewness, kurtosis, Sharpe ratio and turnover).

The analysis is conducted on a dataset of fourteen stocks from the French Stock Market Index (CAC 40 Index) for a five-year period (January 1999 to December 2003). These data were provided by the EUROFIDAI (European Financial Data Institute), and were not subjected to any kind of sample selection. These fourteen stocks are constituents of the French Stock Market Index (CAC 40 Index) at the current date (July, 2017). The empirical evidence is very clear: the high frequency portfolios outperform the low frequency portfolios for every measure and for every risk aversion coefficient.

5.3.1 Empirical analysis

Data description

We compared the performance of the low frequency portfolio (solution of Problem (2.34)) with that of the high frequency portfolio (solution of Problem (5.16)), using a dataset from the CAC 40 Index (Euronext Paris). The dataset was provided by the EUROFIDAI and corresponds to fourteen French stocks (see Table 5.1). These stocks were traded during all the sample period in the French Stock Market (Euronext Paris) and belonged to the CAC 40 Index at least once (but not necessarily always). All stocks are currently constituents of the CAC 40 Index (July, 2017). For each stock, we have access to intraday data gathered during each trading session (from 09:00 a.m. to 17:30 p.m., local time) for a total of 1260 trading days, from January 1999 to December 2003. In these files, for each stock, among many other information, we just retained the transactions timestamps, the stock trading prices and the traded number of securities.

Table 5.1: The fourteen stocks from the France Stock Market Index (CAC 40)

Stock Designation	
AIR LIQUIDE	LVMH
AXA	MICHELIN
CARREFOUR	PERNOD RICARD
DANONE	SAINT-GOBAIN
ESSILOR INTL	SANOFI-AVENTIS
FRANCE TELECOM	TOTAL
L'OREAL	UNIBAL

This table lists the composition of the dataset used in the empirical analysis. The intraday data, on these stocks, were provided by the EUROFIDAI.

The intraday price observations, for the fourteen stocks, were not synchronized. Refresh-time methods used for synchronizing intraday trading among stocks include the pairwise refresh-time method and the all refresh-time method. The pairwise refresh-time method synchronizes the trading for each pair of stocks separately, allowing us to retain more data points (compared to the all refresh-time method); however, the resulting covariance matrix is not necessarily positive definite. In turn, the all refresh-time method synchronizes all stocks simultaneously and ensures that the resulting covariance matrix is positive definite (see Barndorff-Nielsen et al., 2011, for further details).

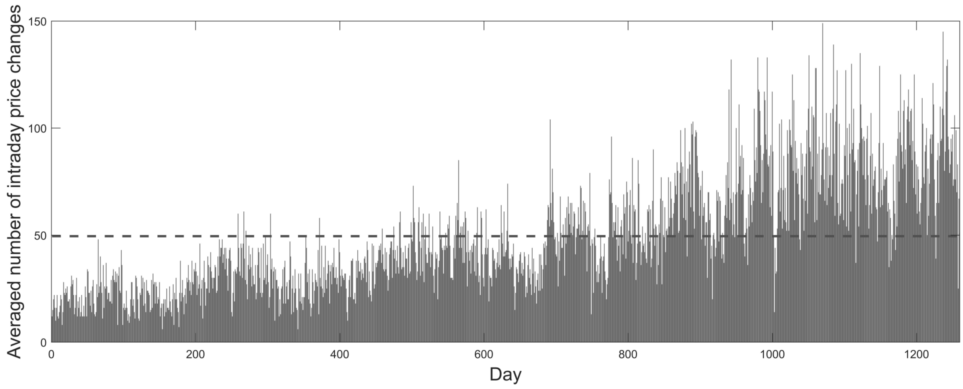
To ensure the positive-definiteness of the covariance matrix, in this Chapter we chose the all refresh-time method (Barndorff-Nielsen et al., 2011). This method was implemented through a `C++` routine. Briefly, this method can be described as follows:

- Let τ_1 be the first intraday period at day $t+1$ where all the available stocks have changed their price at least once since the market opening;
- Let τ_2 be the first intraday period at day $t+1$ where all the available stocks have changed their price at least once since τ_1;
- Proceeding in this way, allows the sequential definition of timestamps τ_q, with $q \in \{1, \dots, Q\}$, until τ_q is defined, corresponding to the market closure;
- Then we can compute the intraday returns for each stock $i \in \{1, \dots, N\}$, in irregularly spaced but perfectly synchronous intervals

$$r_{i,t+\tau_q} = \ln(P_{i,t+\tau_q}) - \ln(P_{i,t+\tau_{q-1}}), \text{ with } q = 2, \dots, Q. \tag{5.18}$$

After overcoming the nonsynchronous trading problem, we obtained an average of about 49 synchronized price changes per day, which corresponds to an average duration of around 10-minutes (see Figure 5.1). From this figure it also visible an increasing trend in the trading frequency during the period under analysis. Hereafter, when estimating the realized moments, it is assumed that microstructure noise does not exist.

Figure 5.1: Averaged (on the fourteen stocks) number of intraday price changes per day



This figure reports the averaged (over the fourteen stocks) number of changes in the intraday price observations for each day. The horizontal axis corresponds to the number of trading days. In the vertical axis is the average number of intraday price changes. The horizontal dashed line represents the averaged (on the overall sample) number of price changes per day (equal to 49.4584 price changes).

Performance of the Models

To compare the performance of the low frequency portfolio ($w^{(low)}$) with that of the high frequency portfolio ($w^{(high)}$), we used a rolling window approach (see Figure 2.5) for a total of 255 evaluation periods (days), for the year of 2003. The first estimation window ranges from the first trading day of January 1999 to the last trading day of December 2002. The low frequency portfolio (solution of Problem (2.34)) and the high frequency portfolio (solution of Problem (5.16)) were computed for ten risk aversion levels, $\gamma = 1, \ldots, 10$. From the recorded OOS daily returns for each portfolio ($w^{(low)}$ and $w^{(high)}$) we computed the OOS utility, $\widehat{g}$, given by

$$\widehat{g} = \begin{cases} \dfrac{(1+\widehat{m})^{1-\gamma} - 1}{1-\gamma} & \text{if} \quad \gamma \ > \ 1, \\ \\ \log(1+\widehat{m}) & \text{if} \quad \gamma \ = \ 1, \end{cases} \tag{5.19}$$

where $\widehat{m}$ represents the OOS mean return. The results are reported in Table 5.2. We can observe that, for all the ten different risk aversion levels, the high frequency portfolio always outperforms the low frequency portfolio in terms of OOS utility.

Table 5.2: OOS portfolio utility ($\widehat{g}$)

Risk Aversion Level	Low Frequency Portfolio ($w^{(low)}$)	High Frequency Portfolio ($w^{(high)}$)
$\gamma = 1$	-13.0449	-6.9156
$\gamma = 2$	-1.3885	4.8674
$\gamma = 3$	1.1692	9.3592
$\gamma = 4$	2.3758	12.4280
$\gamma = 5$	7.6371	17.5120
$\gamma = 6$	12.4940	20.9860
$\gamma = 7$	16.0260	22.9050
$\gamma = 8$	18.3800	24.8660
$\gamma = 9$	20.5960	25.8180
$\gamma = 10$	22.2090	26.7570

This table reports the OOS utility $\widehat{g}$ of each low frequency portfolio ($w^{(low)}$) and high frequency portfolio ($w^{(high)}$) for ten different risk aversion levels. All the OOS utility values are multiplied by a factor of 10^5.

The investor wants to achieve the portfolio with the highest mean and skewness and the lowest variance and kurtosis, therefore the superiority of the high frequency portfolios may be the result of its dominance in any of these dimensions. Strikingly, regardless of the risk aversion coefficient, the high frequency portfolio is able to outperform the low frequency portfolio in terms of OOS mean (see Table 5.3), OOS variance (see Table 5.4), OOS skewness (see Table 5.5) and OOS kurtosis (see Table 5.6).

Table 5.3: OOS portfolio mean ($\widehat{m}$)

Risk Aversion Level	Low Frequency Portfolio ($w^{(low)}$)	High Frequency Portfolio ($w^{(high)}$)
$\gamma = 1$	-13.0441	-6.9154
$\gamma = 2$	-1.3884	4.8676
$\gamma = 3$	1.1692	9.3605
$\gamma = 4$	2.3759	12.4310
$\gamma = 5$	7.6385	17.5195
$\gamma = 6$	12.4988	20.9990
$\gamma = 7$	16.0351	22.9230
$\gamma = 8$	18.3935	24.8906
$\gamma = 9$	20.6154	25.8485
$\gamma = 10$	22.2339	26.7929

This table reports the OOS mean ($\widehat{m}$) of each low frequency portfolio ($w^{(low)}$) and high frequency portfolio ($w^{(high)}$) for ten different risk aversion levels. All the $\widehat{m}$ values are multiplied by a factor of 10^5.

Table 5.4: OOS portfolio variance ($\widehat{v}$)

Risk Aversion Level	Low Frequency Portfolio ($w^{(low)}$)	High Frequency Portfolio ($w^{(high)}$)
$\gamma = 1$	22.2365	19.4203
$\gamma = 2$	19.1773	17.7894
$\gamma = 3$	18.2156	17.1653
$\gamma = 4$	17.4455	16.6957
$\gamma = 5$	16.6210	15.7214
$\gamma = 6$	16.1818	15.0494
$\gamma = 7$	15.9180	14.5053
$\gamma = 8$	15.7023	14.0070
$\gamma = 9$	15.5145	13.5888
$\gamma = 10$	15.3684	13.2560

This table reports the OOS variance ($\widehat{v}$) of each low frequency portfolio ($w^{(low)}$) and high frequency portfolio ($w^{(high)}$) for ten different risk aversion levels. All the $\widehat{v}$ values are multiplied by a factor of 10^5.

These results present a quite strong evidence in the sense that for any possible OOS performance measure, involving any of the four moments (mean, variance, skewness

Table 5.5: OOS portfolio skewness ($\widehat{s}$)

Risk Aversion Level	Low Frequency Portfolio ($w^{(low)}$)	High Frequency Portfolio ($w^{(high)}$)
$\gamma = 1$	-50.3308	-18.4611
$\gamma = 2$	-24.1452	-8.8852
$\gamma = 3$	-18.0263	-6.4945
$\gamma = 4$	-13.4717	-5.6325
$\gamma = 5$	-11.0233	-5.0495
$\gamma = 6$	-9.7057	-4.7543
$\gamma = 7$	-8.9234	-4.4501
$\gamma = 8$	-8.3188	-4.0176
$\gamma = 9$	-7.7254	-3.7165
$\gamma = 10$	-7.3257	-3.5138

This table reports the OOS skewness ($\widehat{s}$) of each low frequency portfolio ($w^{(low)}$) and high frequency portfolio ($w^{(high)}$) for ten different risk aversion levels. All the $\widehat{s}$ values are multiplied by a factor of 10^7.

Table 5.6: OOS portfolio kurtosis ($\widehat{k}$)

Risk Aversion Level	Low Frequency Portfolio ($w^{(low)}$)	High Frequency Portfolio ($w^{(high)}$)
$\gamma = 1$	77.6808	30.0367
$\gamma = 2$	34.1095	15.9259
$\gamma = 3$	24.5314	12.3103
$\gamma = 4$	17.9085	10.5034
$\gamma = 5$	14.1409	8.9305
$\gamma = 6$	12.3113	8.0342
$\gamma = 7$	11.2616	7.3969
$\gamma = 8$	10.5364	6.8483
$\gamma = 9$	9.9493	6.4514
$\gamma = 10$	9.5555	6.1645

This table reports the OOS kurtosis ($\widehat{k}$) of each low frequency portfolio ($w^{(low)}$) and high frequency portfolio ($w^{(high)}$) for ten different risk aversion levels. All the $\widehat{k}$ values are multiplied by a factor of 10^8.

and kurtosis), the high frequency portfolio always exhibits a better performance than the low frequency portfolio. For instance, we can consider the Sharpe ratio, $\widehat{SR}$ (see Equation (3.14)). When the numerator (the OOS mean) of $\widehat{SR}$ is negative, the ratio is refined according to Israelsen (2005) and corresponds to Equation (4.8). Table 5.7 presents the results for the refined Sharpe ratio. The results show that the low frequency portfolio always underperforms the high frequency portfolio, for any of the considered risk aversion level.

