Robust Asset Allocation Under Model Ambiguity

Sandrine Tobelem - Foldvari

Department of Statistics

The London School of Economics and Political Sciences

A thesis submitted for the degree of

Doctor of Philosophy

London, September 2010

Acknowledgements

Starting a PhD is like embarking upon a very long and uncertain journey. Support is a key component for success, and I am extremely grateful I have been helped so much throughout the long six years I spent writing this thesis.

I would like to thank the Statistics Department at the London School of Economics and Political Sciences that allowed me to undertake an especially difficult part-time PhD. I would like to acknowledge in particular the academic help I received from Pr Ragnar Norberg, Pr Qiwei Yao, Pr Umut Cetin and Pr Henry Wynn. I also would like to thank the administrative staff of the department and in particular Ian Marshall. Thank to my fellow LSE PhD students; even if I did not meet with them often, they still made me feel I was part of the research department.

I would like to acknowledge the financial support of the CREST in Paris for the grant I received for the first two years of my PhD, and especially Pr Christian Gourieroux, Pr Laurence Lescourret and Pr Nizar Touzi.

I also would like to thank my current employer Credit Suisse, which financed the remaining four years of my research, and gave me enough flexibility in my job to be able to complete my thesis, despite a very challenging and demanding environment. I particularly would like to thank all my colleagues and my managers who were supportive and allowed me to take my Fridays to work on my PhD for three years.

I especially would like to thank Yves Bentz, my previous manager at Credit Suisse, who initially offered me to work at the bank while completing my PhD. He has been a real inspiration and support in all respects.

I also would like to acknowledge the help and advice I received from Pr Monique Jean-Blanc, Pr Bernard Sinclair-Desgagne and especially Pr Leonard Smith and Pr Rudiger Kiesel.

I thank all my friends, and especially David Le Bris and Judith Bydlowski who devoted time reading parts of my PhD and giving me sensible and most welcome advice.

But first and foremost, I would like to thank my PhD supervisor, Pr Pauline Barrieu. I will never be able to express enough gratitude. She has been incredibly inspiring, guiding me through every step, never letting me give up and encouraging me in my most difficult times. She became a very dear and important person to me and I owe her much more than the great achievement of finishing my PhD thesis.

I thank warmly my family, my parents and my brother and my parents in law who gave me confidence and emotional support.

Finally, I would like to thank my husband Peter. He supported me the most completely and always believed I would finish my PhD thesis even when I was convinced I could never do it. He gave me an infinite supply of time, trust and love and I would like to dedicate my research to him, as I could not have done it without his unconditional support. I also dedicate my PhD thesis to our two lovely baby daughters Laura and Emma.

Abstract

A decision maker, when facing a decision problem, often considers several models to represent the outcomes of the decision variable considered. More often than not, the decision maker does not trust fully any of those models and hence displays ambiguity or model uncertainty aversion.

In this PhD thesis, focus is given to the specific case of asset allocation problem under ambiguity faced by financial investors. The aim is not to find an optimal solution for the investor, but rather come up with a general methodology that can be applied in particular to the asset allocation problem and allows the investor to find a tractable, easy to compute solution for this problem, taking into account ambiguity.

This PhD thesis is structured as follows: First, some classical and widely used models to represent asset returns are presented. It is shown that the performance of the asset portfolios built using those single models is very volatile. No model performs better than the others consistently over the period considered, which gives empirical evidence that: no model can be fully trusted over the long run and that several models are needed to achieve the best asset allocation possible. Therefore, the classical portfolio theory must be adapted to take into account ambiguity or model uncertainty. Many authors have in an early stage attempted to include ambiguity aversion in the asset allocation problem. A review of the literature is studied to outline the main models proposed. However, those models often

lack flexibility and tractability. The search for an optimal solution to the asset allocation problem when considering ambiguity aversion is often difficult to apply in practice on large dimension problems, as the ones faced by modern financial investors. This constitutes the motivation to put forward a novel methodology easily applicable, robust, flexible and tractable. The Ambiguity Robust Adjustment (ARA) methodology is theoretically presented and then tested on a large empirical data set. Several forms of the ARA are considered and tested. Empirical evidence demonstrates that the ARA methodology improves portfolio performances greatly.

Through the specific illustration of the asset allocation problem in finance, this PhD thesis proposes a new general methodology that will hopefully help decision makers to solve numerous different problems under ambiguity.

Contents

List of Figures

Chapter 1

Introduction

"*Success is not final, failure is not fatal: it is the courage to continue that counts.*"

Sir Winston Churchill.

A decision-maker, when facing a decision problem, often considers several models in order to represent the possible outcomes of the decision variables considered. More often than not, the decision-maker does not fully trust any of the models considered, and, hence, displays ambiguity, or model uncertainty aversion. In this PhD thesis, the generic terminology 'model uncertainty', or 'ambiguity' is used to refer to any situation where a decision-maker has to consider different models, different scenarii, or, has to rely on different experts (that may all be wrong, or at least subject to error), to come to a decision. The general term 'prior' describes a model, scenario, or expert's opinion. The question of how ambiguity aversion should be taken into account in the decision-making problem has, therefore, arisen, and has now become crucial in many scientific fields (including but not limited to: economics, biology, physics, climatology and finance).

There is a substantial body of literature on the problem of decision-making under ambiguity. The classical approaches have significant limitations: they often prove to be very difficult and challenging to implement in practice. To overcome these limitations, a methodology to operate a trade-off between robustness and optimisation has been proposed. Since the complexity of most practical frameworks makes this task almost impossible with current limitations, the objective has not been to find the "optimal" decision for the decision-maker; the focus has been, rather, on a robust approach that allows the decision-maker to combine different priors in a practical and tractable way - and to take the best decision - in a robust sense (i.e. that can be easily adapted to the different types and number of priors considered). The question is, really, about finding a solution that encompasses all the different information given by the different priors, but also the ambiguity the decision-maker faces regarding the set of priors; in other words, finding a robust decision rule, as defined by Levin & Williams (2003): a rule that "although not exactly optimal for any prior, yields outcomes that are acceptable to all priors". The contribution of this PhD thesis is therefore to provide the decision maker with a novel methodology that deals with model ambiguity in an original fashion.

In this thesis, focus is given to the financial field to illustrate decision-making under ambiguity. The novel methodology proposed in this PhD theisi is applied to the asset allocation problem of an investor, when ambiguous about the models used to describe the asset returns distribution, providing a practical, systematic algorithm to trade a large portfolio of assets.

The Modern Portfolio Theory, initiated with the classical Markowitz framework, aims at solving the asset allocation problem. Many authors have, since then, considered more complex settings, allowing the investor to take into account several models for risky asset return distributions. Indeed, many different models can be used in finance to represent asset returns: very quantitative models, as well as more qualitative ones. The uncertainty about which model to use adds complexity to the asset allocation problem. The main idea underlying the asset allocation problem is that an investor needs to balance the risk they are willing to take and the return expected from the invested portfolio. Ideally, the aim of the investor is to come up with the optimal portfolio (i.e. generally speaking the preferred portfolio allocation, depending on the investor own preferences), which perfectly represents the risk-return equilibrium required by the investor.

However, an optimal solution is all the more hard to find, as the investor considers different models to represent asset returns (and, therefore, the anticipated portfolio performance).

During this research, the aim has not been to find an optimal solution for the investor facing an asset allocation problem, but, rather, to come up with a general methodology that can be applied, in particular, to the asset allocation problem - that allows the investor to find a tractable, easy to compute solution for this problem - taking into account aversion to model uncertainty, otherwise called ambiguity.

This PhD thesis is structured as follows. First, some classical and widely used models that represent asset returns are presented and discussed. It is shown that the performance of the asset portfolios built using those single models is very volatile. No model performs consistently better than the others over the period considered, which gives empirical evidence that: no model can be fully trusted over the long-run, and that several models are needed to achieve the best asset allocation possible. Therefore, the classical portfolio theory must be adapted to take into account ambiguity or model uncertainty. Many authors have, in the early stages, attempted to include ambiguity aversion in the asset allocation problem. However, those models often lack flexibility and tractability. A review of the literature is performed to outline the main models proposed. The search for an optimal solution to the asset allocation problem, when considering ambiguity aversion, is, in practice, often difficult to apply to large dimension problems, such as the ones faced by modern financial investors. This constitutes the motivation to put forward a novel methodology that is easily applicable, robust, flexible and tractable. The remaining chapters of this PhD thesis present and test this new approach. The Ambiguity Robust Adjustment (ARA) methodology is presented theoretically, and, then tested on a large empirical data-set. Several forms of the ARA are considered and tested. Empirical evidence demonstrates that the ARA methodology improves portfolio performances greatly.

This PhD thesis is organised according to six different chapters:

- In the second Chapter, focus is given to the Modern Portfolio Theory; a general framework, for this PhD thesis, is outlined, and some classical asset allocation problems, involving a sole model to represent asset returns such as the Markowitz mean-variance optimal allocation, the Sharpe Capital Asset Pricing Model or the Ross Asset Pricing Theory, are described. Performance measures, used to evaluate and compare different portfolio allocations, are also presented.

- In the third Chapter, other types of asset return models, widely used among practitioners are detailed, such as fundamental or statistical factor models, encompassing the Capital Asset Pricing Model (CAPM) and different versions of the Asset Pricing Theory (APT). The performance of the models is tested by a number of different performance measures. Empirically, none of the models can be considered as the best over a long time-period: the performance measures vary greatly over time, which provides further evidence of the model uncertainty problem faced by a financial investor.

- In the fourth Chapter, focus is given to the theoretical approaches presented in the literature to date, which include model ambiguity aversion into asset allocation problems. A more formal definition for the concept of ambiguity is given, and the main models used to incorporate ambiguity in portfolio allocation problems are recalled.

- In the fifth Chapter, the novel Ambiguity Robust Adjustment (ARA) methodology is presented. The central idea is that it is extremely challenging to compute a closed form solution, or numerical solution, for the asset allocation optimisation problem when several priors are considered. More often than not, the priors considered do not belong to the same class of models (different parametric/ non parametric models) and, therefore, it can be, even, impossible to precisely define the optimisation problem under a theoretical form. That is why a more ad hoc, practical methodology, is proposed that is altogether easier to compute, more flexible (in terms of the type and number of prior models that can be considered) and tractable (the ARA methodology allows the investor to measure precisely aversion to ambiguity towards a specific prior but also towards the overall set of priors considered). In principle, the ambiguity aversion is decomposed into two types of ambiguity: the absolute ambiguity aversion towards a given model and the relative ambiguity aversion towards the set of models considered. More precisely, this two-step methodology takes, as input, the allocations inferred by the different priors as if they were the only model to consider (those weights can be computed through optimisation, they can be inferred by a qualitative approach). Those weights are first adjusted through an Absolute Robust Ambiguity Adjustment function (ARAA), which allows the investor to express absolute ambiguity towards a given model. Then, the different set of weights, corresponding to the different models, are mixed through a Relative Ambiguity Robust Adjustment (RARA) function that expresses the overall ambiguity of the investor toward the set of priors considered. The ARA methodology is compared to recent approaches of optimisation under ambiguity and a theoretical example is proposed as an illustration.