Table 5.7: OOS portfolio refined Sharpe ratio ($\widehat{SR}_{ref}$)

Risk Aversion Level	Low Frequency Portfolio ($w^{(low)}$)	High Frequency Portfolio ($w^{(high)}$)
$\gamma = 1$	-0.0019	-0.0010
$\gamma = 2$	-0.0002	3.6495
$\gamma = 3$	0.8663	7.1446
$\gamma = 4$	1.7988	9.6206
$\gamma = 5$	5.9249	13.9725
$\gamma = 6$	9.8255	17.1175
$\gamma = 7$	12.7095	19.0331
$\gamma = 8$	14.6785	21.0312
$\gamma = 9$	16.5510	22.1740
$\gamma = 10$	17.9350	23.2709

This table reports the OOS refined Sharpe ratios ($\widehat{SR}_{ref}$) of each low frequency portfolio ($w^{(low)}$) and high frequency portfolio ($w^{(high)}$) for ten different risk aversion levels. All the $\widehat{SR}_{ref}$ values are multiplied by a factor of 10^3.

Finally, we also compare the low and high frequency portfolio's turnover. The turnover is here defined as the average, over all time periods, of the absolute changes in weights across the N available stocks (see Equation (4.9)).

The results are reported in Table 5.8. The same pattern, presented in the previous OOS performance evaluation measures, was found, i.e., for the ten different relative risk aversion levels the high frequency portfolios outperform the low frequency portfolios. So, in the presence of proportional transaction costs, the high frequency portfolios provide a saving in trading costs, implying that the superiority of these portfolios increase after trading cost are taking into account.

We also highlight that, for all the performance evaluation measures, a surprising pattern was found: the OOS performances, both for the low and high frequency portfolios are increasing functions of the risk aversion level (γ). A possible explanation for this puzzling pattern may lie on the fact that with the increase of the risk aversion level, the constructed portfolios become closer to the minimum variance portfolio. The minimum variance portfolio tends to exhibit a superior OOS performance (see, e.g., Jagannathan and Ma, 2003; DeMiguel et al., 2009b). Furthermore, it has been

Table 5.8: Portfolio turnover

Risk Aversion Level	Low Frequency Portfolio ($w^{(low)}$)	High Frequency Portfolio ($w^{(low)}$)
$\gamma = 1$	0.0952	0.0797
$\gamma = 2$	0.0640	0.0575
$\gamma = 3$	0.0593	0.0508
$\gamma = 4$	0.0550	0.0446
$\gamma = 5$	0.0483	0.0382
$\gamma = 6$	0.0427	0.0335
$\gamma = 7$	0.0379	0.0319
$\gamma = 8$	0.0348	0.0300
$\gamma = 9$	0.0319	0.0277
$\gamma = 10$	0.0291	0.0254

This table reports the turnover of each low frequency portfolio ($w^{(low)}$) and high frequency portfolio ($w^{(high)}$) for ten different risk aversion levels.

documented in the literature that stocks with higher variance tend to underperform low variance stocks (Baker et al., 2011).

5.4 Portfolio choice with high frequency data: CRRA preferences and the liquidity effect

In this section, we suggest the construction of expected utility-liquidity portfolios. By solving the proposed expected utility-liquidity (EU/L) problem, the investor will be able to identify the portfolios, which have the maximum expected utility among all that provide at least a certain expected liquidity level. We also assume that the investor has a CRRA-utility. However, it is worth noticing that the proposed methodology is applicable to any other type of utility function. In this study, we consider the fourth order Taylor expansion of the expected utility, around the portfolio expected return. Thus the expected utility is a function of the portfolio expected return, variance, skewness and kurtosis. Relying on intraday transaction data, we use the daily estimates of the portfolio's moments as inputs for the optimization model, using as estimators the portfolio realized variance, realized skewness and realized kurtosis. In addition, since we are interested on the relationship between liquidity and the behaviour of stock prices, following Goyenko et al. (2009) and Chiang and Zheng (2015), the daily illiquidity level is measured by the intraday Amihud illiquidity ratio (Amihud, 2002).

For the empirical application we also use intraday data on the same set of fourteen French stocks used in the previous section. Nevertheless, we extend the time period to the maximum length which we have access, i.e., seven years (from January 1999 to December 2005, which corresponds to 14.5GB of raw data). IS we compute the EU/L Pareto frontier for a moderately risk averse investor. The EU/L Pareto frontier shows the existence of a positive relationship between the expected utility and the expected illiquidity. OOS, we compute three different EU/L optimal portfolios (according to three different, pre-established, illiquidity levels) and compare their performance with two hard to beat benchmark portfolios: the minimum variance and the equally weighted portfolios. The EU/L optimal portfolios are very competitive and in most cases are able to consistently beat the benchmarks. This is observable in terms of OOS utility, liquidity and certainty equivalent. These results hold for different risk aversion levels, which indicates that the proposed EU/L model is quite robust.

5.4.1 The EU/L Problem

As referred by Goyenko et al. (2009) and Chiang and Zheng (2015), the Amihud illiquidity ratio is one of the best proxies to measure stock liquidity since it has a strong correlation with several other measures of liquidity. We define the Amihud illiquidity ratio for stock i at day $t+1$, $AR_{i,t+1}$ as

$$AR_{i,t+1} = \frac{1}{Q}\sum_{q=1}^{Q}\frac{|r_{i,t+(q/Q)}|}{vol_{i,t+(q/Q)}}, \tag{5.20}$$

where $r_{i,t+(q/Q)}$ represents the return of stock i in the intraday period q, and $vol_{i,t+(q/Q)}$ the corresponding trading volume in euros.

In this section, we suggest the construction of efficient portfolios, where the investor maximize her expected utility while taking into account the liquidity level associated to those portfolios. Thereby, motivated by Problem (5.16) and using the Amihud illiquidity ratio as a liquidity measure, we propose the following expected utility-liquidity (EU/L) problem

$$\begin{aligned} \max_{w_t \in \mathbb{R}^N} \quad & \theta_1 \left[E_t\left(r_{p,t+1}\right)\right] - \theta_2 \left[E_t\left(r_{p,t+1}\right)\right] rv_t + \theta_3 \left[E_t\left(r_{p,t+1}\right)\right] rs_t - \theta_4 \left[E_t\left(r_{p,t+1}\right)\right] rk_t \\ \text{subject to} \quad & \iota_t^\top w_t \ \leq \ \iota_{target}, \\ & w_t \in P_t, \end{aligned} \tag{5.21}$$

where ι_t represents the vector of dimension $N \times 1$, with elements equal to the expected Amihud illiquidity ratio (computed according to Equation (5.20)) of each stock $i = 1, \ldots, N$, and ι_{target} is a given illiquidity upper limit. The objective function of the EU/L problem (Problem (5.21)) is a continuous nonlinear but smooth function, all constraints are linear and the feasible space is compact (it is a bounded and closed space). Given these properties, the existence of a maximum for the EU/L problem is guaranteed by the well-known Weierstrass theorem. By solving Problem (5.21) for different values of ι_{target}, which can be done by using any standard nonlinear optimization software for constrained optimization, one can identify the efficient EU/L frontier, i.e. those portfolios which have the maximum expected utility among all feasible portfolios that provide at least a certain level of expected liquidity.

5.4.2 Empirical Application

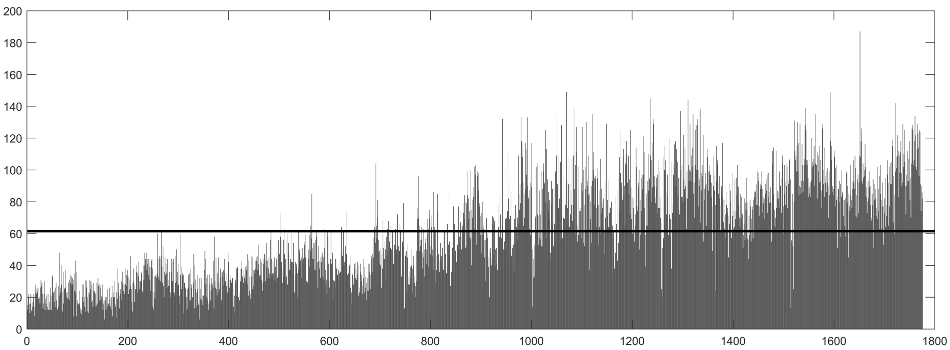
The Data

The empirical exercise was conducted over the same set of stocks presented in Section 5.3.1, i.e., fourteen French stocks (see Table 5.1) from the CAC 40 Index (the data was provided by the EUROFIDAI). Nevertheless, in this case we have access to intraday data for a total of 1777 trading days, from January 1999 to December 2005. Since the intraday price observations were not synchronized, we used the all refresh-time method (Barndorff-Nielsen et al., 2011) as presented in Section 5.3.1. After the synchronization procedure, there are on average about 61 prices changes per day (see Figure 5.2), which corresponds to an average trading frequency of 8-minutes. From Figure 5.2 it also visible that the trading intensity has increased on average about five times during the period under scrutiny.

The proposed methodology (described in Section 5.4.1) was implemented in `MATLAB`. By default, `MATLAB` has a 16 digit precision, which ensures the inexistence of relevant rounding errors when computing the realized moments and co-moments[13]. Additionally, we have computed the condition number of the realized covariance (Equation (5.4)), realized coskewness (Equation (5.7)) and realized cokurtosis (Equation (5.12)) matrixes (for all the sample period) and we have obtained the values of 8.56, 10.26 and 22.88, respectively. This suggests that the estimates of the realized moments and

[13]The most critical co-moment, in terms of possible rounding errors, is cokurtosis, since it involves summing returns raised to the fourth power. Concerning this co-moment, we start by noticing that, in the dataset, we have double digit stock prices (no higher). With double digit stocks prices and ticks of 1c, high frequency returns can be as low as 10^{-4} and their fourth power can be of the 10^{-16} order. In turn, the highest value that the realized cokurtosis takes is of the 10^{-4} order. Therefore, since we work with a 16 digit precision, in the computation of the realized cokurtosis at least 4 significant digits of the fourth power of the high frequency returns are preserved.

Figure 5.2: Averaged number of intraday price changes



This figure reports the averaged (over the fourteen stocks) number of changes in the intraday price observations for each day. The horizontal axis corresponds to the number of trading days. In the vertical axis is the average number of intraday price changes. The horizontal line represents the averaged (on the overall sample) number of price changes per day (equal to 61.4347 price changes).

co-moments are relatively stable.