- In the sixth Chapter, an empirical study is conducted on European empirical data; the performance of the classical portfolios presented in Chapter 2, as well as the Savage Subjected Expected Utility portfolio (basically, a linear blending of the classical portfolios) and the ARA portfolio, are displayed. Due to the high-dimensionality of the asset allocation problem, in practice (financial investors often consider portfolios of hundreds of assets), a simple, tractable methodology is needed. Effectively, the Ambiguity Robust Adjustment is easily applicable to large dimension, complex empirical problems. It has been found, through the empirical study, that the SEU portfolio outperforms almost all of the single strategies by all performance measures considered. This means, that blending the different strategies allows the investor to achieve a smoother, more reliable portfolio performance. It is also shown, that the ARA portfolio beats the SEU portfolio performances, consistently, proving that the ARA methodology is easily applicable to the large-dimension problem considered in this study; and taking into account ambiguity in the asset allocation problem, greatly improves the empirical portfolio performances.

- In the last Chapter, the novel ARA methodology is enhanced by investigating forms for the RARA function, that are more complex than the linear form proposed in the precedent empirical study; the RARA function is calibrated through the non-parametric Support Vector Machine (SVM) methodology, or fitted, a priori, with respect to some nonlinear properties. Indeed, ambiguity aversion implies some nonlinear effects, and taking them into account allows the investor to further enhance portfolio performances, as shown in the empirical tests, which are conducted and presented in this chapter.

This PhD thesis proposes a, new, general methodology that is designed to contribute to discourses and practices of decision-making under ambiguity. The robust approach proposed illustrates and is applied to financial fields, but is not restrictive. Indeed, the approach could be used, for instance, to meet the specific attributes and needs of various research and practice areas, including, but not limited to, financial and actuarial risk management, environmental policy, monetary policy and technology management. Individual decision-making and collective decision-making can be undertaken using the proposed methodology, since it does not rely on specificities of any particular choice criterion.

Chapter 2

Classical Approaches to the Asset Allocation Problem

"*The process of selecting a portfolio may be divided into two stages. The first stage starts with observation and experience and ends with beliefs about the future performance of available securities. The second stage starts with the relevant beliefs about future performances and ends with the choice of portfolio.*"

Harry Markowitz, Portfolio Selection, The Journal of Finance, 1952.

When facing asset allocation problems, financial investors aim to allocate their initial wealth optimally across financial assets, i.e they want find the allocation that best fits their preferences. To define the portfolio of assets, an investor requires a model to represent asset returns. In practice investors can employ a variety of models, including the classical models presented in this Chapter. The following section will comprehensively describe the three most established Modern Portfolio Theories: the Efficient Frontier developed by Markowitz (1952), the Capital Asset Pricing Model developed by Sharpe (1964) and the Asset Pricing Theory more recently proposed by Ross (1976).

Modern portfolio theories aim to solve asset allocation problems faced by financial investors. As Markowitz notes, the portfolio selection problem is constituted by several phases: the observational phase, where investors empirically observe price dynamics; the modelling phase, reflecting inferred investor beliefs from the initial observation; and finally a decision phase, where informed investors express their preferences:

1 **Observation phase**: signifies investor observes financial asset prices

2 **Modelling phase**: describes investor beliefs concerning financial market uncertainty, and encompassing:

   – *the set of possible states of the world*: i.e. the set of definitions for asset prices.

   – *the asset price dynamics*: i.e. the distribution measure the investor believes to lead asset prices.

   – *how asset prices reflect the flow of information*: market efficiency is commonly assumed, i.e. asset prices fully reflect available information, representing true investment values.

3 **Decision phase**: defines the investor decision-making procedure under risk and uncertainty; specifying:

   – *the investor preferences*: commonly defined through a utility function, that takes into account investor risk aversion.

   – *the investor valuation function*: often expressed through the expected utility framework as the investor discounted expected final wealth utility.

A persistent and major assumption within portfolio optimisation problem settings has been that investors are able to accurately model uncertainty by attributing the right probability measure leading asset prices. However, the addition of a fourth phase to the portfolio selection procedure introduced in the early stages of modern finance created a fundamental distinction between uncertainty and ambiguity (see Knight (1921), as discussed in Chapter 4):

4 **Ambiguity adjustment phase**: investors determine a method to account for ambiguity (i.e. investor acts upon doubts both of investment beliefs expressed in the modelling phase and the ability to perfectly model asset prices dynamics).

The following chapter will focus on the (second) modelling phase and (third) decision phase of the classical portfolio selection problem. The (first) observation and (fourth) ambiguity adjustment phases will be further discussed in Chapters 3 and 5 respectively.

The first section of this chapter will outline the key definitions utilised throughout this thesis. Furthermore it will describe the settings of the portfolio selection problems considered throughout this research. Important results are proved in the text; and additional proofs can be found in the Appendix. The second section will describe the Efficient Frontier (which represents the set of efficient portfolios) proposed by Markowitz (1952). A particular focus will be placed upon the introduction of two efficient portfolios: the fully invested Minimum Variance portfolio and the Maximum Sharpe portfolio, tested in Chapter 3. The third section will describe the equilibrium theory and the Capital Asset Pricing Model developed by Sharpe (1964), also tested in Chapter 3. The fourth section will focus on the more general Arbitrage Pricing Theory introduced by Ross (1976), which forms the basis for all modern factor models, including some models considered in the next chapter. The final section will introduce the performance measures that are used to compare the portfolio performances discussed throughout the remainder of this thesis.

2.1 Framework

The aim of this section is to define the portfolio allocation problem in greater detail. First the framework and appropriate notations will be precisely described for the financial market considered throughout this PhD thesis. The second section will provide the definition of a trading strategy (i.e. the formal description of a portfolio allocation). The third sub-section will describe the classical decision under risk procedures that are used by investors to effectively discriminate between different trading strategies. Finally, the classical optimisation problem will be formally specified.

2.1.1 Notation and setting

Unless otherwise specified, the following key assumptions and notations are used throughout the PhD thesis.

- **Time horizon :** Static one period models are considered; the corresponding investment horizon is taken to be a finite and unique time horizon T. It is assumed that there is one single period $[0; T]$. At time 0, the investors make their investment decisions, and at time T they observe the value of their portfolio.

- **Financial Market :** It is assumed that financial market uncertainty is modelled using a standard probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, where:

  - The set Ω represents the set of all possible states of the world.
  - The σ-field $\mathfrak{F}$ represents the structure of available information on the financial market at time T.
  - $\mathbb{P}$ stands for the true probability measure of the financial market considered according to the set of possible states of the world:

    $$\mathbb{P} \in \mathcal{M}(\Omega, \mathfrak{F})$$

    where $\mathcal{M}(\Omega, \mathfrak{F})$ stands for the set of measurable functions from Ω to $\mathfrak{F}$. In the context of a risky, non-ambiguous framework, the objective probability $\mathbb{P}$ is known. However, in the context of a risky, ambiguous framework the objective probability is inferred as it is not known by investors.

- **Financial Assets :** There are $N + 1$ primary assets traded between date 0 and T, consisting of two different types:

  - **Risky assets :** It is assumed that there are N risky assets in the financial market. Their prices at time T, denoted by $\mathbf{s}_T = (s_T^1, ..., s_T^N)$, are $\mathfrak{F}$-measurable.

- **Risk-free asset:** A risk-free asset also exists. The risk-free asset price at time T is denoted by s_T^0. It is assumed that the price of the risk-free asset is deterministic (and in particular: non-ambiguous). The constant instantaneous risk-free rate is denoted by r_f.

As a standard assumption, financial assets are considered to be exchanged in a friction-free financial market. The following additional standard assumptions are made:

- No transaction costs are generated when buying or selling financial assets: exchanges and broker's fees are disregarded[1].
- Asset prices are infinitely divisible. Price granularity, such as lot size (minimum amount of shares to be exchanged in one transaction) or tick size (minimum price granularity authorised by the exchanges) is disregarded.
- There is an unlimited and costless liquidity: any amount of financial assets can be exchanged, bought or sold at market price without any price impact. In particular, an agent can short-sell any asset without cost (i.e. borrowing costs and short selling regulation limiting the amount of asset shares to be sold without cover are disregarded).

Finally, a number of general notations will be applied consistently throughout the thesis. The following denotes:

- A scalar, or one-dimensional random variable by a simple letter (e.g. a),
- A vector, or multi-dimensional random variable by a bold letter (e.g. $\mathbf{a}$)[2],
- A matrix by a bold capital letter (e.g. $\mathbf{A}$),
- A time index as a subscript (e.g. a_t is the value of the variable a at time t)
- An asset or portfolio index as an upper script (e.g. a^i is the value of the variable a for the asset i)

[1] However, it should be noted that in empirical tests some transaction costs are introduced to give more realistic results.

[2] $\mathbf{a}'$ denotes the transpose of the vector $\mathbf{a}$

- For the probability measure $\mathbb{P}$, the expectancy operator is denoted as $\mathbb{E}_{\mathbb{P}}$, the variance operator as $\mathbb{V}_{\mathbb{P}}$ and the covariance operator as $\mathbb{COV}_{\mathbb{P}}$. Moreover, if the reference probability $\mathbb{P}$ is obvious, they are denoted as $\mathbb{E}$, $\mathbb{V}$ and $\mathbb{COV}$.

- To simplify notations whenever needed, the expected value of the variable a can be denoted μ^a and its standard deviation σ^a. The covariance of two variables a and b can be denoted as $\sigma^{a,b}$ and their correlation coefficient as $\rho^{a,b}$.

2.1.2 Trading Strategy

This section provides a definition of portfolio allocation (more formally described as a trading strategy). Put simply, a financial investor with a given initial wealth denoted by x_0, wants to allocate their initial wealth among the different financial assets available. A given allocation $\boldsymbol{\phi}$ (that is also called a trading strategy) is defined as:

Definition 2.1. ***Trading Strategy***

The trading strategy $\boldsymbol{\phi}$ assigns the set of weights or asset allocation $(\phi^0, ..., \phi^N)$ to the $N+1$ financial assets at date 0. The various components ϕ^i for $i = 0, 1, ...N$ represent the proportional cash units invested in the security i. Negative as well as positive real values are assumed, reflecting assumptions concerning short-selling and asset divisibility. The allocation $\boldsymbol{\phi}$ is defined according to the information available up to the initial date of portfolio rebalance (i.e., when the investor defines their asset allocation at time 0 for the period $[0;T]$). The value at time T of the portfolio $\boldsymbol{\phi}$ will be denoted by $x_T^{\boldsymbol{\phi}}$:

$$x_T^{\boldsymbol{\phi}} = \sum_{i=0}^{N} \phi^i s_T^i$$

In order to decide which trading strategy is the best, an investor needs to specify a set of decision preferences. The following section describes the classical procedure of decision under risk. Preferences are expressed through a utility function and the value function is defined as the expected utility of the investor's terminal wealth (i.e. the value of the investor's portfolio at the horizon time T).