We can look to the EU/L problem (Problem (5.21)) as a biobjective problem

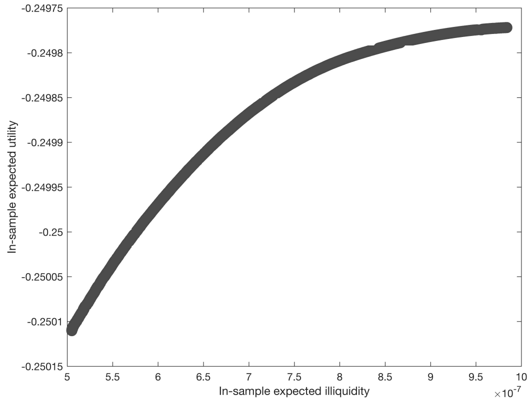
$$\begin{array}{ll} \max\limits_{w_t \in \mathbb{R}^N} & \theta_1 \left[E_t\left(r_{p,t+1}\right)\right] - \theta_2 \left[E_t\left(r_{p,t+1}\right)\right] rv_t + \theta_3 \left[E_t\left(r_{p,t+1}\right)\right] rs_t - \theta_4 \left[E_t\left(r_{p,t+1}\right)\right] rk_t \\ \min\limits_{w_t \in \mathbb{R}^N} & \iota^\top w_t \\ \text{subject to} & w_t \in P_t. \end{array} \tag{5.22}$$

The solution of Problem (5.22) is given in the form of a Pareto frontier in the expected utility-illiquidity space, allowing the investor to directly analyze the efficient tradeoff between these two dimensions. Problem (5.22) can be solved using a MOO algorithm. Motivated by previous works (see Brito et al., 2016; 2017c) and since the first objective, $\theta_1 \left[E_t\left(r_{p,t+1}\right)\right] - \theta_2 \left[E_t\left(r_{p,t+1}\right)\right] rv_t + \theta_3 \left[E_t\left(r_{p,t+1}\right)\right] rs_t - \theta_4 \left[E_t\left(r_{p,t+1}\right)\right] rk_t$, is a highly nonlinear function, we have decided to use a derivative-free solver based on DMS (see Appendix B).

Setting the IS period equal to all the available time window (January 1999 to December 2005), we applied the solver `dms` (see Appendix B) to determine the EU/L Pareto frontier. Figure 5.3 contains the plot of the EU/L Pareto frontier for an investor with a constant relative risk aversion parameter equal to 5 (see Equation (2.19)).

From the analysis of Figure 5.3, we can observe a positive relationship between the

Figure 5.3: EU/L Pareto frontier



This figure reports the solution of the EU/L biobjective Problem (Problem (5.22)). The vertical axis corresponds to the first objective function (expected utility) and the horizontal axis represents the second objective function (expected illiquidity). This solution is for a moderate risk aversion parameter ($\gamma = 5$).

expected utility and the expected illiquidity[14].

Portfolio constraints commonly used in practice (e.g. turnover constraints, buy-in threshold constraints, cardinality constraints) can be introduced and explored in the proposed EU/L model. For example, if we introduce a turnover constraint[15] we obtain a different EU/L frontier (see Figure 5.4). From the comparison between the EU/L Pareto frontier (see Figure 5.3) and the EU/L Pareto frontier with a turnover constraint (see Figure 5.4), it is possible to verify that the inclusion of a turnover constraint leads to a deterioration of the portfolios in the two considered objectives (expected utility and expected illiquidity) although the frontier maintains its convexity.

OOS Results and Sensitivity to Risk Aversion

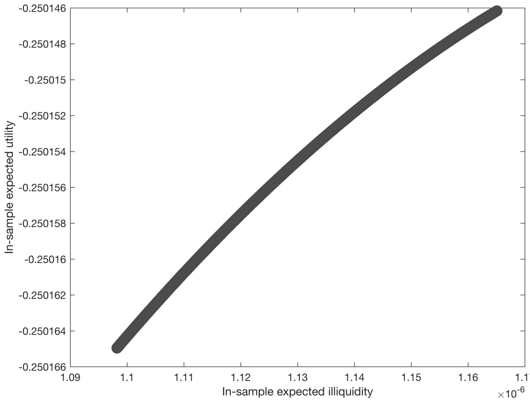
In order to analyze the robustness of the efficient portfolios on the EU/L space, this subsection compares the OOS performance of three different EU/L optimal portfolios in relation to two benchmark portfolios.

The three chosen EU/L optimal portfolios are: $w_{\iota_{10}}$, $w_{\iota_{50}}$ and $w_{\iota_{90}}$, the portfolios that correspond to the solution of Problem (5.21) for $\iota_{target} = \iota_{10}$, $\iota_{target} = \iota_{50}$ and $\iota_{target} = \iota_{90}$, respectively, where ι_y represents the yth percentile of the expected illiq-

[14]The same pattern was found for two different choices of the relative risk aversion parameter ($\gamma = 1$ and $\gamma = 10$).

[15]A turnover constraint can be formulated as $\sum_{i=1}^{N} |w_{i,t+1} - w^0_{i,t}| \leq h$, where $w^0_{i,t}$ is the reference portfolio and h is the turnover upper bound.

Figure 5.4: EU/L Pareto frontier with a turnover constraint



This figure reports the solution of the EU/L biobjective Problem (Problem (5.22)) with the inclusion of a turnover constraint (the reference portfolio was the equally weighted portfolio and the turnover upper bound was set to 5%). The vertical axis corresponds to the first objective function (expected utility) and the horizontal axis represents the second objective function (expected illiquidity). This solution is for a moderate risk aversion parameter ($\gamma = 5$).

uidity across all stocks. Hence, by construction, the aggregate level of liquidity of the resulting efficient portfolios diminishes with the increasing percentile.

Previous works, such as Jagannathan and Ma (2003); DeMiguel et al. (2009b); Brito et al. (2016), have showed the good OOS performance of the well-known minimum variance portfolio, *mv* portfolio (solution of Problem (2.16)), and the equally weighted portfolio, *ew* portfolio (defined by Equation (2.15)). In this study, we have thus decided to use these two portfolios as benchmark portfolios.

We used a rolling window approach (see Figure 2.5) for the OOS performance evaluation. We considered an estimation window of 1520 days, while the remaining 257 days are used for evaluating the OOS performance measures. The first estimation window is from January 1999 to December 2004, and January 3, 2005 is the first day where we evaluate the out-of-sample performance. For each estimation window, we computed the two benchmark portfolios, *mv* portfolio and *ew* portfolio, using the daily returns and three optimal EU/L portfolios, $w_{\iota_{10}}$, $w_{\iota_{50}}$ and $w_{\iota_{90}}$, for each of three different levels of risk aversion ($\gamma = 1$, $\gamma = 5$ and $\gamma = 10$), using intraday data. Then each portfolio was held fixed and its daily returns were observed over the next day. The estimation window was then moved forward one day, and the daily returns were computed for the next day of the evaluation period. The process was thus repeated until exhausting the 257 trading days of 2005.

Table 5.9 presents some OOS descriptive statistics. The returns of all portfolios present above normal kurtosis and are skewed. For $\gamma = 1$ and $\gamma = 5$, all the EU/L portfolios present a higher OOS mean than the benchmark portfolios (with higher OOS standard deviation, also). It is interesting to notice that the most liquid EU/L portfolio, $w_{\iota_{10}}$, has a higher OOS mean than the benchmark portfolios, for two different risk aversion levels ($\gamma = 1$ and $\gamma = 5$).

Table 5.9: OOS descriptive statistics

Descriptive statistics	$\widehat{m}$	$\widehat{\sigma}$	$\widehat{s}_{stand}$	$\widehat{k}_{stand}$
Benchmark portfolios				
mv portfolio	0.00072746	0.0068	-0.0407	3.3707
ew portfolio	0.00069979	0.0064	-0.1181	3.5758
EU/L portfolios				
$\gamma = 1$				
$w_{\iota_{10}}$	0.00077098	0.0085	0.2823	4.5101
$w_{\iota_{50}}$	0.00075402	0.0084	0.2350	4.5533
$w_{\iota_{90}}$	0.00077375	0.0085	0.2836	4.5155
$\gamma = 5$				
$w_{\iota_{10}}$	0.00074280	0.0075	-0.0281	3.4576
$w_{\iota_{50}}$	0.00074384	0.0075	-0.0260	3.4545
$w_{\iota_{90}}$	0.00074411	0.0075	-0.0258	3.4547
$\gamma = 10$				
$w_{\iota_{10}}$	0.00070052	0.0073	-0.2337	3.7339
$w_{\iota_{50}}$	0.00071437	0.0073	-0.2407	3.7289
$w_{\iota_{90}}$	0.00070040	0.0073	-0.2340	3.7342

This table reports the OOS mean ($\widehat{m}$), standard deviation ($\widehat{\sigma}$), standardized skewness ($\widehat{s}_{stand}$) and standardized kurtosis ($\widehat{k}_{stand}$) for each portfolio. The benchmark portfolios, mv and ew portfolios, refer to the minimum variance and equally weighted portfolios, computed using daily data. The three portfolios $w_{\iota_{10}}$, $w_{\iota_{50}}$ and $w_{\iota_{90}}$, denote the optimal expected utility/liquidity portfolios considering percentiles 10, 50 and 90 of the overall illiquidity spectrum across all stocks, and are computed using intraday data. For the computation of the EU/L optimal portfolios we considered three different levels of risk aversion: $\gamma = 1$, $\gamma = 5$ and $\gamma = 10$.

Given the time series of daily OOS returns for each portfolio (two benchmark portfolios and nine EU/L optimal portfolios), we computed three performance evaluation measures. The OOS utility, $\widehat{u}$, defined as

$$\hat{u} = \begin{cases} \dfrac{(1+\widehat{m})^{1-\gamma}}{1-\gamma} & \text{if} \quad \gamma > 1, \\ \\ \log(1+\widehat{m}) & \text{if} \quad \gamma = 1, \end{cases} \tag{5.23}$$

where $\widehat{m}$ corresponds to the OOS mean return. The OOS illiquidity, $\widehat{\iota}$, defined as the averaged OOS portfolio illiquidity, and the certainty equivalent, $\widehat{CE}$ (see Equation(4.7)). Recall that, the certainty equivalent can be interpreted as the risk-free rate that an investor is willing to accept in order to give up a particular risky portfolio.

The results for these three performance evaluation measures are presented in Tables 5.10 to 5.12.

The three EU/L optimal portfolios ($w_{\iota_{10}}$, $w_{\iota_{50}}$ and $w_{\iota_{90}}$) have a consistently lower illiquidity level than the benchmark portfolios (the *mv* and *ew* portfolios), for any of the considered risk aversion levels. A possible explanation for this result is that liquidity has a persistent nature, thus the most liquid stocks *ex-ante* tend to be the most liquid ones *ex-post.* This result highlights the robustness of the EU/L model in building reliably liquid portfolios.

The robustness of the EU/L model is also reflected in the OOS utility, $\widehat{u}$, results. For $\gamma = 1$ and $\gamma = 5$, all the EU/L portfolios consistently show a significant (at a 5% significance level) higher utility than the benchmarks portfolios. In turn, when the investor is more sensitive to losses, the case of $\gamma = 10$, all the EU/L portfolios significantly underperform one of the benchmark portfolios (the *mv* portfolio) and slightly outperform the other (the *ew* portfolio).

In terms of OOS certainty equivalent return, $\widehat{CE}$, for a low ($\gamma = 1$) and a moderate ($\gamma = 5$) risk aversion levels, the three chosen EU/L portfolios present a competitive certainty equivalent (compared with the benchmark portfolios). However, this pattern does not hold for $\gamma = 10$: here the EU/L portfolios clearly underperform the benchmark portfolios in terms of certainty equivalent return. We must note that both the utility and the certainty equivalent do not take the liquidity into account; since the EU/L portfolios must also take into account liquidity, it is not surprising that they underperform (or at least some of them underperform) the benchmark portfolios.