2.1.3 Decision under risk

Investors need to be able to compare differing asset allocations and choose which one is best. They consider a preference relation between different investment alternatives. This preference relation allows them to discriminate the different investment options they have. Therefore, they can choose the strategy that maximises their preferences. The utility functions translate those preferences into numerical values that can then be used in optimisation problem modelling.
Under uncertainty (i.e. the risky financial asset prices are random), the investor needs to evaluate a certainty equivalent value of their preferences. Indeed, the investor has to be able to compare the different outcomes of their portfolio choices so that they can make a decision. In the classical framework presented in this chapter, the investor relies on the certainty equivalent to compare random payouts, and uses the von Neumann-Morgenstern expected utility maximisation as a decision criterion.
More precisely, the investor terminal wealth (the quantity x_T^{ϕ}) is random, and depends on the vector of asset prices $\mathbf{s}_T$ at horizon T. However, the expected utility of this quantity is certain. Therefore, the certainty equivalent of x_T^{ϕ} denoted $c(x_T^{\phi})$ is defined such that:

$$u[c(x_T^{\phi})] = \mathbb{E}_{\mathbb{P}}[u(x_T^{\phi})]$$

where u is a concave, increasing utility function. And the criterion used by the investor can be described as the value function:

$$V(\boldsymbol{\phi}) \equiv \mathbb{E}_{\mathbb{P}}[u(x_T^{\phi})]$$

To obtain the optimal portfolio under these settings, the investor needs to maximise the value function V over all the possible asset allocations $\boldsymbol{\phi}$. The classical portfolio optimisation problem can therefore be described as follows.

2.1.4 The classical portfolio optimisation problem

The portfolio optimisation problem is solved by Markowitz (1952), assuming the investor knows the true probability distribution $\mathbb{P}$. What differentiates investors is their attitude towards risk (represented by their risk aversion parameter λ). In the case of a classical von Neumann-Morgenstern utility maximisation setting, the decision maker problem can be formalised as:

$$\max_{\phi} \mathbb{E}_{\mathbb{P}}[u(x_T^{\phi}, \lambda)] \tag{2.1}$$

where either u is a quadratic function ($u(x, \lambda) = x - \lambda x^2$), or the dynamic of x is normal, so that the von Neumann-Morgenstern value function defined as:

$$V(x_T^{\phi}) \equiv \mathbb{E}[u(x_T^{\phi}, \lambda)]$$

only depends on the first two moments (the mean and variance) of the terminal wealth x_T^{ϕ} distribution, as is detailed in the next section.

Now that the framework of study and the classical asset allocation problem faced by financial investors has been described, a detailed investigation of the three most famous approaches of portfolio selection can take place; starting with the Efficient Frontier of Markowitz. The construction of particular portfolio allocations is proved in the Appendix section.

2.2 The Efficient Frontier, Markowitz (1952)

The Markowitz Efficient Frontier provides the foundation for single-period investment theory. It explicitly addresses the trade-off between the expected and variance values for the rate of return of a given portfolio. Any efficient portfolio lying on the Efficient Frontier can be expressed as a convex combination of two given efficient portfolios ("Two Fund Theorem") or as a linear combination of the tangent portfolio and the risk-free asset, if such a risk-free asset exists ("One Fund Theorem"). The section is organised as follows. In a first sub-section, the hypothesis of the Markowitz framework is detailed: the mean-variance paradigm and the diversification effect. Then, the Efficient Frontier Equation, and a formal description of two classical efficient portfolios are given: the Minimum Variance portfolio and the Maximum Sharpe portfolio. In the third sub-section, a risk-free asset is introduced. Furthermore, the -in this case- simplified equation of the

Efficient Frontier is given, which can be entirely expressed with the knowledge of a single portfolio: the Tangential portfolio.

2.2.1 Background

The Markowitz Frontier solves the asset allocation problem under the assumption that any investor believes in the mean-variance paradigm. More specifically, only the first two moments of a portfolio return (the mean and variance) are significant to define the best allocation. The diversification effect justifies the mean-variance paradigm as explained in the following.

2.2.1.1 The mean-variance paradigm

The grounds for the Markowitz mean-variance paradigm can be expressed through the following concept:
Higher expected returns come with greater risk, and lower expected returns come with lesser risk, where the risk is measured by the variance of an investor portfolio.
Assuming the investor intends to optimise their portfolio asset allocation, there are two equivalent ways of proceeding: either by maximising the expected return of the portfolio under the constraint that the portfolio variance remains below a certain risk tolerance level, or by minimising the risk (i.e. the variance of the portfolio), given the level of portfolio return intended to be achieved.
There is a strong underlying assumption required to justify the mean-variance paradigm. This is that either the investor preferences are described by a quadratic utility function (only the first two moments of the returns distribution are significant), or that the asset returns are normally distributed (their distribution is entirely defined by their first two moments).
Markowitz offers justification for the mean-variance paradigm on the basis that it complies with the benefits of diversification (as the number of uncorrelated assets with identical return distribution in the portfolio increases, the portfolio standard deviation decreases, whereas; the portfolio expected return converges towards the assets common expected return).

2.2.1.2 The concept of diversification

Two simple illustrations will be used to provide a formal demonstration of the diversification effect. The random return of the asset i over the period $[0;T]$ will be denoted by $r^i \equiv \frac{s^i_T - s^i_0}{s^i_0}$. Let us denote by $\boldsymbol{\mu}$ the $N \times 1$ vector of risky asset returns mean and by $\boldsymbol{\Sigma}$ the $N \times N$ covariance matrix of the asset returns. μ^i denotes the mean of the return r^i and σ^i denotes its standard deviation.
Two simple examples will be considered:

Situation A: It is assumed that the asset returns are mutually independent and follow a normal distribution $N(\mu^i, \sigma^i)$ with mean μ^i and standard deviation σ^i. It is assumed that the mean and standard deviation are bounded for any risky asset i:

$$\begin{cases} \mu_{min} < \mu^i < \mu_{max} \\ \sigma_{min} < \sigma^i < \sigma_{max} \end{cases}$$

The Equally Weighted portfolio, where for any risky asset i, $\phi^i = \frac{1}{N}$ will be considered. The return of the portfolio $\boldsymbol{\phi}$ will be denoted by $r^\phi \equiv \frac{x^\phi_T - x^\phi_0}{x^\phi_0}$. For the sake of simplicity, all initial prices are assumed to be set to 1, resulting in:

$$\mathbb{E}(r^\phi) = \sum_{i=1}^N \frac{\mu^i}{N} > \mu_{min} \text{ and } \mathbb{V}(r^\phi) = \sum_{i=1}^N \frac{(\sigma^i)^2}{N^2} < \frac{\sigma^2_{max}}{N}$$

When N becomes very large, the portfolio return is bounded from below by μ_{min} and its variance is bounded from above by a quantity that converges to 0. The effect of diversification is fully observed: when the number of uncorrelated assets in the portfolio increases, the expected return of the portfolio converges to a value greater than the minimum expected asset return, while the portfolio standard deviation (assimilated to its risk) converges to zero.

Situation B : now, it will be assumed that the asset returns are correlated. The covariance between the returns of the assets i and j are denoted by $\sigma^{i,j}$; it is assumed that for any risky assets i and j:

$$\sigma^{i,j} > \sigma^2_{min}$$

Therefore:

$$\mathbb{E}(r^{\phi}) = \sum_{i=1}^{N} \frac{\mu^i}{N} > \mu_{min} \text{ and } \mathbb{V}(r^{\phi}) = \sum_{i=1}^{N} \frac{(\sigma^i)^2}{N^2} + \sum_{i=1}^{N} \sum_{j \neq i} \frac{\sigma^{i,j}}{N^2} > \sigma^2_{min}$$

The diversification effect is limited: the portfolio standard deviation is bounded below by σ^2_{min}, whatever the number of assets N included in the portfolio $\boldsymbol{\phi}$. The correlation of the asset returns limits the diversification effect. However, as long as the asset returns are not 100% correlated, the diversification effect still diminishes the risk of the portfolio when the number of assets increases.

Diversification allows the investor to reduce variance; therefore, reducing risk, and increasing the investor's future wealth expected utility.

2.2.1.3 Efficient Frontier Definition

In the Markowitz framework, investors choose their portfolio among a common set of efficient portfolios defined as the efficient frontier, according to their risk aversion.

Definition 2.2. ***Efficient Portfolio***

A portfolio is said to be efficient if its variance is smaller than the variance of all the portfolios with the same expected return. Formally speaking, it is said that the portfolio represented by the asset allocation $\boldsymbol{\phi}$ is efficient if for any other allocation $\tilde{\boldsymbol{\phi}}$ the following is found:

$$\mathbb{E}(x_T^{\tilde{\boldsymbol{\phi}}}) = \mathbb{E}(x_T^{\boldsymbol{\phi}}) \Rightarrow \mathbb{V}(x_T^{\tilde{\boldsymbol{\phi}}}) > \mathbb{V}(x_T^{\boldsymbol{\phi}})$$

This leads to the following general definition of the Markowitz Efficient Frontier:

Definition 2.3. ***Efficient Frontier***

The Efficient Frontier is the set of all efficient portfolios.

More precisely, Markowitz establishes a mapping of expected returns and standard deviation (or risk) for any fully invested portfolio $\boldsymbol{\phi}$ (i.e. the weights $(\phi^i)_{1 \leq i \leq N}$ sum to one). The optimal asset allocation is found by the investor by moving along the Efficient Frontier according to either, their risk aversion (i.e.

the amount of risk he/she is willing to take), or the expected return they want to achieve. Therefore, the portfolio asset allocation problem can be synthesised by either one of the two optimisation problems:

Maximise expected return for a given risk level σ^P

$$\begin{cases} \max_{\boldsymbol{\phi}} \mathbb{E}(r^{\boldsymbol{\phi}}) \\ \text{s.t } \mathbb{V}(r^{\boldsymbol{\phi}}) = \sigma^P \\ \text{and } \boldsymbol{\phi}'\mathbf{1} = 1 \end{cases}$$

or

Minimise risk for a given expected return μ^P

$$\begin{cases} \min_{\boldsymbol{\phi}} \mathbb{V}(r^{\boldsymbol{\phi}}) \\ \text{s.t } \mathbb{E}(r^{\boldsymbol{\phi}}) = \mu^P \\ \text{and } \boldsymbol{\phi}'\mathbf{1} = 1 \end{cases}$$

The expected return and variance of the portfolio $\boldsymbol{\phi}$ are denoted by:

$$\mathbb{E}(r^{\boldsymbol{\phi}}) = \mu^{\boldsymbol{\phi}} \equiv \boldsymbol{\mu}'\boldsymbol{\phi} \text{ and } \mathbb{V}(r^{\boldsymbol{\phi}}) = (\sigma^{\boldsymbol{\phi}})^2 \equiv \boldsymbol{\phi}'\boldsymbol{\Sigma}\boldsymbol{\phi}$$

To recall, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ stand respectively for the empirical mean and covariance matrix of the asset returns.
In the following section, the formal equation of the Efficient Frontier is provided; cases with and without a risk-free asset will be discussed.

2.2.2 The Efficient Frontier without a risk-free asset

To present the equation for the efficient frontier, a description will first be made of the Minimum Variance portfolio (MN) and then the fully invested Maximum Sharpe portfolio (MS); two efficient portfolios of particular interest.

2.2.2.1 The Minimum Variance portfolio

The efficient portfolio that has the smallest variance of all efficient portfolios will be considered. This portfolio represents the minimum amount of risk an investor must be ready to take when investing on the financial markets.

Proposition 2.1 (Minimum Variance Portfolio).
If the Minimum Variance portfolio allocation is denoted by $\boldsymbol{\phi}^{MN}$, then $\boldsymbol{\phi}^{MN}$ is the solution for the following problem:

$$\begin{cases} \min_{\boldsymbol{\phi}} \frac{1}{2}\boldsymbol{\phi}'\boldsymbol{\Sigma}\boldsymbol{\phi} \\ s.t\ \boldsymbol{\phi}'\mathbf{1} = 1 \end{cases}$$

The first two moments of the portfolio return $r^{\boldsymbol{\phi}^{MN}}$ are:

$$\mathbb{E}(r^{\boldsymbol{\phi}^{MN}}) = \frac{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}} \text{ and } \mathbb{V}(r^{\boldsymbol{\phi}^{MN}}) = \frac{1}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}$$

-and the Minimum Variance Portfolio allocation is $\boldsymbol{\phi}^{MN} = \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}$.

Proof. See Appendix. □

2.2.2.2 The Maximum Sharpe ratio portfolio

The Sharpe ratio, as detailed in Sharpe (1994), is one of the most popular measures of portfolio performance. The Sharpe ratio represents the average return per unit of risk (risk being defined as the portfolio standard deviation) and therefore complies with the mean-variance paradigm (only the first two moments of the portfolio return distribution matter to evaluate the portfolio performance). For any portfolio $\boldsymbol{\phi}$, the Sharpe ratio is defined as:

$$Sharpe^{\boldsymbol{\phi}} = \frac{\mathbb{E}(r^{\boldsymbol{\phi}})}{\sqrt{\mathbb{V}(r^{\boldsymbol{\phi}})}} = \frac{\mu^{r^{\boldsymbol{\phi}}}}{\sigma^{r^{\boldsymbol{\phi}}}}$$

More details will be given in Section (2.5), when additional performance measures will also be introduced. The efficient frontier is actually the set of portfolios that maximise the Sharpe ratio for any given expected return (fixing μ^{ϕ}, the Sharpe is maximised when σ^{ϕ} is minimised).

If the portfolio expected return $\mathbb{E}(r^{\phi}) = \mu^P$ is fixed, the maximum Sharpe ratio allocation associated with this portfolio is defined as the allocation that minimises the standard deviation of a portfolio of expected value μ^P:

Proposition 2.2 (Maximum Sharpe Ratio Portfolio)**.**
If $\boldsymbol{\phi}^{MS}$ denotes the Maximum Sharpe ratio portfolio allocation with expected return μ^{MS}, then $\boldsymbol{\phi}^{MS}$ is the solution of the following problem:

$$\begin{cases} \min_{\boldsymbol{\phi}} \frac{1}{2}\boldsymbol{\phi}'\boldsymbol{\Sigma}\boldsymbol{\phi} \\ s.t\ \boldsymbol{\phi}'\mu = \mu^{MS} \end{cases}$$

The first two moments of the portfolio return $\boldsymbol{\phi}^{MS}$ are:

$$\mathbb{E}(r^{\boldsymbol{\phi}^{MS}}) = \mu^{MS} \text{ and } \mathbb{V}(r^{\boldsymbol{\phi}^{MS}}) = \frac{(\mu^{MS})^2}{c}$$

Where $c = \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$.
The Maximum Sharpe Portfolio allocation is $\boldsymbol{\phi}^{MS} = \frac{\mu^{MS}}{c}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$.

Proof. See Appendix. □

Note that the set of Maximum Sharpe portfolios forms a straight line in the risk-return map (σ, μ), with the Equation:

$$\sigma = \frac{1}{\sqrt{c}}\mu \tag{2.2}$$

One particular Maximum Sharpe portfolio will be described. The Fully Invested Maximum Sharpe portfolio is the portfolio that is said to lay on the tangent of both: the Efficient Frontier of fully invested portfolios with the Equation (2.3), that is explicitly outlined below, and the Maximum Sharpe portfolios line with Equation (2.2). In the remainder of this thesis, MS will denote the Fully Invested Maximum Sharpe portfolio.

Proposition 2.3 (Fully Invested Maximum Sharpe Portfolio)**.**
If $\boldsymbol{\phi}^{MS}$ denotes the fully invested Maximum Sharpe portfolio allocation, then the mean μ^{MS} and standard deviation σ^{MS} of this portfolio must respect the Equations (2.3) and (2.2):

$$\begin{cases} (1)\ \sigma^{MS} = \sqrt{\frac{a}{d}(\mu^{MS} - \frac{b}{a})^2 + \frac{1}{a}} \\ (2)\ \sigma^{MS} = \frac{1}{\sqrt{c}}\mu^{MS} \end{cases}$$

Where it is denoted:

$$\begin{cases} a = \mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1} \\ b = \mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} = \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\mathbf{1} \\ c = \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \\ d = ac - b^2 \end{cases}$$

The first two moments of the portfolio return $\boldsymbol{\phi}^{MS}$ *are:*

$$\mathbb{E}(r^{\phi^{MS}}) = \tfrac{c}{b} \textit{ and } \mathbb{V}(r^{\phi^{MS}}) = \tfrac{c}{b^2}$$

The fully invested Maximum Sharpe Portfolio allocation is: $\phi^{MS} = \frac{\Sigma^{-1}\boldsymbol{\mu}}{\mu'\Sigma^{-1}\mathbf{1}}$.

Proof. See Appendix. □

2.2.2.3 The Two Fund Theorem.

The solution to either the variance-minimisation- or mean-maximisation-problem can be used to find the relation that the Efficient Frontier Equation establishes between the mean $\mathbb{E}(r^{\phi})$ and variance $\mathbb{V}(r^{\phi})$ of any fully invested efficient portfolio $\boldsymbol{\phi}$ when there is no risk-free asset.

Proposition 2.4 (Efficient Frontier Equation)**.**
Considering the variance minimisation problem:

$$\begin{cases} \min_{\boldsymbol{\phi}} \mathbb{V}(r^{\boldsymbol{\phi}}) \\ s.t\ \mathbb{E}(r^{\boldsymbol{\phi}}) = \mu^P \\ \textit{and } \boldsymbol{\phi}'\mathbf{1} = 1 \end{cases}$$

-it is found that the Efficient Frontier Equation is:

$$\mathbb{V}(r^{\phi^P}) = \frac{a}{d}[\mathbb{E}(r^{\phi^P}) - \frac{b}{a}]^2 + \frac{1}{a} \tag{2.3}$$

Proof. See Appendix. □

It is now possible to state the Two Fund Theorem proved by Merton (1973):

Theorem 2.1 (Two Fund Theorem)**.**
Any efficient portfolio can be duplicated in terms of mean and variance as a linear combination of any two efficient portfolios.
In particular, any efficient portfolio can be defined by the knowledge of the Minimum Variance and the fully invested Maximum Sharpe portfolios:

$$\mathbb{V}(r^{\phi^P}) = \mathbb{V}(r^{\phi^{MN}}) + \frac{\mathbb{V}(r^{\phi^{MN}})}{\mathbb{E}(r^{\phi^{MN}})[\mathbb{E}(r^{\phi^{MS}}) - \mathbb{E}(r^{\phi^{MN}})]}[\mathbb{E}(r^{\phi^P}) - \mathbb{E}(r^{\phi^{MN}})]^2$$

Proof. Apply Equation (2.3) to the Minimum Variance and Fully Invested Maximum Sharpe portfolios mean and variance. □

This result has dramatic implications: according to the Two Fund Theorem, an investor can replicate any efficient investment by investing solely in two efficient portfolios without purchasing individual stocks. However, this conclusion is based on the strong assumptions that investors only attribute significance to the mean and variance of their investment, and that only a single period is appropriate.
Figure (2.1) plots the efficient frontier for Eurostoxx constituent returns that have been computed from 2000 to 2010; the fully invested Maximum Sharpe portfolio (or Tangential portfolio); and the Minimum Variance portfolio. Individually speaking, it can be seen that the assets are all sub-efficient (they all lie below the efficient frontier); however, this is an illustration of the diversification benefit: indeed, an efficient portfolio can be constituted of non-efficient assets. Depending on either their expected return or risk aversion, the investor can choose any portfolio on the red curve.

2.2.3 Introduction of a risk-free asset

Assuming that the investor can freely borrow or lend money at a risk-free rate, denoted by r_f: a proportion of the investor initial wealth can be borrowed or lent at the risk-free rate. Thus, the portfolio optimisation problem for the investor becomes:

$$\begin{cases} \min_{\phi} \frac{1}{2}\boldsymbol{\phi}'\boldsymbol{\Sigma}\boldsymbol{\phi} \\ \text{s.t } \boldsymbol{\phi}'\mu + (1-\boldsymbol{\phi}'\mathbf{1})r_f = \mu^P \end{cases}$$

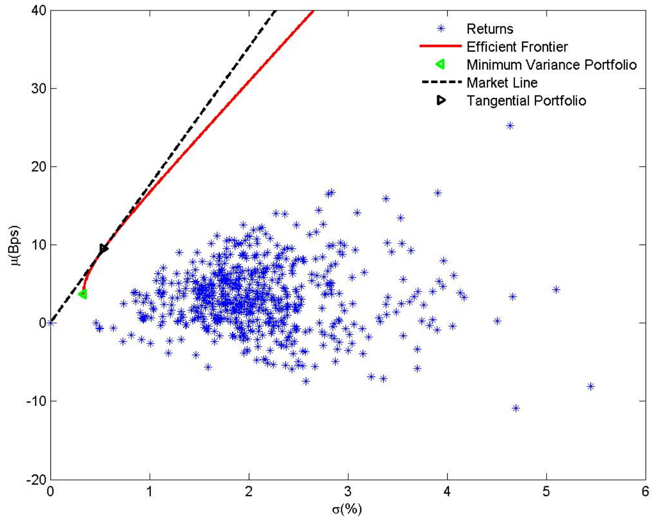


Figure 2.1: Efficient Frontier

Proposition 2.5 (Efficient Frontier Equation with a risk-free asset)**.** *It is found that in the case where a risk-free asset exists, the Markowitz efficient frontier is a straight line with equation:*

$$\mu^P = \sqrt{e}\sigma^P + r_f \tag{2.4}$$

with $e \equiv (\mu - r_f\mathbf{1})'\mathbf{\Sigma}^{-1}(\mu - r_f\mathbf{1})$.

Proof. See Appendix. □

2.2.3.1 Tangential portfolio

The Efficient Frontier line joins the risk-free rate with the Tangential portfolio. The Tangential portfolio, denoted by $\boldsymbol{\phi}^T$, has coordinates that comply with Equations (2.3) and (2.4).

Proposition 2.6 (Tangential Portfolio)**.**
If $\boldsymbol{\phi}^T$ *denotes the Tangential portfolio allocation; the mean and standard deviation of this portfolio are given by:*

$$\mathbb{E}(r^{\phi^T}) = \frac{e}{\sqrt{ae-d}} + r_f \text{ and } \mathbb{V}(r^{\phi^T}) = \frac{e}{ae-d}$$

-where:

$$\begin{cases} a = \mathbf{1}'\mathbf{\Sigma}^{-1}\mathbf{1} \\ b = \mathbf{1}'\mathbf{\Sigma}^{-1}\boldsymbol{\mu} = \boldsymbol{\mu}'\mathbf{\Sigma}^{-1}\mathbf{1} \\ c = \boldsymbol{\mu}'\mathbf{\Sigma}^{-1}\boldsymbol{\mu} \\ d = ac - b^2 \\ e = (\boldsymbol{\mu} - r_f\mathbf{1})'\mathbf{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1}) \end{cases}$$

Proof. See Appendix. □

2.2.3.2 The One Fund Theorem.

As any efficient portfolio lies on the efficient frontier line, it can be expressed in terms of mean and variance as a linear combination of the Tangential portfolio and the risk-free asset. This property is called the One Fund Theorem.