In addition to the three aforementioned performance measures, we decided to include in this OOS analysis the Sharpe ratio[16] (see Equation (3.14)) and net Sharpe ratio[17] (see Equation (4.11)). Since the Sharpe ratio is one of the most referenced OOS

[16]When the numerator was negative, the ratio was refined in order to achieve a correct rank of the portfolios, according the Israelsen (2005) methodology (see Equation (4.8)).

[17]As commonly assumed in the literature (see, e.g., Balduzzi and Lynch, 1999; DeMiguel et al.,

Table 5.10: OOS performance evaluation for $\gamma = 1$

Performance measures	$\widehat{u}$	$\widehat{\iota}$	$\widehat{CE}$
Benchmark portfolios			
mv portfolio	7,271.9328	3.7452	7,041.6351
ew portfolio	6,995.4703	4.7819	6,790.0271
EU/L portfolios			
$w_{\iota_{10}}$	7,706.8555	2.7257	7,346.4346
	$(0.0079)^{mv}(0.0024)^{ew}$	$(0.0000)^{mv}(0.0000)^{ew}$	$(0.0028)^{mv}(0.0017)^{ew}$
$w_{\iota_{50}}$	7,537.3783	2.7632	7,186.4844
	$(0.0204)^{mv}(0.0021)^{ew}$	$(0.0000)^{mv}(0.0000)^{ew}$	$(0.0049)^{mv}(0.0191)^{ew}$
$w_{\iota_{90}}$	7,734.4585	2.7253	7,374.1947
	$(0.0011)^{mv}(0.0045)^{ew}$	$(0.0000)^{mv}(0.0000)^{ew}$	$(0.0041)^{mv}(0.0054)^{ew}$

This table reports the OOS utility ($\widehat{u}$), illiquidity ($\widehat{\iota}$) and certainty equivalent ($\widehat{CE}$), for each portfolio. All the presented OOS values are multiplied by a factor of 10^7. In parenthesis are the bootstrap p-values of the difference between the respective performance measure of each EU/L portfolio and the benchmark portfolios, denoted by the superscripts mv and ew, respectively. These bootstrap p-values were computed according the classical methodology proposed by Efron and Tibshirani (1994).

Table 5.11: OOS performance evaluation for $\gamma = 5$

Performance measures	$\widehat{u}$	$\widehat{\iota}$	$\widehat{CE}$
Benchmark portfolios			
mv portfolio	-2,492,738.6332	3.7452	6,109.4671
ew portfolio	-2,493,014.3079	4.7819	5,957.4286
EU/L portfolios			
$w_{\iota_{10}}$	-2,492,585.7349	3.1594	6,019.1712
	$(0.0075)^{mv}(0.0030)^{ew}$	$(0.0000)^{mv}(0.0000)^{ew}$	$(0.0524)^{mv}(0.0918)^{ew}$
$w_{\iota_{50}}$	-2,492,575.3643	3.1595	6,029.8099
	$(0.0051)^{mv}(0.0000)^{ew}$	$(0.0000)^{mv}(0.0000)^{ew}$	$(0.0599)^{mv}(0.0353)^{ew}$
$w_{\iota_{90}}$	-2,492,572.6882	3.1596	6,032.4879
	$(0.0049)^{mv}(0.0000)^{ew}$	$(0.0000)^{mv}(0.0000)^{ew}$	$(0.0847)^{mv}(0.0141)^{ew}$

This table reports the OOS utility ($\widehat{u}$), illiquidity ($\widehat{\iota}$) and certainty equivalent ($\widehat{CE}$), for each portfolio. All the presented OOS values are multiplied by a factor of 10^7. In parenthesis are the bootstrap p-values of the difference between the respective performance measure of each EU/L portfolio and the benchmark portfolios, denoted by the superscripts mv and ew, respectively. These bootstrap p-values were computed according the classical methodology proposed by Efron and Tibshirani (1994).

performance evaluation measures in the literature and the net Sharpe ratio allows the

2009b; DeMiguel and Olivares-Nadal, 2016) and according to Equation (4.12), we set the proportional transaction costs equal to 50 basis points per transaction. Thus the cost of a trade over all stocks is

Table 5.12: OOS performance evaluation for $\gamma = 10$

Performance measures	$\widehat{u}$	$\widehat{\iota}$	$\widehat{CE}$
Benchmark portfolios			
mv portfolio	-1,103,862.9229	3.7452	4,943.2142
ew portfolio	-1,104,137.6161	4.7819	4,914.1740
EU/L portfolios			
$w_{\iota_{10}}$	-1,104,130.3797	3.4059	4,318.0065
	$(0.0000)^{mv}(0.2212)^{ew}$	$(0.0000)^{mv}(0.0000)^{ew}$	$(0.0000)^{mv}(0.0000)^{ew}$
$w_{\iota_{50}}$	-1,104,130.8790	3.2626	4,316.6101
	$(0.0000)^{mv}(0.3077)^{ew}$	$(0.0000)^{mv}(0.0000)^{ew}$	$(0.0000)^{mv}(0.0000)^{ew}$
$w_{\iota_{90}}$	-1,104,130.1621	3.2626	4,317.1349
	$(0.0000)^{mv}(0.2402)^{ew}$	$(0.0000)^{mv}(0.0000)^{ew}$	$(0.0000)^{mv}(0.0000)^{ew}$

This table reports the OOS utility ($\widehat{u}$), illiquidity ($\widehat{\iota}$) and certainty equivalent ($\widehat{CE}$), for each portfolio. All the presented OOS values are multiplied by a factor of 10^7. In parenthesis are the bootstrap p-values of the difference between the respective performance measure of each EU/L portfolio and the benchmark portfolios, denoted by the superscripts mv and ew, respectively. These bootstrap p-values were computed according the classical methodology proposed by Efron and Tibshirani (1994).

analysis of the impact of transaction costs, we decided to investigate how the EU/L portfolios behave in these two performance evaluation measures comparatively to the benchmark portfolios.

Table 5.13 shows the results for these two additional performance evaluation measures, the Sharpe ratio ($\widehat{SR}$) and the net Sharpe ratio ($\widehat{SR}_{tc}$). Regarding the Sharpe ratios, we can observe that it is for a moderate risk aversion level ($\gamma = 5$) that the EU/L portfolios achieve the highest values. However, for the different risk aversion levels the benchmark portfolios always outperform the EU/L portfolios. As we noticed for the utility and the certainty equivalent measures, this is not surprising, since the EU/L portfolios must also take into account liquidity and higher moments.

When we consider the transaction costs (by computing the net Sharpe ratio), we observe that, all the EU/L portfolios are not able to beat the two benchmark portfolios. This was somehow expected, since the EU/L portfolios present a higher turnover (see Equation (4.9)) than the benchmarks (see Table 5.14). It is well-known in the literature that the mv and ew portfolios present a much lower turnover than other alternative strategies (see, e.g, DeMiguel et al., 2009b).

defined as

$$tc = \sum_{t=1}^{\#periods} 0.5\% \sum_{i=1}^{N} \left(\left|w_{i,t+1} - w^h_{i,t}\right|\right).$$

Table 5.13: OOS Sharpe ratios and net refined Sharpe ratios

	Sharpe ratios	Net refined Sharpe ratios
Benchmark portfolios		
mv portfolio	1,031.7464	-0.6939
ew portfolio	1,053.0251	-0.5152
EU/L portfolios		
$\gamma = 1$		
$w_{\iota_{10}}$	862.2577	-7.8601
	$(0.2713)^{mv}(0.3219)^{ew}$	$(0.0010)^{mv}(0.0010)^{ew}$
$w_{\iota_{50}}$	854.8575	-11.9545
	$(0.2627)^{mv}(0.3189)^{ew}$	$(0.0010)^{mv}(0.0010)^{ew}$
$w_{\iota_{90}}$	865.6784	-8.0057
	$(0.2591)^{mv}(0.3291)^{ew}$	$(0.0010)^{mv}(0.0010)^{ew}$
$\gamma = 5$		
$w_{\iota_{10}}$	952.1998	-2.8132
	$(0.2855)^{mv}(0.5435)^{ew}$	$(0.0010)^{mv}(0.0010)^{ew}$
$w_{\iota_{50}}$	953.6620	-2.8126
	$(0.3095)^{mv}(0.5525)^{ew}$	$(0.0010)^{mv}(0.0010)^{ew}$
$w_{\iota_{90}}$	954.0161	-2.8065
	$(0.2933)^{mv}(0.5559)^{ew}$	$(0.0010)^{mv}(0.0010)^{ew}$
$\gamma = 10$		
$w_{\iota_{10}}$	921.3735	-1.6135
	$(0.2130)^{mv}(0.3673)^{ew}$	$(0.0010)^{mv}(0.0010)^{ew}$
$w_{\iota_{50}}$	940.7258	-2.5132
	$(0.2649)^{mv}(0.4453)^{ew}$	$(0.0010)^{mv}(0.0010)^{ew}$
$w_{\iota_{90}}$	921.0850	-1.6095
	$(0.2175)^{mv}(0.3703)^{ew}$	$(0.0010)^{mv}(0.0010)^{ew}$

This table reports the OOS Sharpe ratio ($\widehat{SR}$) and net Sharpe ratio, ($\widehat{SR}_{tc}$), for each portfolio. All the OOS Sharpe ratios and net Sharpe ratios are multiplied by a factor of 10^4. Note that, for the net Sharpe ratios case, the presented values refer to the refined ratios according to the Israelsen (2005) methodology. In parenthesis are the bootstrap p-values of the difference between the respective performance measure of each EU/L portfolio and those of the benchmarks: the mv portfolio and the ew portfolio, in the first and second parenthesis, respectively. These p-values were computed according the robust methodology, developed specifically for the Sharpe ratio, of Ledoit and Wolf (2008).

Table 5.14: Portfolio turnover

	Turnover
Benchmark portfolios	
mv portfolio	0.8462
ew portfolio	0.6747
EU/L portfolios	
$\gamma = 1$	
$w_{\iota_{10}}$	7.2378
$w_{\iota_{50}}$	11.1241
$w_{\iota_{90}}$	7.3724
$\gamma = 5$	
$w_{\iota_{10}}$	2.9733
$w_{\iota_{50}}$	2.9729
$w_{\iota_{90}}$	2.9666
$\gamma = 10$	
$w_{\iota_{10}}$	1.7698
$w_{\iota_{50}}$	2.7296
$w_{\iota_{90}}$	1.7653

This table reports the portfolio turnover for each portfolio. All the turnover values are multiplied by a factor of 10^2. *The portfolio turnover is computed according Equation (4.9).*

5.5 Conclusions

Nowadays the use of big data seems to offer a competitive advantage in many fields. Particularly in Finance, the increasing availability of huge sets of high frequency financial data encourages the emergence of new investment strategies built on all that information.

In this chapter, we have analyzed the practical benefits of using intraday information in portfolio choice. We have considered a general framework where the investor wants to maximize her CRRA-utility. The expected utility was modeled using not only the two first moments of the returns distribution but also the skewness and the kurtosis. Within this framework, for a given risk aversion level, we have constructed two portfolios: a low frequency portfolio, solution of the portfolio choice problem where the inputs are obtained from daily data, and a high frequency portfolio, solution of the portfolio choice problem where the inputs are obtained from intraday data.

The empirical results, based on fourteen stocks from the CAC 40 Index, showed a superior daily OOS performance of the high frequency portfolio over the low frequency portfolio. For ten different risk aversion levels, each high frequency portfolio outperformed the corresponding low frequency portfolio in terms of several OOS measures (utility, mean, variance, skewness, kurtosis, Sharpe ratio and turnover). This empirical evidence suggests the existence of practical real gains when high frequency data is taken into account in portfolio choice. This is in accordance with one elementary principle in statistics: *ceteris paribus*, more data is desirable to less.

In this chapter we also proposed a new methodology for portfolio choice. We suggest that the investor may build her portfolios according to the utility maximization criteria, but, at the same time, taking into account a desired level of liquidity. The proposed EU/L model allows the investor to identify the portfolios which have the maximum expected utility among all that provide at least a certain expected level of liquidity. The investor can thus directly analyze the efficient tradeoff between expected utility and expected liquidity and, accordingly, make her choices in the expected utility-liquidity space.

The empirical application, using high frequency data on fourteen stocks from the CAC 40 Index, showed a positive relationship between the expected utility and the expected illiquidity.

The analysis of the OOS performance, for different levels of risk aversion, revealed that the EU/L portfolios are usually competitive with the minimum variance and equally weighted portfolios. These two benchmarks were always beaten in terms of liquidity, but they sometimes performed better in terms of utility and certainty equivalent, and they were always better in terms of Sharpe ratio. This shows that, in order to achieve a higher liquidity, the EU/L portfolios must sometimes sacrifice other performance measures. Finally, the results for the net Sharpe ratio show that the EU/L portfolios tend to exhibit a higher turnover than the two considered benchmark portfolios.

The sample used in the EU/L analysis, already comprises a period of sharp decreasing prices (the dot-com bubble in 2000). As future work, it would be interesting to test the proposed model in face of more recent "black swan events" (the global financial crisis (2008-2009) and the sovereign debt crisis in Europe (2011)). Although it is not easy to have access to reliable intraday data, as future work we would like to test the informational impact of using high frequency data in comparison with using daily data, especially after these events. We are also interested in refining the computation of trading costs and to test the model's robustness to other utility functions (beyond the considered CRRA-utility).

Chapter 6

Final conclusions

With this thesis we tried to contribute, theoretically and empirically, to the literature on portfolio choice under uncertainty. Motivated by the innumerable empirical evidence against the Gaussian returns' distribution assumption, we emphasized the importance of considering higher moments (instead of just considering the mean and the variance) in portfolio choice. In this way we can avoid the Gaussian returns' distribution assumption, which is breached in several financial markets. Following Brandt et al. (2009), we have shown how straightforward it is to incorporate higher moments in the investor's utility maximization problem, with the investor's preferences being characterized by a CRRA-utility. Thus, we dealt with the return and the risk in a different way from what is traditionally done in the MPT. Based on the existing literature, we argued that the rational investor is not only interested in maximizing the portfolio return and minimizing the portfolio variance, but he also seeks to maximize the portfolio skewness and to minimize the portfolio kurtosis.

We began by discussing the limitations of MPT, in Chapter 2. We discussed the need to abandon the Gaussian returns' distribution assumption, when the data tell us that it does not fit reality. The MV optimization model, being based on this assumption is a flawed model. In addition, the documented (in several studies) poor OOS performance of the MV optimization model, suggests that we have to develop other methodologies that lead to the birth of a "new" MPT. Based on these two main limitations, we can say that the MV optimization model has weak descriptive and predictive powers. In order to overcome these limitations, we proposed the inclusion of higher moments (namely, skewness and kurtosis) in the investor's portfolio choice problem. In the subsequent chapters of this thesis, we discussed the potential of some methodologies, all of which considering the inclusion of higher moments.

In Chapter 3 we proposed a methodology to construct portfolios in a skewness/ semivariance bidimensional space. The use of skewness is justified by the intuitive investor's preference for positive skewness (limited losses and higher probability of an extreme favorable outcome) and by her willingness to sacrifice expected return for positive skewness (reason why we have not considered the expected return as an objective in the proposed model). In turn, the use of the portfolio semivariance in this model shows two important aspects: 1) assuming that just one measure is used to quantify the portfolio risk, there are more plausible measures than the traditional portfolio variance to quantify this risk; 2) the reference point (benchmark return in the semivariance computation) used to define losses is defined according to the investor's beliefs (risk has a subjective nature!). We dealt with the semivariance explicitly and we overcame the endogeneity problem of the cosemivariance matrix through a derivative-free algorithm. The empirical application evidenced a competitive OOS performance of the portfolios constructed according to the skewness/semivariance biobjective model, for three different benchmark returns (used in the computation of the semivariance). Some of the skewness/semivariance portfolios were even able to consistently outperform the equally weighted portfolio in terms of Sharpe ratio. Thereby, the skewness/semivariance model offers a way to directly analyze the tradeoff between skewness and semivariance, with the ability to generate portfolios exhibiting a competitive OOS performance. The model is very flexible, allowing the investor to choose the benchmark return according to her own beliefs. Given this flexibility and the promising results, as future work, this model can be deeply explored.

In Chapter 4 we extended the study of the cardinality impact, on portfolio performance, from a MV framework to MVS and MVSK frameworks. The models corresponding to the frameworks with higher moments (the MVS and MVSK frameworks) constitute an innovative contribution to the existing literature, and their potential should be explored in future studies. Furthermore, all the proposed models, for each framework, lead to NP-hard problems, and other algorithms (in addition to the used DMS algorithm) must be tested on these problems and must be compared in terms of algorithmic efficiency. Regarding the analysis that we have conducted, the empirical results showed that cardinality plays an important role in every one of the considered frameworks. For each cardinality level, the portfolios formed in each framework presented a similar performance (in terms of certainty equivalent and Sharpe ratio) IS while, OOS, we found evidence of gains in using higher moments, for certain cardinality levels. We have thus drawn attention to the fact that cardinality can be seen as a non-negligible source of risk.

Finally, in Chapter 5 we began by studying the potential gains of using high frequency data (intraday data), through the joint use of the realized variance, realized skewness and realized kurtosis. In direct comparison with the use of low frequency data (daily data), the analysis showed, for different risk aversion levels and for several performance evaluation measures, the superiority (in terms of OOS performance) of the portfolios formed according the methodology based on high frequency data. It is important to highlight that the analysis was only conducted on a particular set of stocks, and hence it is not possible to draw very general conclusions about the benefits of using high frequency data. Nevertheless, the strong pattern found in favor of the use of high frequency data, in this specific case, contributes to the existing literature that suggests the existence of significant gains in using high frequency data. Looking to liquidity as another important source of risk, in the second part of Chapter 5 we have introduced, in the investor's problem with higher realized moments, a liquidity constraint. According to the suggested methodology, the investor can make her choices directly in the expected utility/liquidity bidimensional space. The empirical results are promising, showing that the portfolios formed according the proposed model are robust (in terms of liquidity) and competitive (in comparison with the equally weighted and the minimum variance portfolios).

All the suggested methodologies present limitations and should be improved in future works. In addition to those previously discussed, one of the main limitations that is transversal to all of the suggested methodologies is the presence of estimation error. To deal with this problem, there are several research lines that can be followed. The use of Bayesian estimation, shrinkage estimators, robust optimization techniques or factor models are just some of the ways to improve the models proposed in this thesis. Related with the latter, there are few portfolio optimization models based on firms' characteristics. Based on Brandt et al. (2009) we are starting to think in a new way of explore the information about the cross-sectional characteristics of securities.

The various methodologies proposed in this thesis seek to better fit reality. None should therefore be regarded as the "best" investment strategy. Rather, they should be used holistically in the design of the "best" investment strategy. This thesis leaves many open questions but, as the great Richard Feynman said: "We absolutely must leave room for doubt or there is no progress and no learning. There is no learning without having to pose a question. And a question requires doubt. People search for certainty. But there is no certainty".

Appendices

A The Financial Modelers' Manifesto (by Emanuel Derman and Paul Willmott)

The Financial Modelers' Manifesto

by Emanuel Derman and Paul Wilmott

Preface

A spectre is haunting Markets – the spectre of illiquidity, frozen credit, and the failure of financial models.

Beginning with the 2007 collapse in subprime mortgages, financial markets have shifted to new regimes characterized by violent movements, epidemics of contagion from market to market, and almost unimaginable anomalies (who would have ever thought that swap spreads to Treasuries could go negative?). Familiar valuation models have become increasingly unreliable. Where is the risk manager that has not ascribed his losses to a once-in-a-century tsunami?

To this end, we have assembled in New York City and written the following manifesto.

Manifesto

In finance we study how to manage funds – from simple securities like dollars and yen, stocks and bonds to complex ones like futures and options, subprime CDOs and credit default swaps. We build financial models to estimate the fair value of securities, to estimate their risks and to show how those risks can be controlled. How can a model tell you the value of a security? And how did these models fail so badly in the case of the subprime CDO market?

Physics, because of its astonishing success at predicting the future behavior of material objects from their present state, has inspired most financial modeling. Physicists study the world by repeating the same experiments over and over again to discover forces and their almost magical mathematical laws. Galileo dropped balls off the leaning tower, giant teams in Geneva collide protons on protons, over and over again. If a law is proposed and its predictions contradict experiments, it's back to the drawing board. The method works. The laws of atomic physics are accurate to more than ten decimal places.

It's a different story with finance and economics, which are concerned with the mental world of monetary value. Financial theory has tried hard to emulate the style and elegance of physics in order to discover its own laws. But markets are made of people, who are influenced by events, by their ephemeral feelings about events and by their expectations of other people's feelings. The truth is that there are no fundamental laws in finance. And even if there were, there is no way to run repeatable experiments to verify them.

You can hardly find a better example of confusedly elegant modeling than models of CDOs. The CDO research papers apply abstract probability theory to the price co-movements of thousands of mortgages. The relationships between so many mortgages can be vastly complex. The modelers, having built up their fantastical theory, need to make it useable; they resort to sweeping under the model's rug all unknown dynamics; with the dirt ignored, all that's left is a single number, called the default correlation. From the sublime to the elegantly ridiculous: all uncertainty is reduced to a single parameter that, when entered into the model by a trader, produces a CDO value. This over-reliance on probability and statistics is a severe limitation. Statistics is shallow description, quite unlike the deeper cause and effect of physics, and can't easily capture the complex dynamics of default.

Models are at bottom tools for approximate thinking; they serve to transform your intuition about the future into a price for a security today. It's easier to think intuitively about future housing prices, default rates and default correlations than it is about CDO prices. CDO models turn your guess about future housing prices, mortgage default rates and a simplistic default correlation into the model's output: a current CDO price.

Our experience in the financial arena has taught us to be very humble in applying mathematics to markets, and to be extremely wary of ambitious theories, which are in the end trying to model human behavior. We like simplicity, but we like to remember that it is our models that are simple, not the world.

Unfortunately, the teachers of finance haven’t learned these lessons. You have only to glance at business school textbooks on finance to discover stilts of mathematical axioms supporting a house of numbered theorems, lemmas and results. Who would think that the textbook is at bottom dealing with people and money? It should be obvious to anyone with common sense that every financial axiom is wrong, and that finance can never in its wildest dreams be Euclid. Different endeavors, as Aristotle wrote, require different degrees of precision. Finance is not one of the natural sciences, and its invisible worm is its dark secret love of mathematical elegance and too much exactitude.

We do need models and mathematics – you cannot think about finance and economics without them – but one must never forget that models are not the world. Whenever we make a model of something involving human beings, we are trying to force the ugly stepsister’s foot into Cinderella’s pretty glass slipper. It doesn't fit without cutting off some essential parts. And in cutting off parts for the sake of beauty and precision, models inevitably mask the true risk rather than exposing it. The most important question about any financial model is how wrong it is likely to be, and how useful it is despite its assumptions. You must start with models and then overlay them with common sense and experience.

Many academics imagine that one beautiful day we will find the ‘right’ model. But there is no right model, because the world changes in response to the ones we use. Progress in financial modeling is fleeting and temporary. Markets change and newer models become necessary. Simple clear models with explicit assumptions about small numbers of variables are therefore the best way to leverage your intuition without deluding yourself.