Theorem 2.2 (One Fund Theorem)**.**
When a risk-free asset exists, the mean and variance of any efficient portfolio can be defined as a linear combination of any efficient portfolio (in particular the Tangential portfolio) and the risk-free asset.
Thus, the Markowitz frontier Equation becomes:

$$\mathbb{E}(\boldsymbol{r}^{\phi^P}) = r_f + \frac{(\mathbb{E}(\boldsymbol{r}^{\phi^T}) - r_f)}{\sqrt{\mathbb{V}(\boldsymbol{r}^{\phi^T})}}\sqrt{\mathbb{V}(\boldsymbol{r}^{\phi^P})} \tag{2.5}$$

Proof. Apply Equation (2.4) to the Tangential portfolio mean and variance. □

This is illustrated in Figure (2.2), which sets $r_f = 1.5$ basis points.
Extending from the Markowitz Efficient Frontier, Sharpe deduces an equilibrium model that can be used to provide the correct price of a risky asset within the framework of the mean-variance setting. The model is described in the following section.

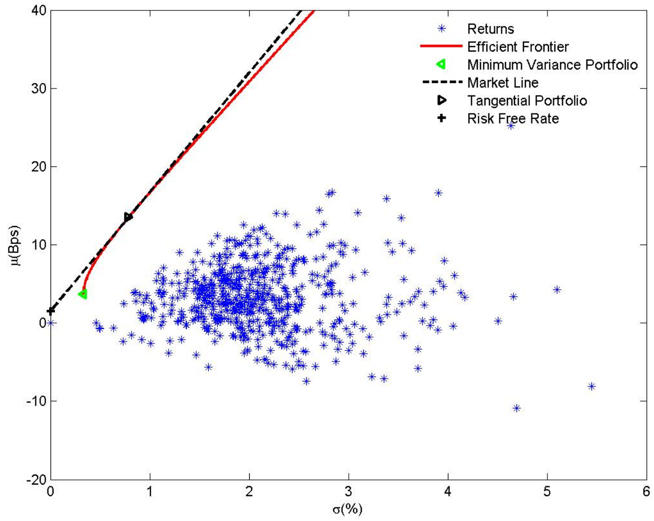


Figure 2.2: Efficient Frontier with a risk-free asset

2.3 The Capital Asset Pricing Model (CAPM), Sharpe (1964)

In what is considered to be a landmark paper, Sharpe (1964) extends the Markowitz model to a multi-agent setting. Proposing a global equilibrium model, known as the Capital Asset Pricing Model (CAPM), a mapping between risk, return and asset prices is enabled. It is shown, that if all investors anticipate similar expected returns and standard deviation of asset prices (and if the assumption of the Markowitz model is satisfied), then all asset returns must lie on the Security Market Line, which links expected return to risk. Thus, the CAPM gives a standard of comparison under the strong consensus assumption that all investors share the same view upon the distribution of asset returns.

The first sub-section will describe the Market portfolio and the Capital Market Line, which effectively corresponds to the the Tangential portfolio and the Markowitz Efficient Frontier described in Section 2.2.3. The equation of the equilibrium value for any risky asset in the context of the CAPM is then given. Finally, a description is formed of the CAPM model, and the specific performance measure developed byJensen (1969). This allows the CAPM portfolio tested in the following chapter to be constructed along with the Minimum Variance and Maximum Sharpe portfolios described in the previous section.

2.3.1 The Capital Market Line (CML)

Assuming that investors rely on the mean-variance paradigm, and there is complete agreement on the return distribution for the risky assets, it becomes possible to compute a unique equilibrium price for any efficient portfolio. The Markowitz frontier when computed for a representative agent, is from this point on referred to as the Capital Market Line, and is considered to apply to all investors in the financial market. In addition, the Tangential portfolio has been renamed the Market portfolio. Below, this portfolio will be described, after which the formal equation of the Capital Market Line will be outlined.

2.3.1.1 The Market Portfolio

By reference to the One Fund Theorem (if a risk-free asset exists), it is known that any investor can purchase a single portfolio, which is typically the Tangential portfolio. In addition, the investor can freely borrow or lend money at a risk-free rate to replicate any efficient portfolio. Furthermore, since in the CAPM assumptions all investors use the same probability measure $\mathbb{P}$ to represent the risky assets distribution: the same representative portfolio will be considered. This common "One Fund" or representative portfolio is referred to as the Market Portfolio. In actual fact, it represents a weighted average of all the risky assets weighted by their proportional market capitalisation. This result is based on an equilibrium argument: if the representative portfolio were not identical for investors sharing the same view on asset returns distribution, the price of assets in higher demand would rise and the price of assets in lower demand would fall. Ultimately, this will lead to a re-computation of the investors' representative portfolio converging towards the Market Portfolio. More formally, we have:

Proposition 2.7 (Market Portfolio). *In the CAPM model, the Tangential portfolio is effectively unique and common to all investors. It can be described as the Market portfolio (i.e. the allocation that weights risky assets according to their market capitalisation).*

Proof. The uniqueness of One Fund stems from the fact that all investors have identical anticipations about the distribution $\mathbb{P}$; and therefore, solve the same optimisation problem as specified in section 2.2.3. The market capitalisation will be denoted by q^i of the risky asset i, where $i \in [1, N]$ (i.e. the number of shares trading for the asset i multiplied by the price of the asset i). The Tangential portfolio allocation will be denoted by $\boldsymbol{\phi}^T$ and the Market portfolio allocation by $\boldsymbol{\phi}^M$. According to the definition of the Market portfolio the following is given:

$$\forall i \in [1, N],\ \phi^{M,i} \equiv \frac{q^i}{\sum_{i=1}^{N} q^i}$$

Furthermore, it will be assumed that there are J investors in the financial market. The amount detained by the investor $j \in [1, J]$, of any asset $i \in [1, N]$, is denoted as $\theta^{i,j}$. Therefore:

$$\sum_{j=1}^{J} \theta^{i,j} = q^i$$

The total market capitalisation of the asset i is equal to the sum of the individual investments made by investors in the risky asset i. If the initial wealth of the investor j invested in the Tangential portfolio $\boldsymbol{\phi}^T$ is denoted by x_0^j, the following is given:

$$\sum_{j=1}^{J} x_0^j = \sum_{i=1}^{N} q^i$$

In fact, the sum of wealth invested in the risky assets must be equal to the total market capitalisation.
According to the One Fund Theorem the following is given:

$$\forall i \in [1, N],\ \theta^{i,j} = \phi^{T,i} x_0^j$$

Therefore, by summing over all the investors, the following is obtained:

$$\forall i \in [1, N],\ \sum_{j=1}^{J} \theta^{i,j} = \phi^{T,i} \sum_{j=1}^{J} x_0^j \Rightarrow \phi^{T,i} = \frac{q^i}{\sum_{j=1}^{J} x_0^j} = \frac{q^i}{\sum_{i=1}^{N} q^i}$$

-which is precisely the allocation of the Market portfolio for the asset i.

□

The Market Portfolio is effectively the Market Index. Note: all tests carried out in this research are run on European data, and the Market Index is taken to be the Eurostoxx 600 index.

2.3.1.2 The CML Equation

In the CAPM framework, the Efficient Frontier in the $(\sigma - \mu)$ map is a straight line, emanating through the risk-free asset and passing through the Market portfolio. The Efficient Frontier is then termed the Capital Market Line; the equation of which is given by the following formula:

Proposition 2.8. *The CML Equation shows the relation between the expected return and the risk of return for any efficient portfolio of assets P. More formally, this is given as:*

$$\mathbb{E}(\boldsymbol{r}^{\phi^P}) = r_f + \frac{(\mathbb{E}(\boldsymbol{r}^{\phi^M}) - r_f)}{\sqrt{\mathbb{V}(\boldsymbol{r}^{\phi^M})}} \sqrt{\mathbb{V}(\boldsymbol{r}^{\phi^P})} \tag{2.6}$$

Proof. This is a direct consequence of the One Fund Theorem as applied to the Market Portfolio. □

The expected return of any efficient portfolio belonging to the CML is a linear function of its standard deviation. The slope factor: $\frac{\mathbb{E}(r^{\phi^M}) - r_f}{\sqrt{\mathbb{V}(r^{\phi^M})}}$ is called the market price of risk. To simplify notations in the following, instead of $\boldsymbol{r}^{\phi^M}$, the Market portfolio returns are denoted by $\boldsymbol{r}^M$.

2.3.2 The Security Market Line (SML)

The CAPM goes further, and signifies how the expected return of a single asset should relate to its individual risk. This gives a precise pricing formula for any risky asset in the CAPM framework. The price of risk is commonly referred to as Beta, as shown in a first sub-section. The Security Market Line Equation and pricing formula for any risky asset are given in a subsequent section.

2.3.2.1 The CAPM Betas

Definition 2.4 (CAPM Beta)**.** *The Beta of a risky asset i is denoted by $\beta^{i,M}$. It represents the price of risk of the asset i, and is effectively the covariance of the returns of the asset i and the returns of the Market portfolio r^M (in the empirical example the Eurostoxx index), adjusted by the variance of the Market portfolio returns. The Beta can be represented more formally as:*

$$\beta^{i,M} = \frac{\mathbb{COV}(\boldsymbol{r}^i, \boldsymbol{r}^M)}{\mathbb{V}(\boldsymbol{r}^M)}$$

The Beta of an asset i represents the relative contribution of the asset return i to the variance of the market return $\boldsymbol{r}^M$.

2.3.2.2 The SML Equation

The main result of the CAPM is that Sharpe extends the CML to a general relationship between any single risky asset expected return (that is not necessarily efficient and therefore does not lay automatically on the CML) and the Market portfolio return:

Proposition 2.9 (Security Market Line Equation)**.**
The expected return of any asset i is given by:

$$\mathbb{E}(r^i) = r_f + \beta^{i,M}[\mathbb{E}(r^i) - r_f]$$

Proof.

The portfolio a constituted by the risky asset i in proportion a, and the market portfolio in proportion (1-a) will be considered. The first two moments of this portfolio are expressed as:

$$\begin{cases} \mathbb{E}(r^a) = a\mathbb{E}(r^i) + (1-a)\mathbb{E}(r^M) \\ \mathbb{V}(r^a) = a^2\mathbb{V}(r^i) + (1-a)^2\mathbb{V}(r^M) + 2a(1-a)\mathbb{COV}(r^i, r^M) \end{cases}$$

Thus, it is found that:

$$\begin{cases} \frac{d\mathbb{E}(r^a)}{da} = \mathbb{E}(r^i) - \mathbb{E}(r^M) \\ \frac{d\mathbb{V}(r^a)}{da} = 2[a\mathbb{V}(r^i) + (a-1)\mathbb{V}(r^M) + (1-2a)\mathbb{COV}(r^i, r^M)] \end{cases}$$

When $a = 0$ is taken:

$$\left.\frac{d\sqrt{\mathbb{V}(r^a)}}{da}\right|_{a=0} = \frac{1}{2\sqrt{\mathbb{V}(r^a)}}\frac{d\mathbb{V}(r^a)}{da}\frac{\mathbb{COV}(r^i, r^M) - \mathbb{V}(r^M)}{\sqrt{\mathbb{V}(r^M)}}$$

And then:

$$\frac{d\mathbb{E}(r^a)}{d\sqrt{\mathbb{V}(r^a)}}\Bigg|_{a=0} = \frac{(\mathbb{E}(r^i) - \mathbb{E}(r^M))\sqrt{\mathbb{V}(r^M)}}{\mathbb{COV}(r^i, r^M) - \mathbb{V}(r^M)}$$

Using equality with the CML slope in $a = 0$, the result is finally:

$$\tfrac{(\mathbb{E}(r^i)-\mathbb{E}(r^M))\sqrt{\mathbb{V}(r^M)}}{\mathbb{COV}(r^i,r^M)-\mathbb{V}(r^M)} = \tfrac{\mathbb{E}(r^i)-r_f}{\sqrt{\mathbb{V}(r^M)}}$$

$\square$

Under the equilibrium conditions assumed by the CAPM, any asset (including efficient assets) should fall on the Security Market Line. Therefore, under the CAPM assumptions, the Security Market Line is a universal pricing line.