All models sweep dirt under the rug. A good model makes the absence of the dirt visible. In this regard, we believe that the Black-Scholes model of options valuation, now often unjustly maligned, is a model for models; it is clear and robust. Clear, because it is based on true engineering; it tells you how to manufacture an option out of stocks and bonds and what that will cost you, under ideal dirt-free circumstances that it defines. Its method of valuation is analogous to figuring out the price of a can of fruit salad from the cost of fruit, sugar, labor and transportation. The world of markets doesn’t exactly match the ideal circumstances Black-Scholes requires, but the model is robust because it allows an intelligent trader to qualitatively adjust for those mismatches. You know what you are assuming when you use the model, and you know exactly what has been swept out of view.

Building financial models is challenging and worthwhile: you need to combine the qualitative and the quantitative, imagination and observation, art and science, all in the service of finding approximate patterns in the behavior of markets and securities. The greatest danger is the age-old sin of idolatry. Financial markets are alive but a model, however beautiful, is an artifice. No matter how hard you try, you will not be able to breathe life into it. To confuse the model with the world is to embrace a future disaster driven by the belief that humans obey mathematical rules.

MODELERS OF ALL MARKETS, UNITE! You have nothing to lose but your illusions.

The Modelers' Hippocratic Oath

~ I will remember that I didn't make the world, and it doesn't satisfy my equations.

~ Though I will use models boldly to estimate value, I will not be overly impressed by mathematics.

~ I will never sacrifice reality for elegance without explaining why I have done so.

~ Nor will I give the people who use my model false comfort about its accuracy.
Instead, I will make explicit its assumptions and oversights.

~ I understand that my work may have enormous effects on society and the economy,
many of them beyond my comprehension

.
Emanuel Derman
January 7 2009

Paul Wilmott
January 7 2009

B MOO and DMS

MOO

We often face optimization problems where two or more objective functions have to be optimized, and many times there are tradeoffs between these objectives. A constrained nonlinear MOO problem can be formulated as

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & F(x) \; = \; [\; f_1(x) \; ... \; f_m(x) \;]^\top \\ \text{subject to} \quad & x \in \Omega, \end{aligned} \tag{B.1}$$

where:

- $f_i : \mathbb{R}^n \hookrightarrow \mathbb{R}, \, \forall_i \in \{1, \ldots, m\}$,
- $m \geq 2$ is the number of objective functions, and
- Ω is the decision feasible set.

Therefore, $\mathbb{R}^n$ corresponds to the space of the variables and $\mathbb{R}^m$ is the decision space. In order to find an optimal solution for Problem (B.1), we need to use a criterion for systematically comparing different points. This can be achieved by using the concept of Pareto dominance.

Definition B.1 (Pareto minimizer) *$x \in \Omega$ is a Pareto minimizer of $F(\cdot)$ if*

$$\nexists_{y \in \Omega} : \; f_i(y) \; < \; f_i(x), \; \forall_{i \in \{1,\ldots,m\}}.$$

A Pareto minimizer, $x \in \Omega$, is also known as a nondominated point.

Definition B.2 (The set of Pareto minimizers, X) *The set of Pareto minimizers, S, of $F(\cdot)$, is defined as*

$$S \; = \; \left\{ x \in \Omega : \; \nexists_{y \in \Omega} : \; f_i(y) \; < \; f_i(x), \; \forall_{i \in \{1,\ldots,m\}} \right\}.$$

The set of Pareto minimizers, S, defines the set of nondominated points of Problem (B.1). The Pareto or efficient frontier corresponds to the mapping of $F(\cdot)$ on the set of nondominated points, S.

For example, let us recall the MV optimization model in Chapter 2 (see Problem (2.10) and Problem (2.12)). The solution of Problem (2.10) is in fact a Pareto minimizer (or a nondominated solution) for Problem (2.12).

Let $w_t^* \in P_t = \{w_t \in \mathbb{R}^N : 1_N^\top w_t = 1 \ \wedge \ w_t \geq 0_N\}$ be a solution of Problem (2.10). Now, suppose that w_t^* is not a Pareto minimizer of Problem (2.12), then, by definition (see Definition B.1),

$$\exists_{z_t \in P_t} : \ z_t^\top \Sigma_t z_t \ < \ w_t^* \Sigma_t w_t^* \ \wedge \ -\mu_t^\top z_t \ < \ -\mu_t^\top w_t^*. \tag{B.2}$$

From Equation (B.2) follows that

$$z_t \in P_t : \ z_t^\top \Sigma_t z_t \ < \ w_t^* \Sigma_t w_t^* \ \wedge \ \mu_t^\top z_t \ > \ \mu_t^\top w_t^* > r. \tag{B.3}$$

But, from Equation (B.3), it follows that $w_t^* \in P_t$ is not an optimal solution of Problem (2.10). This way, we have shown by *reductio ad absurdum* that $w_t^* \in P_t$ is in fact a Pareto minimizer or a nondominated solution of Problem (2.12). Denoting Y as the set of the nondominated points (see Definition B.2) of Problem (2.12), the mapping of $v(r_{p,t+1}) = w_t^\top \Sigma_t w_t$ and $m(r_{p,t+1}) = \mu_t^\top w_t$ onto Y corresponds to the efficient frontier plotted in Figure 2.2.

DMS

Derivative-free optimization (see Conn et al., 2009, for a survey) is a class of nonlinear optimization techniques used when we face problems where all, or at least some, of the derivatives are unavailable, unreliable or impractical to obtain. For example, when the function evaluations are expensive or when they are noisy, one cannot approximate the derivatives using methods based on finite differences. Given a feasible point, a derivative-free algorithm only uses the function value as information, which is why many times these algorithms are referred as of black-box type.

DMS (Custódio et al., 2011) is a derivative-free algorithm to solve problems given on the form of Problem (B.1). DMS does not aggregate or scalarize any of the objective functions components. Each iteration of DMS comprises a search step (optional) and a pool step. DMS is thus a generalization of direct search methods of directional type (see, e.g., Kolda et al., 2003; Conn et al., 2009) from single to MOO. The main objective of DMS is to compute an approximate solution for Problem (B.1), i.e., an approximation of the real Pareto frontier. From the pool step, theoretically, it is only possible to prove that there is a limit point in a stationary form of the Pareto frontier, as no aggregation or scalarization method is incorporated (see Custódio et al., 2011, for further details). Nevertheless, as documented in Custódio et al. (2011), in practice

DMS exhibits a very good performance, being able to compute the entire Pareto frontier for a wide variety of problems. In a test set of more than one hundred problems (some with discontinuous and nonconvex Pareto frontiers), DMS performed in a highly competitive way in comparison with other well-known derivative-free solvers, namely AMOSA (Bandyopadhyay et al., 2008), BIMADS (Audet et al., 2008) and NSGA-II (Deb et al., 2002). In direct comparison with the latter, DMS performed the best.

As common in derivative-free optimization, DMS deals with the constraints through an extreme barrier function

$$F_\Omega(x) = \begin{cases} F(x) & \text{if } x \in \Omega, \\ (+\infty, \ldots, +\infty)^\top & \text{otherwise.} \end{cases} \tag{B.4}$$

As explained in Custódio et al. (2011), Equation (B.4) states that when a point, x, is not feasible then each function component assumes the value $+\infty$. This procedure allows to deal with black-box type constraints. Following Custódio et al. (2011) we present the DMS algorithm below.

Algorithm B.1 (DMS for MOO)

Initialization

Choose $x_0 \in \Omega$ with $f_i(x_0) < +\infty, \forall i \in \{1, \ldots, m\}$, $\alpha_0 > 0$, $0 < \beta_1 \leq \beta_2 < 1$, and $\gamma \geq 1$. Let D be a set of positive spanning sets. Initialize the list of nondominated points and corresponding step size parameters ($L_0 = \{(x_0; \alpha_0)\}$ in case of a singleton).

For $k = 0, 1, 2, \ldots$

1. **Selection of an iterate point:** Order the list L_k in some way and select the first item $(x; \alpha) \in L_k$ as the current iterate and step size parameter (thus setting $(x_k; \alpha_k) = (x; \alpha)$).
2. **Search step:** Compute a finite set of points $\{z_s\}_{s \in S}$ and evaluate F_Ω at each element. Set $L_{add} = \{(z_s; \alpha_k), s \in S\}$.
   Form L_{trial} by eliminating dominated points from $L_k \cup L_{add}$. If $L_{trial} \neq L_k$ declare the iteration (and the search step) successful, set $L_{k+1} = L_{trial}$, and skip the poll step.
3. **Poll step:** Choose a positive spanning set D_k. Evaluate F_Ω at the set of poll points $P_k = \{x_k + \alpha_k d : \, d \in D_k\}$. Set $L_{add} = \{(x_k + \alpha_k d; \alpha_k), d \in D_k\}$.
   Form L_{trial} by eliminating dominated points from $L_k \cup L_{add}$. If $L_{trial} \neq L_k$

declare the iteration (and the poll step) as successful and set $L_{k+1} = L_{trial}$. Otherwise, declare the iteration (and the poll step) unsuccessful and set $L_{k+1} = L_k$.

4. **Step size parameter update:** If the iteration was successful then maintain or increase the corresponding step size parameters: $\alpha_{k,new} \in [\alpha_k, \gamma\alpha_k]$ and replace all the new points $(x_k + \alpha_k d; \alpha_k)$ in L_{k+1} by $(x_k + \alpha_k d; \alpha_{k,new})$, when success is coming from the poll step, or $(z_s; \alpha_k)$ in L_{k+1} by $(z_s; \alpha_{k,new})$, when success is coming from the search; replace also $(x_k; \alpha_k)$, if in L_{k+1}, by $(x_k; \alpha_{k,new})$.
   Otherwise decrease the step size parameter: $\alpha_{k,new} \in [\beta_1\alpha_k, \beta_2\alpha_k]$ and replace the poll pair $(x_k; \alpha_k)$ in L_{k+1} by $(x_k; \alpha_{k,new})$.

We will now report all the changes made to the default parameters of the `dms` solver[18], in order to properly solve every problem of this thesis. Along this thesis we have used the `dms` solver to compute the solutions of the following problems: Problem (3.8) (see Chapter 3); Problem (4.4), Problem (4.5) and Problem (4.6) (see Chapter 4); and the solution of Problem (5.22) (see Chapter 5). To solve every one of these problems, we made the following changes to the `dms` solver:

- we needed to increase the maximum number of function evaluations (option `max_fevals`) from 20000 to 350000;
- we needed to require more accuracy by reducing the step size tolerance (option `tol_stop`) from 10^{-3} to 10^{-5};
- instead of using a random sample for the initial list of feasible nondominated points, as it is assumed by default, for Problem (3.8) and Problem (5.22) we set the option `list` equal to 0 (initializing the list with a single point, which we defined as the equally weighted point) and for the problems in Chapter 4 (Problem (4.4), Problem (4.5) and Problem (4.6)) we set the option `list` equal to 3 (initializing the list with points that are considered equally spaced in the line segment, joining the variable upper and lower bounds).

[18]This solver is public and available (at the current date, July 2017) by request at `http://www.mat.uc.pt/dms/`.

References

Aït-Sahalia, Y. and Brandt, M. W. (2001). Variable selection for portfolio choice. *The Journal of Finance*, 56(4):1297–1351.

Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2005a). How often to sample a continuous-time process in the presence of market microstructure noise. *The Review of Financial Studies*, 18(2):351–416.

Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2005b). A tale of two time scales: determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, 100(472):1394–1411.

Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2011). Ultra high frequency volatility estimation with dependent microstructure noise. *Journal of Econometrics*, 160(1):160–175.

Alderfer, C. P. and Bierman, H. (1970). Choices with risk: Beyond the mean and variance. *The Journal of Business*, 43(3):341–53.

Amaya, D., Christoffersen, P., Jacobs, K., and Vasquez, A. (2015). Does realized skewness predict the cross-section of equity returns? *Journal of Financial Economics*, 118(1):135–167.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, 5(1):31–56.

Amihud, Y. and Mendelsen, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics*, 17(2):223–249.

Anagnostopoulos, K. P. and Mamanis, G. (2011). The mean-variance cardinality constrained portfolio optimization problem: An experimental evaluation of five multiobjective evolutionary algorithms. *Expert Systems with Applications*, 38:14208–14217.

Andersen, T. G., Bollerslev, T., Christoffersen, P. F., and Diebold, F. X. (2006). *Volatility and Correlation Forecasting*, volume 1 of *Handbook of Economic Forecasting*, chapter 15, pages 777–878. Elsevier, Amsterdam: North-Holland.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Ebens, H. (2001a). The distribution of realized stock return volatility. *Journal of Financial Economics*, 61(1):43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2001b). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, 96(453):42–55.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71(2):579–625.

Ang, J. S. and Chua, J. H. (1979). Composite measures for the evaluation of investment performance. *Journal of Financial and Quantitative Analysis*, 14(2):361–384.

Arditti, F. D. (1967). Risk and the required return on equity. *The Journal of Finance*, 22(1):19–36.

Arditti, F. D. (1975). Skewness and investor's decisions: A reply. *Journal of Financial and Quantitative Analysis*, 10(1):173–176.

Areal, N. M. P. C. and Taylor, S. J. (2002). The realized volatility of FTSE-100 futures prices. *Journal of Futures Markets*, 22(7):627–648.

Arrow, K. J. (1971). *Essays in the Theory of Risk-Bearing.* Chicago: Markham, Amsterdam and London: North-Holland.

Athayde, G. and Flôres, R. (2004). Finding a maximum skewness portfolio: A general solution to three-moments portfolio choice. *Journal of Economic Dynamics and Control*, 28(7):1335–1352.

Audet, C., Savard, G., and Zghal, W. (2008). Multiobjective optimization through a series of single-objective formulations. *SIAM Journal on Optimization*, 19(1):188–210.

Bachelier, L. (1900). *Théorie de la Spéculation.* PhD thesis, Annales scientifiques de l'École Normale Supérieure, Série 3, 17:21-86.

Bai, X., Russell, J. R., and Tiao, G. C. (2003). Kurtosis of GARCH and stochastic volatility models with non-normal innovations. *Journal of Econometrics*, 114(2):349–360.

Baker, M., Bradley, B., and Wurgler, J. (2011). Benchmarks as limits to arbitrage: Understanding the low-volatility anomaly. *Financial Analysts Journal*, 67(1):40–54.

Balduzzi, P. and Lynch, A. W. (1999). Transaction costs and predictability: Some utility cost calculations. *Journal of Financial Economics*, 52(1):47–78.

Bamberg, G. and Dorfleitner, G. (2013). On a neglected aspect of portfolio choice: The role of the invested capital. *Review of Managerial Science*, 7:85–98.

Ban, G.-Y., Karoui, N. E., and Lim, A. E. B. (2016). Machine learning and portfolio optimization. *Management Science*.

Bandyopadhyay, S., Saha, S., Maulik, U., and Deb, K. (2008). A simulated anneling-based multiobjective optimization algorithm: Amosa. *IEEE Transactions on Evolutionary Computation*, 12(3):269–283.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2011). Multivariate realized kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics*, 162(2):149–169.

Barndorff-Nielsen, O. E. and Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of The Royal Statistical Society*, Series B 64(2):253–280.

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Econometric analysis of realized covariation: high frequency based covariance, regression, and correlation in financial economics. *Econometrica*, 72(3):885–925.

Barndorff-Nielsen, O. E. and Shephard, N. (2005). Variation, jumps, market frictions and high frequency data in financial econometrics. Technical report, Oxford Financial Research Centre, United Kingdom.

Bawa, V. S. (1975). Optimal rules for ordering uncertain prospects. *Journal of Financial Economics*, 2(1):95–121.

Beedles, W. L. (1979). On the asymmetry of market returns. *Journal of Financial and Quantitative Analysis*, 14(3):653–660.

Benartzi, S. and Thaler, R. H. (2001). Naive diversification strategies in defined contribution saving plans. *The American Economic Review*, 91(1):79–98.

Bertsimas, D. and Shioda, R. (2009). Algorithm for cardinality-constrained quadratic optimization. *Computational Optimization and Applications*, 43(1):1–22.

Bienstock, D. (1996). Computational study of a family of mixed-integer quadratic programming problems. *Mathematical Programming*, 74(2):121–140.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81(3):637–654.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327.

Brandt, M. W., Santa-Clara, P., and Valkanov, R. (2009). Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *The Review of Financial Studies*, 22(9):3411–3447.

Brennan, M. J. and Subrahmanyam, A. (1996). Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. *Journal of Financial Economics*, 41(3):441–464.

Brito, R. P., Sebastião, H., and Godinho, P. (2016). Efficient skewness/semivariance portfolios. *Journal of Asset Management*, 17(5):331–346.

Brito, R. P., Sebastião, H., and Godinho, P. (2017a). On the gains of using high frequency data and higher moments in portfolio choice. Technical report, CeBER WP No. 2017-02, Faculty of Economics, University of Coimbra, Portugal.

Brito, R. P., Sebastião, H., and Godinho, P. (2017b). Portfolio choice with high frequency data: CRRA preferences and the liquidity effect. *Portuguese Economic Journal, DOI: 10.1007/s10258-017-0131-3 (Forthcoming article. At the current date, July 2017, an early version is available at* `https://link.springer.com/article/10.1007/s10258-017-0131-3`*).*

Brito, R. P., Sebastião, H., and Godinho, P. (2017c). Portfolio management with higher moments: The cardinality impact. *International Transactions in Operational Research, DOI: 10.1111/itor.12404 (Forthcoming article. At the current date, July 2017, an early version is available at* `http://onlinelibrary.wiley.com/doi/10.1111/itor.12404/full`*).*

Brito, R. P. and Vicente, L. N. (2014). *Efficient cardinality/mean-variance portfolios.* In: C. Pötzsche, C. Heuberger, B. Kaltenbacher, and F. Rendl (eds.) System Modeling and Optimization, Springer series IFIP Advances in Information and Communication Technology, Berlin: Springer Berlin Heidelberg.

Broadie, M. (1993). Computing efficient frontiers using estimated parameters. *Annals of Operations Research*, 45(1):21–58.

Brodie, J., Daubechies, I., Mol, C. D., Giannone, D., and Loris, I. (2009). Sparse and stable markowitz portfolios. *Proceedings of the National Academy of Sciences of the United States of America*, 106(30):12267–12272.

Campbell, J. Y., Lo, A. W., and MacKinlay, A. G. (1997). *The Econometrics of Financial Markets*. Princeton University Press, Chichester.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1):57–82.

Cesarone, F., Scozzari, A., and Tardella, F. (2009). Efficient algorithms for mean-variance portfolio optimization with hard real-world constraints. *Giornale dell'Istituto Italiano degli Attuari*, 72:37–56.

Chang, K., Hameed, A., and Tong, W. (2000a). Profitability of momentum strategies in the international equity markets. *The Journal of Financial and Quantitative Analysis*, 35(2):153–172.

Chang, L. K. C., Jegadeesh, N., and Lakonishok, J. (1996). Momentum strategies. *The Journal of Finance*, 51(5):1681–1713.

Chang, L. K. C., Karceski, J., and Lakonishok, J. (1999). O portfolio optimization: Forecasting covariances and choosing the risk model. *The Review of Financial Studies*, 12:937–974.

Chang, T. J., Meade, N., Beasley, J. E., and Sharaiha, Y. M. (2000b). Heuristics for cardinality constrained portfolio optimisation. *Computers & Operations Research*, 27(13):1271–1302.

Chiang, T. C. and Zheng, D. (2015). Liquidity and stock returns: Evidence from international markets. *Global Finance Journal*, 27(C):73–97.

Chopra, V. K. and Ziemba, W. T. (1993). The effect of errors in means, variances, and covariances on optimal portfolio choice. *The Journal of Portfolio Management*, 19(2):6–11.

Chordia, T., Roll, R., and Subrahmanyam, A. (2001). Market liquidity and trading activity. *The Journal of Finance*, 56(2):501–530.

Chunhachinda, P., Dandapani, K., Hamid, S., and Prakash, A. J. (1997). Portfolio selection and skewness: Evidence from international stock markets. *Journal of Banking & Finance*, 21(2):143–167.

Conn, A. R., Scheinberg, K., and Vicente, L. N. (2009). *Introduction to Derivative-Free Optimization*. MPS-SIAM Series on Optimization. SIAM, Philadelphia, PA.

Cornnuejols, G. and Tütüncü, R. H. (2007). *Optimizations Methods in Finances*. Cambridge University Press, Cambridge.

Curto, J., Reis, E., and Esperança, J. (2003). Stable paretian distributions: an unconditional model for PSI20, DAX and DJIA indexes. *Review of Financial Markets*, 5(1):5–18.

Custódio, A. L., Madeira, J. F. A., Vaz, A. I. F., and Vicente, L. N. (2011). Direct multisearch for multiobjective optimization. *SIAM Journal on Optimization*, 21(3):1109–1140.

Datar, V. T., Naik, N. Y., and Radcliffe, R. (1998). Liquidity and stock returns: An alternative test. *Journal of Financial Markets*, 1(2):203–219.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197.

DeMiguel, V., Garlappi, L., Nogales, F. J., and Uppal, R. (2009a). A generalized approach to portfolio optimization: Improving performance by constrained portfolio norms. *Management Science*, 55:798–812.

DeMiguel, V., Garlappi, L., and Uppal, R. (2009b). Optimal versus naive diversification: How inefficient is the $1/N$ portfolio strategy? *The Review of Financial Studies*, 22(5):1915–1953.

DeMiguel, V., Martin-Utrera, A., Nogales, F. J., and Uppal, R. (2017). A portfolio perspective on the multitude of firm characteristics. Technical report, London Business School, United Kingdom.

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

DeMiguel, V., Nogales, F. J., and Uppal, R. (2014). Stock return serial dependence and out-of-sample portfolio performance. *The Review of Financial Studies*, 27(4):1031–1073.

DeMiguel, V. and Olivares-Nadal, A. V. (2016). A robust perspective on transaction costs in portfolio optimization. Technical report, London Business School, United Kingdom.

DeMiguel, V., Plyakha, Y., Uppal, R., and Vilkov, G. (2013). Improving portfolio selection using option-implied volatility and skewness. *Journal of Financial and Quantitative Analysis*, 48(6):1813–1845.

Derman, E. (2011). Metaphors, models & theories. *Quarterly Journal of Finance*, 1(1):109–126.

Ding, Z., Granger, C. W. J., and Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1(1):83–106.

Dittmar, R. (2002). Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. *The Journal of Finance*, 57(1):369–403.

Easley, D., Hvidkjaer, S., and O'Hara, M. (2002). Is information risk a determinant of asset returns? *The Journal of Finance*, 57(5):2185–2221.

Efron, B. and Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, New York.

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica*, 50(4):987–1008.

Estrada, J. (2008). Mean-semivariance optimization: A heuristic approach. *Journal of Applied Finance*, 18(1):57–72.