2.3.3 The CAPM

The CAPM proposes a model for asset price returns that complies with the SML. In a first sub-section, the CAPM will be formally described, and then the method of building the CAPM portfolio based on the Jensen measure will be presented. The CAPM portfolio will be empirically tested in Chapter 3, among other classical portfolios.

2.3.3.1 The CAPM Equation

The CAPM states that any random asset return r^i can be separated into a systematic component and a residual component:

$$r^i = r_f + \beta^i(r^M - r_f) + \epsilon^i$$

If no assumption is made on the distribution of ϵ^i, this equation is arbitrary. To be coherent with the SML (taking the expected value on both sides of the equation), the CAPM assumes that the idiosyncratic risk is uncorrelated with the market risk, and its expected value is zero. The CAPM theorem can now be stated precisely:

Theorem 2.3 (CAPM). *Any risky asset return can be expressed with respect to its price for risk:*

$$r^i = r_f + \beta^i (r^M - r_f) + \epsilon^i$$

-where ϵ^i is such that:

$$\forall i,\ \mathbb{E}(\epsilon^i) = 0 \textit{ and } \forall i,\ \mathbb{COV}(\epsilon^i, r^M) = 0$$

This leads to the following:

$$\forall i,\ \mathbb{E}(r^i) = \beta^i \mathbb{E}(r^M) \text{ and } \forall i,\ \mathbb{V}(r^i) = (\beta^i)^2 \mathbb{V}(r^M) + \mathbb{V}(\epsilon^i)$$

To summarise the workings undertaken in the chapter so far, the CAPM decomposes any risky asset excess return in a systematic component, defined as $\beta^i(r^M - r_f)$, and an idiosyncratic component, which is defined as ϵ^i. The first equation states that the expected return of a risky asset is the product of the risky asset Beta and the expected return of the market. This relation defines the "Security Market Line". Figure (2.3) plots the Security Market Line for the Eurostoxx 600 constituents (the Betas are estimated over the whole period January 2000 to April 2010). It can be observed that the CAPM relationship is not well respected empirically (the empirical observations presented are well dispersed around the theoretical SML...). The second equation states that any asset risk can be decomposed into a systematic risk (the market risk adjusted by the asset Beta) and a specific risk. Under the CAPM assumption, although the market risk is inescapable, the idiosyncratic risk is escapable through diversification. As such, the idiosyncratic risk is self-imposed by the investor as a trade off between return and risk (i.e. the investor's risk aversion).
Figure (2.3) plots the SML that has been deduced from the empirical estimation of European stock returns, where the Market portfolio is the Eurostoxx 600. When comparing the empirical asset returns with their empirical Betas, The points are rather scattered around the Security Market Line: empirically the strict CAPM equilibrium of asset returns and their Beta times the market return does not hold.

2.3.3.2 The Jensen Alpha

Jensen (1969) provides a performance analysis of investment funds benchmarked to the CAPM. Jensen defines the Alpha as a deviation from the equilibrium induced by the CAPM. The Alpha is formally defined as:

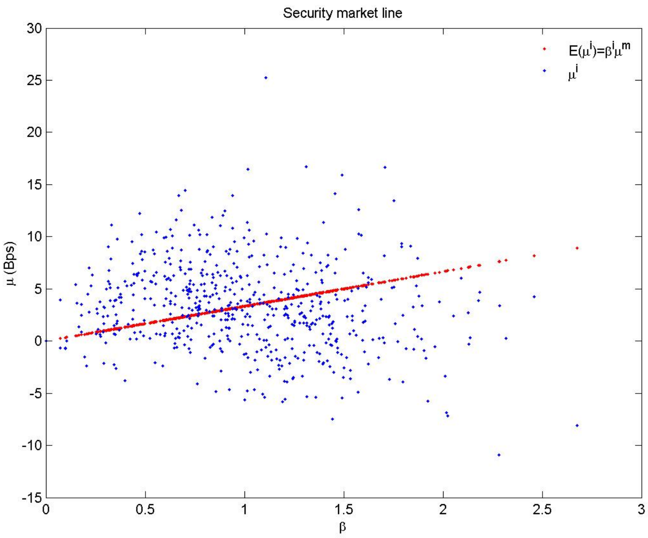


Figure 2.3: Security Market Line

Definition 2.5 (Jensen Alpha). *The Alpha of an asset i is denoted by α^i, and defined as:*

$$\alpha^i \equiv r_f + \beta^i(\overline{r^M} - r_f) - \overline{r^i}$$

-where $\overline{r^i}$ stands as the empirical mean of the asset i returns, and $\overline{r^M}$ as the empirical mean of the Market portfolio.

According to the CAPM, the value of α^i should be zero. Hence, it measures how much the observed performance of the asset i (i.e. its empirical mean) has deviated from its theoretical value. If $\alpha^i > 0$, the return of stock i is below its equilibrium CAPM value, and if the CAPM holds, the return of stock i will be expected to increase. If $\alpha^i < 0$, the return of stock i is above its equilibrium CAPM value, and the return of stock will be expected to decrease. The CAPM portfolio allocation $\boldsymbol{\phi}^{CAPM}$ is constructed as:

$$\forall i,\ \phi^{CAPM,i} \equiv \frac{\alpha^i}{\sum_{j=1}^{N} |\alpha^j|}$$

-which corresponds to a relative Alpha weight, normalised by the sum of absolute Alphas, and additionally, that correspond to the scale of the portfolio (how many Alpha units in absolute terms the investors need to invest in the portfolio). The performances of the CAPM, Minimum-Variance and Maximum-Sharpe portfolios will be tested in the next chapter.

The CAPM is a powerful model which provides a general equilibrium theory for asset prices. Under the CAPM framework, all the investors have the same expectations and differ only in their tolerance for risk. Of course, the very strong assumptions implied by the CAPM make it subject to qualification, and caution must be applied. Nonetheless, the CAPM represents a powerful benchmark that is widely used in practice: investors assume real asset returns should respect the hypothesis and theoretical framework of the CAPM (i.e. they should be proportional to their Beta times the market return). The main drawback of the CAPM can be attributed to the difficulty involved in estimating the β^i in practice. This is all the more so when investors have access to varied information, making a divergence on their estimation of the β^i more likely.
If the Market can effectively explain a good proportion of the asset returns, the strong assumptions of the CAPM are a deterrent. As such, a number of alternative models have been proposed, including the Asset Pricing Theory, which is the most widely known alternative to the CAPM.

2.4 The Asset Pricing Theory (APT), Ross (1976)

The Asset Pricing Theory is reliant upon the factor model framework and forms an alternative theory of asset pricing based on the no-arbitrage principle. The theory omits the assumption that investors rely on the mean-variance paradigm. Thus, in this sense, the APT is more general than the CAPM, which is limited by the reliance upon both the mean-variance framework, and a strong version of equilibrium that assumes that all investors use the same framework.
In the first sub-section, the APT framework is described in detail. Then the APT is presented more formally, before finally providing the methodology for portfolios constructed using factor models such as the APT.

2.4.1 The APT Framework

Ross (1976) considers a multi-factor model, which is defined by the use of more than one factor to take into account and represent asset returns. The APT is a model of expected returns. As such, it is an equilibrium model that is based upon no-arbitrage, and excludes investor preference. Ross argues that if equilibrium prices offer no arbitrage opportunities over a static portfolio of the assets, then the expected returns on the assets are approximately linearly-related to the factor loadings. The APT, unlike the CAPM, is not a consensus model as it depends on the selection of factors made by each investor.
The APT model can be formalised through the following equation:

$$r^i = \mathbb{E}(r^i) + \sum_{k=1}^{K} f^k \beta^{k,i} + \epsilon^i \tag{2.8}$$

-where K is the number of factors selected [1], f^k stands for the centred return of the factor k (i.e. the return of the factor k minus its mean) and $\beta^{k,i}$ is the loading of the asset i on the factor k.
The systematic component of the asset return is therefore defined as the factors exposure $\sum_{k=1}^{K} f^k \beta^{k,i}$, and the idiosyncratic risk of the asset return i is defined as ϵ^i. At this point, if no assumption is made on ϵ^i, the equation is arbitrary. The APT relies upon an assumption concerning the distribution of the idiosyncratic risk ϵ^i, as it assumes that it is uncorrelated with every one of the K factors, and its expected value is zero. Also, the factors can be assumed to be uncorrelated (so that $dim(f^1, ..., f^k) = K$)[2]. More formally, the APT assumes the following conditions are satisfied:

APT Assumptions (A) :

$$\begin{cases} (a)\, \forall i,\, \mathbb{E}(\epsilon^i) = 0 \text{ and } \forall j \neq i,\, \mathbb{COV}(\epsilon^i, \epsilon^j) = 0 \\ (b)\, \forall i,\, \forall k,\, \mathbb{COV}(\epsilon^i, f^k) = 0 \\ (c)\, \forall k,\, \forall k' \neq k,\, \mathbb{COV}(f^k, f^{k'}) = 0 \\ (d) K << N \end{cases}$$

[1]Note that $K < N - 2$ is a necessary condition for the APT to hold true.
[2]Note that this is not a key assumption of the APT.

The APT theorem can now be formally enunciated.

2.4.2 APT Theorem

Theorem 2.4 (APT). *If considering the following Equation (2.8):*

$$r^i = \mathbb{E}(r^i) + \sum_{k=1}^{K} f^k \beta^{k,i} + \epsilon^i$$

-under the hypothesis (A), the following result holds true:

$$\forall i,\ \mathbb{E}(r^i) = r_f + \sum_{k=1}^{K} \mathbb{E}(f^k - r_f)\beta^{k,i} \tag{2.9}$$

-which leads to:

$$\forall i,\ \mathbb{V}(r^i) = \sum_{k=1}^{K} \mathbb{V}(f^k)(\beta^{k,i})^2 + \mathbb{V}(\epsilon^i) \tag{2.10}$$

The proof of Equation (2.9) is based on a no-arbitrage argument.

Step 1: Firstly, a portfolio ϕ with a null initial value is constructed:

$$\sum_{i=1}^{N} \phi^i = 0$$

According to Equation (2.8), the return of this portfolio is defined as:

$$\sum_{i=1}^{N} \phi^i r^i = \sum_{i=1}^{N} \phi^i \mathbb{E}(r^i) + \sum_{i=1}^{N} \phi^i \sum_{k=1}^{K} f^k \beta^{k,i} + \sum_{i=1}^{N} \phi^i \epsilon^i$$

Step 2: Secondly, a particular value of $\boldsymbol{\phi}$ is chosen. The non-risky portfolio is considered. The exposure of this portfolio to the different K factors is eliminated:

$$\forall k,\ \sum_{i=1}^{N} \phi^i \beta^{k,i} = 0$$

Because according to assumption (b), the ϵ^i are independent of the factors, it is only necessary to eliminate the Beta exposure to eliminate the factors risk. When N is large enough, the residual risk can eliminated. Indeed, thanks to the diversification effect, when N is large the following occurs:

$$\sum_{i=1}^{N} \epsilon^i = 0$$

Step 3:
From Step 1 and 2, because the initial value of the portfolio is null and the portfolio is non-risky, the return of this portfolio should also be null under the no arbitrage principle; therefore:

$$\sum_{i=1}^{N} \phi^i \mu'^i = 0$$

where $\boldsymbol{\mu}$ is the vector of risky asset mean returns (i.e. $\forall i,\ \mu^i \equiv \mathbb{E}(r^i)$).
It is deduced that $\boldsymbol{\phi}$ belongs to R^T, which is the orthogonal space of R generated by $\boldsymbol{\mu}$.
From Step 2, it is also deduced that $\boldsymbol{\phi}$ also belongs to P^T, the orthogonal space of P, which is the space generated by the vectors $(\mathbf{1}, \boldsymbol{\beta}^1, ..., \boldsymbol{\beta}^K)$.

Step 4:
The non-arbitrage argument implies that: $P^T \subset R^T$ and therefore $R \subset P$. Therefore, there exists a $K+1$ vector $\boldsymbol{\lambda}$ (this vector is unique if the factors are independent, i.e. $dim(P) = K + 1$[1]) such that:

$$\boldsymbol{\mu} = \lambda^0 \mathbf{1} + \sum_{k=1}^{K} \lambda^k \boldsymbol{\beta}^k$$

-or for any asset i the following is given:

$$\mu^i = \lambda^0 + \sum_{k=1}^{K} \lambda^k \boldsymbol{\beta}^{k,i}$$

[1]Note that if $dim(P) > N$ the problem has no solution.

In the particular case of the risk-free asset, because $\forall k$, $\beta^{k,f} = 0$ (the risk-free asset is independent of all the risky assets and in particular it is independent of all the APT factors, i.e. $\mathbb{COV}(r_f, f^k) = 0$), it can be deduced that:

$$\lambda^0 = r_f$$

Applying the equality to any factor k, and given that $\forall k' \neq k$, $\beta^{k,k'} = 0$ (assumption (c)) and $\beta^{k,k} = 1$, the following is given:

$$\mathbb{E}(f^k - r_f) = \lambda^k$$

Finally, the following is obtained:

$$\mathbb{E}(r^i) = \mu^i = r_f + \sum_{k=1}^{K} \mathbb{E}(f^k - r_f)\boldsymbol{\beta}^{k,i}$$

The proof of Equation (2.10) is a direct result of the conditions (a), (b) and (c). In many cases it is assumed that the factor relationships are more stable than the stock relationships, and are therefore more predictable. The APT is not a general equilibrium model as it is subjective in the factors selection process; in contrast to the CAPM, it is investor specific. The APT does not provide an investment rule as the One Fund Theorem provided by the CAPM. However, it is less constraining than the CAPM because it is not constrained by quadratic preferences and there is no assumption made on the factors return distribution. Note that the CAPM is a particular case of the APT where a unique factor is considered: the Market factor. The main problem connected with the APT is the identification of the factors, as will be discussed in the next chapter.

2.4.3 Construction of a factor model portfolio

In the next chapter, several different factor models based on the APT principle are tested. Given the choice of a set of factors $(f^1, ..., f^K)$, the factor model Alpha is defined as:

$$\alpha^{APT,i} \equiv r_f + \sum_{k=1}^{K} \mathbb{E}(f^k - r_f)\beta^{k,i} - \overline{r^i}$$

According to the APT, the value of $\alpha^{APT,i}$ should be zero. Hence, it measures how much the observed performance of the asset i (i.e. its empirical mean) has deviated from its theoretical value. The APT portfolio allocation $\boldsymbol{\phi}^{APT}$ is constructed:

$$\forall i,\ \phi^{APT,i} \equiv \frac{\alpha^{APT,i}}{\sum_{j=1}^{N} |\alpha^{APT,j}|}$$

Different types of factor model portfolios are tested in the next chapter. In particular, a form of the famous empirical application of the APT model developed by Fama & French (1992) will be tested. In addition, tests will be performed on factor models that are based on pure statistical models such as Principal Component Analysis or Independent Component Analysis.

In order to evaluate and compare different models, numerous performance measures have been proposed in the literature. As presented above, the Sharpe ratio is a measure linked to the Efficient Frontier. Other measures are also considered in order to compare the performance of the different models throughout this PhD thesis.

2.5 Performance measures

The excess return of any risky asset i, will be denoted by $\underline{r}^i$:

$$\underline{r}^i \equiv r^i - r_f$$

The Sharpe is by far the most widely known measure of performance. Many improvements have been proposed in the literature. One such example is the Sortino which only considers the standard deviation of the losses (or negative returns) as risk. The Sharpe and Sortino represent an expected average return per unit of risk, expressed as:

- **Sharpe Ratio (see Sharpe (1994)):** represents the ratio of the mean excess return of an investment over the standard deviation of its returns.

This is the most famous risk measure used in finance. It effectively indicates the average expected return of an investment per unit of risk.

$$Sharpe \equiv \frac{\mathbb{E}(\underline{r}^i)}{\sigma(\underline{r}^i)}$$

- **Sortino Ratio (see Sortino & Price (1994)):** represents the ratio of the mean excess return of an investment over the standard deviation of its negative returns. It is an adjusted Sharpe ratio: the idea here is that only downside risk matters to the investor (i.e. the variations of the negative returns). Therefore, the standard deviation of the returns is replaced by the standard deviation of the truncated distribution of the returns (only taking into account negative returns).

$$Sortino \equiv \frac{\mathbb{E}(\underline{r}^i)}{\sigma(\underline{r}^i/\underline{r}^i < 0)}$$

The main problem with the precedent measures, is the implicit assumption that the asset returns are either normally distributed, or that only the first two moments of their distribution are significant (i.e. the investor needs to fully believe in the mean-variance paradigm), however it has been shown in many empirical studies that, contrary to theoretical assumptions, real asset returns are not normally distributed (see for instance Longin & Solnik (2001)). The following two measures are more empirical and do not form assumptions concerning the distribution of the returns. The Win/Lose ratio (or hit ratio), and the similar Gain/Loss ratio are widely used in the hedge fund industry, as they are easily interpretable and model-independent. (There is no implicit assumption made on the return distribution, as is the case for the other measures that consider the returns normally distributed or at least solely defined by its first two moments). These measures represent a percentage of comparison at 50%: if above, the number of positive returns or gain surpasses the number of negative returns or losses. In the following, the empirical cumulative distribution of a given random variable x is denoted by $F(x)$.

- **Gain Loss Ratio (see Bernardo & Ledoit (2001)):** it is the ratio of total positive returns over total negative returns. It gives an idea of the proportion of gains against losses when investing in a strategy.

$$GainLoss \equiv \frac{\int_{\underline{r}^i>0} \underline{r}^i dF(\underline{r}^i)}{\int_{\underline{r}^i} \underline{r}^i dF(\underline{r}^i)}$$

- **Winner Loser Ratio (WinLose):** similar to the Gain Loss ratio, it is the ratio of the number of total positive returns over the number of total negative returns. Also called the "Hit Ratio", it proportionately indicates the number of times an investment is profitable.

$$WinLose \equiv \frac{\int_{\tilde{r}^i>0} dF(\tilde{r}^i)}{\int_{\tilde{r}} dF(\tilde{r}^i)}$$

- **Certain Equivalent Ratio (CER) :** corresponds to the equivalent risk-free return of the index return. This measure can be considered more academic, as practitioners often find it difficult to parametrize their risk aversion.

$$CER \equiv \mathbb{E}(\underline{r}^i) - \lambda\sigma^2(\underline{r}^i)$$

  where λ stands for the investor risk aversion parameter[1].

In the next Chapter, these performance measures will be used to study the empirical performance of a number of commonly used portfolios, including the classical models described above, and a number of other factor model based portfolios.

[1] In the following, $\lambda = 1$ is taken, as in DeMiguel *et al.* (2007)

2.6 Appendix

2.6.1 Efficient Frontier without a risk-free asset

Proposition 2.10 (Minimum Variance Portfolio).
The Minimum Variance portfolio allocation denoted by $\boldsymbol{\phi}^{MN}$ is the solution for the following problem:

$$\begin{cases} \min_{\boldsymbol{\phi}} \frac{1}{2}\boldsymbol{\phi}'\boldsymbol{\Sigma}\boldsymbol{\phi} \\ s.t\ \boldsymbol{\phi}'\mathbf{1} = 1 \end{cases}$$

- The first two moments of the portfolio return $r^{\boldsymbol{\phi}^{MN}}$ are:

$$\begin{cases} \mathbb{E}(r^{\boldsymbol{\phi}^{MN}}) = \frac{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}} = \frac{b}{a} \\ \mathbb{V}(r^{\boldsymbol{\phi}^{MN}}) = \frac{1}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}} = \frac{1}{a} \end{cases}$$

- The Minimum Variance Portfolio allocation is $\boldsymbol{\phi}^{MN} = \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}$.

Proof.
The Lagrangian associated with the above equation is:

$$L(\boldsymbol{\phi}, \lambda) \equiv \frac{1}{2}\boldsymbol{\phi}'\boldsymbol{\Sigma}\boldsymbol{\phi} - \lambda(\boldsymbol{\phi}'\mathbf{1} - 1)$$

- The first order conditions read:

$$\begin{cases} (1)\ \frac{\delta L(\boldsymbol{\phi},\lambda)}{\delta \boldsymbol{\phi}} = 0 \Leftrightarrow \boldsymbol{\phi} = \lambda\boldsymbol{\Sigma}^{-1}\mathbf{1} \\ (2)\ \frac{\delta L(\boldsymbol{\phi},\lambda)}{\delta \lambda} = 0 \Leftrightarrow \mathbf{1}'\boldsymbol{\phi} = 1 \end{cases}$$

By multiplying (1) by $\mathbf{1}'$ and using (2), it is found that $\boldsymbol{\phi}^{MN} = \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}$.
- The value of $\mathbb{E}(\boldsymbol{\phi}^{MN})$ and $\mathbb{V}(\boldsymbol{\phi}^{MN})$ is deduced:

$$\begin{cases} \mathbb{E}(\boldsymbol{\phi}^{MN}) = \boldsymbol{\mu}'\boldsymbol{\phi}^{MN} = \frac{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}} = \frac{b}{a} \\ \mathbb{V}(\boldsymbol{\phi}^{MN}) = \boldsymbol{\phi}^{MN\prime}\Sigma\boldsymbol{\phi}^{MN} = \frac{\boldsymbol{\phi}^{MN\prime}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}} = \frac{1}{a} \end{cases}$$

□

Proposition 2.11 (Maximum Sharpe Ratio Portfolio).
If the Maximum Sharpe ratio portfolio allocation with expected return μ^{MS} is denoted by $\boldsymbol{\phi}^{MS}$, then $\boldsymbol{\phi}^{MS}$ is the solution for the following problem:

$$\begin{cases} \min_{\boldsymbol{\phi}} \frac{1}{2}\boldsymbol{\phi}'\boldsymbol{\Sigma}\boldsymbol{\phi} \\ s.t\ \boldsymbol{\phi}'\mu = \mu^{MS} \end{cases}$$

- The first two moments of the portfolio return $\boldsymbol{\phi}^{MS}$ are:

$$\begin{cases} \mathbb{E}(r^{\boldsymbol{\phi}^{MS}}) = \mu^{MS} \\ \mathbb{V}(r^{\boldsymbol{\phi}^{MS}}) = \frac{(\mu^{MS})^2}{c} \end{cases}$$

- In addition, the Maximum Sharpe Portfolio allocation is $\boldsymbol{\phi}^{MS} = \frac{\mu^{MS}}{c}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$.