Evans, J. L. and Archer, S. H. (1968). Diversification and the reduction of dispersion: An empirical analysis. *The Journal of Finance*, 23(5):761–767.

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

Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, 38(1):34–105.

Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56.

Fama, E. F. and French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *The Journal of Finance*, 51(1):55–84.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22.

Fan, J., Li, Y., and Yu, K. (2012). Vast volatility matrix estimation using high-frequency data for portfolio selection. *Journal of The American Statistical Association*, 107(497):412–428.

Fernandes, B., Street, A., ao, D. V., and Fernandes, C. (2016). An adaptive robust portfolio optimization model with loss constraints based on data-driven polyhedral uncertainty sets. *European Journal of Operational Research*, 256(3):961–970.

Fieldsend, J. E., Matatko, J., and Peng, M. (2004). Cardinality constrained portfolio optimisation. In *Intelligent Data Engineering and Automated Learning - IDEAL 2004*, volume 3177 of *Lecture Notes in Computer Science*, pages 788–793. Springer, Berlin.

Fishburn, P. C. (1977). Mean-risk analysis with risk associated with below-target returns. *American Economic Review*, 67(2):116–126.

Flemming, J., Kirby, C., and Ostdiek, B. (2003). The economic value of volatility timing using "realized" volatility. *Journal of Financial Economics*, 67(3):473–509.

Goldfarb, D. and Iyengar, G. (2003). Robust portfolio selection problems. *Mathematics of Operations Research*, 28(1):1–38.

Gotoh, J.-Y., Shinozaki, K., and Takeda, A. (2013). Robust portfolio techniques for mitigating the fragility of cvar minimization and generalization to coherent risk measures. *Quantitative Finance*, 13(10):1621–1635.

Goyenko, R. Y., Holden, C. W., and Trzcinka, C. A. (2009). Do liquidity measures measure liquidity? *Journal of Financial Economics*, 92(1):153–181.

Hansen, P. R. and Lunde, A. (2006). Realized variance and market microstructure noise. *Journal of Business & Economics Statistics*, 24(2):127–161.

Harlow, W. V. and Rao, R. K. S. (1989). Asset pricing in a generalised mean-lower partial moment framwork: theory and evidence. *Journal of Financial and Quantitative Analysis*, 24(3):285–311.

Harvey, C. R., Liechty, J., Liechty, M. W., and Mueller, P. (2010). Portfolio selection with higher moments. *Quantitative Finance*, 10(5):469–485.

Harvey, C. R. and Siddique, A. (2000). Conditional skewness in asset pricing tests. *The Journal of Finance*, 55(3):1263–1295.

Hasbrouck, J. and Seppi, D. J. (2001). Common factors in prices, order flows, and liquidity. *Journal of Financial Economics*, 59(3):383–411.

Hautsch, N., Kyj, L. M., and Oomen, R. C. A. (2012). A blocking and regularization approach to high-dimensional realized covariance estimation. *Journal of Applied Econometrics*, 27(4):625–645.

Hogan, W. W. and Warren, J. M. (1974). Toward the development of an equilibrium capital-market model based on semivariance. *Journal of Financial and Quantitative Analysis*, 9:1–11.

Hsieh, D. A. (1991). Chaos and nonlinear dynamics: Application to financial markets. *The Journal of Finance*, 46(5):1839–1877.

Hsu, D. A. (1984). The behavior of stock returns: Is it stationary or evolutionary? *The Journal of Financial and Quantitative Analysis*, 19(1):11–28.

Huang, T.-C., Tu, Y.-C., and Chou, H.-C. (2015). Long memory and the relation between options and stock prices. *Finance Research Letters*, 12:77–91.

Huang, X. (2008). Mean-semivariance models for fuzzy portfolio selection. *Journal of Computational and Applied Mathematics*, 217(1):1–8.

Huberman, G. and Halka, D. (2001). Systematic liquidity. *The Journal of Financial Research*, 24(2):161–178.

Huo, L., Kim, T.-H., and Kim, Y. (2012). Robust estimation of covariance and its application to portfolio optimization. *Finance Research Letters*, 9(3):121–134.

Ikeda, Y. and Kubokawa, T. (2016). Linear shrinkage estimation of large covariance matrices using factor models. *Journal of Multivariate Analysis*, 152:61–81.

Israelsen, C. L. (2005). A refinement to the sharpe ratio and information ratio. *Journal of Asset Management*, 5(6):423–427.

Jagannathan, R. and Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance*, 58(4):1651–1684.

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

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

Jobst, N. J., Hornimam, M. D., Lucas, C. A., and Mitra, G. (2010). Computational aspects of alternative portfolio selection models in the presence of discrete asset choice constraints. *Quantitative Finance*, 1:489–501.

Jorion, P. (1986). Bayes-stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis*, 21(3):279–292.

Jorion, P. (2000). Risk management lessons from long-term capital management. *European Fianancial Management*, 6(3):277–300.

Joro, T. and Na, P. (2006). Portfolio performance evaluation in a mean-variance-skewness framework. *European Journal of Operational Research*, 175(1):446–461.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292.

Kan, R. and Zhou, G. (2007). Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis*, 42(3):621–656.

Kimball, M. (1993). Standard risk aversion. *Econometrica*, 61(3):589–611.

Klein, R. W. and Bawa, V. S. (1976). The effect of estimation risk on optimal portfolio choice. *Journal of Financial Economics*, 3(3):215–231.

Koekebakker, S. and Zakamouline, V. (2009). A generalization of the mean-variance analysis. *European Fianancial Management*, 15(5):934–970.

Kolda, T. G., Lewis, R. M., and Torczon, V. (2003). Optimization by direct search: New perspectives on some classical and modern methods. *SIAM Review*, 45(3):385–482.

Koopman, S. J., Jungbacker, B., and Hol, E. (2005). Forecasting daily variability of the S&P 100 stock index using historical, realized and implied volatility measurements. *Journal of Empirical Finance*, 12(3):445–475.

Kraus, A. and Litzenberger, R. H. (1976). Skewness preference and the valuation of risk assets. *The Journal of Finance*, 31(4):1085–1100.

Ledoit, O. and Wolf, M. (2004a). Honey, I shrunk the sample covariance matrix. *The Journal of Portfolio Management*, 30(4):110–119.

Ledoit, O. and Wolf, M. (2004b). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2):365–411.

Ledoit, O. and Wolf, M. (2008). Robust performance hypothesis testing with the sharpe ratio. *Journal of Empirical Finance*, 15(5):850–859.

Lewellen, J. (2002). Momentum and autocorrelation in stock returns. *The Review of Financial Studies*, 15(2):533–563.

Liu, Q. (2009). On portfolio optimization: How and when do we benefit from high-frequency data? *Journal of Applied Econometrics*, 24(4):560–582.

Mandelbrot, B. (1963). The variation of certain speculative prices. *The Journal of Business*, 36(4):394–419.

Mandelbrot, B. (1967). The variation of some other speculative prices. *The Journal of Business*, 40(4):393–413.

Mandelbrot, B. and Hudson, R. L. (2004). *The (Mis)behavior of Markets*. Basic Books, New York.

Maringer, D. and Parpas, P. (2009). Global optimization of higher order moments in portfolio selection. *Journal of Global Optimization*, 43(2):219–230.

Markowitz, H. M. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.

Markowitz, H. M. (1959). *Portfolio Selection: Efficient Diversification of Investments.* John Wiley & Sons, New York.

Markowitz, H. M. (2014). Mean-variance approximations to expected utility. *European Journal of Operational Research*, 234(2):346–355.

Martellini, L. and Ziemann, V. (2010). Improved estimates of higher-order comoments and implications for portfolio selection. *The Review of Financial Studies*, 23(4):1467–1502.

Mas-Colell, A., Whinston, M. D., and Green, J. R. (1995). *Microeconomic Theory.* Oxford University Press, Oxford.

McAleer, M. and Medeiros, M. (2008). Realized volatility: A review. *Econometric Reviews*, 27(1-3):10–45.

Meddahi, N., Mykland, P., and Shephard, N. (2011). Special issue on realized volatility. *Journal of Econometrics*, 160(1):1–288.

Mencia, J. and Sentana, E. (2009). Multivariate location-scale mixtures of normals and mean-variance-skewness portfolio allocation. *Journal of Econometrics*, 153(2):105–121.

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

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

Moreno, D. and Rodríguez, R. (2009). The value of coskewness in mutual fund performance evaluation. *Journal of Banking & Finance*, 33(9):1664–1676.

Nawrocki, D. N. (1999). A brief history of downside risk measures. *The Journal of Investing*, 8(3):9–25.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, 59(2):347–370.

Neuberger, A. (2012). Realized skewness. *The Review of Financial Studies*, 25(11):3423–3455.

Osborne, M. F. M. (1959). Brownian motion in the stock market. *Operations Research*, 7:145–173.

Pastor, L. (2000). Portfolio selection and asset pricing models. *The Journal of Finance*, 55(1):179–223.

Pastor, L. and Stambaugh, R. F. (2000). Comparing asset princing models: An investment perspective. *Journal of Financial Economics*, 56(3):335–81.

Pastor, L. and Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*, 111(3):642–685.

Quirk, J. P. and Saposnik, R. (1962). Admissability and measurable utility functions. *The Review of Economic Studies*, 29(2):140–146.

Rachev, S. T. and Mittnik, S. (2000). *Stable Paretian Models in Finance*. Wiley, Chichester.

Rege, S., Teixeira, J. C. A., and Menezes, A. G. (2013). The daily returns of the portuguese stock index: a distributional characterization. *Journal of Risk Model Validation*, 7(4):53–70.

Roy, A. D. (1952). Safety first and the holding of assets. *Econometrica*, 20(3):431–449.

Sadique, S. and Silvapulle, P. (2001). Long-term memory in stock market returns: International evidence. *International Journal of Finance and Economics*, 6(1):59–67.

Schwert, G. W. (1989). Why does stock market volatility change over time? *The Journal of Finance*, 44(5):1115–1153.

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

Sing, T. F. and Ong, S. E. (2000). Asset allocation in a downside risk framework. *Journal of Real Estate Portfolio Management*, 6(3):213–224.

Statman, M. (1987). How many stocks make a diversified portfolio? *Journal of Financial and Quantitative Analysis*, 22(3):353–363.

Stărică, C. and Granger, C. (2005). Nonstationarities in stock returns. *The Review of Economics and Statistics*, 87(3):503–522.

Tavares, A. B., Curto, J. D., and Tavares, G. N. (2008). Modelling heavy tails and asymmetry using ARCH-type models with stable paretian distributions. *Nonlinear Dynamics*, 51(1):231–243.

Taylor, S. J. (1986). *Modelling Financial Time Series*. Wiley, Chichester.

Tütüncü, R. H. and Koenig, M. (2004). Robust asset allocation. *Annals of Operations Research*, 132(1):157–187.

Vielma, J. P., Ahmed, S., and Nemhauser, G. L. (2008). A lifted linear programming branch-and-bound algorithm for mixed-integer conic quadratic programs. *INFORMS Journal on Computing*, 20(3):438–450.

Von Neumann, J. and Morgenstern, O. (1953). *Theory of games and economic behavior*. Princeton University Press, Princeton.

Williams, J. B. (1938). *The theory of investment value*. Harvard University Press, Cambridge.

Woodside-Oriakhi, M., Lucas, C., and Beasley, J. E. (2011). Heuristic algorithms for the cardinality constrained efficient frontier. *European Journal of Operational Research*, 213(3):538–550.

Zakoian, J. M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, 18(5):931–955.

Zhu, S. and Fukushima, M. (2009). Worst-case conditional value-at-risk with application to robust portfolio management. *Operations Research*, 57(5):1155–1168.