Proof.
The Lagrangian associated with the problem is:
$L(\phi, \lambda) \equiv \frac{1}{2}\boldsymbol{\phi}'\boldsymbol{\Sigma}\boldsymbol{\phi} - \lambda(\boldsymbol{\phi}'\boldsymbol{\mu} - \mu^{MS})$

- The first order conditions read:

$$\begin{cases} (1)\frac{\delta L(\boldsymbol{\phi},\lambda)}{\delta\boldsymbol{\phi}} = 0 \Leftrightarrow \boldsymbol{\phi} = \lambda\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \\ (2)\frac{\delta L(\boldsymbol{\phi},\lambda)}{\delta\lambda} = 0 \Leftrightarrow \boldsymbol{\mu}'\boldsymbol{\phi} = \mu^P \end{cases}$$

By multiplying (1) by $\boldsymbol{\mu}'$ and using (2), it is found that $\boldsymbol{\phi}^P = \frac{\mu^P}{c}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$. Therefore, $\mathbb{V}(\boldsymbol{\phi}^P) = \boldsymbol{\phi}^{P\prime}\Sigma\boldsymbol{\phi}^P = \frac{(\mu^P)^2}{c}$

□

Proposition 2.12 (Fully Invested Maximum Sharpe Portfolio)**.**
If the fully invested Maximum Sharpe portfolio allocation is denoted by $\boldsymbol{\phi}^{MS}$*, then the mean* μ^{MS} *and standard deviation* σ^{MS} *of this portfolio must respect the equations (2.3) and (2.2):*

$$\begin{cases} (1)\,\sigma^{MS} = \sqrt{\frac{a}{d}(\mu^{MS} - \frac{b}{a})^2 + \frac{1}{a}} \\ (2)\,\sigma^{MS} = \frac{1}{\sqrt{c}}\mu^{MS} \end{cases}$$

- The first two moments of the portfolio return $\boldsymbol{\phi}^{MS}$ *are:*

$$\begin{cases} \mathbb{E}(r^{\boldsymbol{\phi}^{MS}}) = \frac{c}{b} \\ \mathbb{V}(r^{\boldsymbol{\phi}^{MS}}) = \frac{c}{b^2} \end{cases}$$

- The fully invested Maximum Sharpe Portfolio allocation is: $\phi^{MS} = \frac{\Sigma^{-1}\boldsymbol{\mu}}{\mu'\Sigma^{-1}\mathbf{1}}$.

It is possible to derive (1) and (2) with respect to μ^{MS}, and those two derivatives must be equal:

$$\begin{cases} (1)\,\frac{\delta\sigma^{MS}}{\delta\mu^{MS}} = \frac{ad(\mu^{MS} - \frac{b}{a})}{\sigma^{MS}} \\ (2)\,\frac{\delta\sigma^{MS}}{\delta\mu^{MS}} = \frac{1}{\sqrt{c}} \end{cases}$$

By equaling (1) and (2), $\mu^{MS} = \frac{c}{b}$ and $\sigma^{MS} = \frac{\sqrt{c}}{b}$ are found. The Maximum Sharpe portfolio asset allocation is such that: $\boldsymbol{\mu}'\phi^{MS} = \frac{c}{b}$ and therefore $\phi^{MS} = \frac{\Sigma^{-1}\boldsymbol{\mu}}{\mu'\Sigma^{-1}\mathbf{1}}$.

Proposition 2.13 (Efficient Frontier Equation without a risk free asset).
If the variance minimisation problem is considered:

$$\begin{cases} \min_{\boldsymbol{\phi}} \mathbb{V}(r^{\boldsymbol{\phi}}) \\ s.t\ \mathbb{E}(r^{\boldsymbol{\phi}}) = \mu^P \\ and\ \boldsymbol{\phi}'\mathbf{1} = 1 \end{cases}$$

-it is found that:

$V(r^{\boldsymbol{\phi}^P}) = \frac{a}{d}[\mathbb{E}(R^{\boldsymbol{\phi}^P}) - \frac{b}{a}]^2 + \frac{1}{a}(2.11)$

Proof.

The Lagrangian associated with the problem is:

$$L(\phi, \lambda_1, \lambda_2) \equiv \frac{1}{2}\boldsymbol{\phi}'\boldsymbol{\Sigma}\boldsymbol{\phi} - \lambda_1(\boldsymbol{\phi}'1 - 1) - \lambda_2(\boldsymbol{\phi}'\boldsymbol{\mu} - \mu^P)$$

-the first order conditions read:

$$\begin{cases} (1)\ \frac{\delta L(\boldsymbol{\phi},\lambda_1,\lambda_2)}{\delta \phi} = 0 \Leftrightarrow \boldsymbol{\phi} = \lambda_1\boldsymbol{\Sigma}^{-1}\mathbf{1} + \lambda_2\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \\ (2)\ \frac{\delta L(\boldsymbol{\phi},\lambda_1,\lambda_2)}{\delta \lambda_1} = 0 \Leftrightarrow \mathbf{1}'\boldsymbol{\phi} = 1 \\ (3)\ \frac{\delta L(\boldsymbol{\phi},\lambda_1,\lambda_2)}{\delta \lambda_2} = 0 \Leftrightarrow \boldsymbol{\mu}'\boldsymbol{\phi} = \mu^P \end{cases}$$

By multiplying (1) by $\boldsymbol{\mu}'$ and $\mathbf{1}'$ respectively and using (2) and (3), it is found that:

$$\begin{cases} (4)\ 1 = \lambda_1 a + \lambda_2 b \\ (5)\ \mu^P = \lambda_1 b + \lambda_2 c \end{cases}$$

Then, by multiplying (4) by b and (5) by a and (4) by c and (5) by b, it is found that:

$$\begin{cases} \lambda_1 = \frac{c - b\mu^P}{d} \\ \lambda_2 = \frac{a\mu^P - b}{d} \end{cases}$$

It is deduced that $\mathbb{V}(r^{\boldsymbol{\phi}^P}) = \boldsymbol{\phi}^{P\prime}\Sigma\boldsymbol{\phi}^P = \lambda_1\boldsymbol{\phi}^{P\prime}\mathbf{1} + \lambda_2\boldsymbol{\phi}^{P\prime}\boldsymbol{\mu}$.
Therefore, $\mathbb{V}(r^{\boldsymbol{\phi}^P}) = \frac{a(\mu^P)^2 - 2b\mu^P + c}{d}$ is given. It is finally found:

$$\mathbb{V}(\boldsymbol{\phi}^P) = \frac{a}{d}[\mathbb{E}(\boldsymbol{\phi}^P) - \frac{b}{a}]^2 + \frac{1}{a}$$

□

2.6.2 Efficient Frontier with a risk-free asset

Proposition 2.14 (Efficient Frontier Equation with a risk-free asset).

It is found that in a given case when a risk-free asset exists, the Markowitz efficient frontier is a straight line, with the equation:

$$\mu^P = \sqrt{e}\sigma^P + r_f \tag{2.12}$$

with $e \equiv (\boldsymbol{\mu} - r_f\mathbf{1})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1})$.

Proof.
The Lagrangian associated with the problem is:
$L(\phi,\lambda) \equiv \frac{1}{2}\boldsymbol{\phi}'\boldsymbol{\Sigma}\boldsymbol{\phi} - \lambda(\boldsymbol{\phi}'\mu + (1-\boldsymbol{\phi}'\mathbf{1})r_f - \mu^P)$

- The first order conditions read:
$\begin{cases} (1)\ \frac{\delta L(\phi,\lambda)}{\delta\phi} = 0 \Leftrightarrow \boldsymbol{\phi}' = \lambda(\mu - r_f\mathbf{1})'\boldsymbol{\Sigma}^{-1} \\ (2)\ \frac{L(\phi,\lambda)}{\delta\lambda} = 0 \Leftrightarrow \boldsymbol{\phi}'(\mu - r_f\mathbf{1}) = \mu^P - r_f \end{cases}$
By multiplying (1) by $(\mu - \mathbf{1}r_f)$ and using (2), it is found that $\lambda = \frac{\mu^P - r_f}{(\mu - r_f\mathbf{1})'\boldsymbol{\Sigma}^{-1}(\mu - r_f\mathbf{1})}$.
Therefore: $\mathbb{V}(\boldsymbol{\phi}^P) = (\sigma^P)^2 = \boldsymbol{\phi}'\boldsymbol{\Sigma}\boldsymbol{\phi} = \lambda(\mu - r_f\mathbf{1})'\boldsymbol{\phi} = \lambda(\mu^P - r_f) = \frac{(\mu^P - r_f)^2}{e}$,
where $e \equiv (\mu - r_f\mathbf{1})'\boldsymbol{\Sigma}^{-1}(\mu - r_f\mathbf{1})$.

$\square$

Proposition 2.15 (Tangential Portfolio)**.**
If $\boldsymbol{\phi}^T$ *denotes the Tangential portfolio allocation, then the mean* μ^T *and standard deviations* σ^T *of this portfolio are:*

$$\begin{cases} \mathbb{E}(r^{\boldsymbol{\phi}^T}) = \frac{e}{\sqrt{ae-d}} + r_f \\ \mathbb{V}(r^{\boldsymbol{\phi}^T}) = \frac{e}{ae-d} \end{cases}$$

where $e = (\boldsymbol{\mu} - r_f\mathbf{1})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1})$.

Proof.

$\begin{cases} (1)\ \mathbb{E}(r^{\boldsymbol{\phi}^T}) = \sqrt{\frac{d}{a}(\mathbb{V}(r^{\boldsymbol{\phi}^T}) - \frac{1}{a})} + \frac{b}{a} \\ (2)\ \mathbb{E}(r^{\boldsymbol{\phi}^T}) = \sqrt{e\mathbb{V}(r^{\boldsymbol{\phi}^T})} + r_f \end{cases}$

It is possible to derive (1) and (2) with respect to $\mathbb{V}(\boldsymbol{\phi}^T)$ as the derivative of the two equations must be equal at the tangential point:
$\begin{cases} (1)\ \frac{\delta\mathbb{E}(r^{\boldsymbol{\phi}^T})}{\mathbb{V}(\boldsymbol{\phi}^T)} = \sqrt{(e)} \\ (2)\ \frac{\delta\mathbb{E}(r^{\boldsymbol{\phi}^T})}{\mathbb{V}(\boldsymbol{\phi}^T)} = \sqrt{\frac{d\mathbb{V}(\boldsymbol{\phi}^T)}{a\mathbb{V}(\boldsymbol{\phi}^T)-1}} \end{cases}$

It is finally found that the coordinates of the Tangential portfolio are:
$$\begin{cases} \mathbb{E}(r^{\boldsymbol{\phi}^T}) = \frac{r_f d - be}{d - ae} \\ \mathbb{V}(r^{\boldsymbol{\phi}^T}) = \frac{(r_f d - be)^2}{e(d - ae)^2} \end{cases}$$
-where $e = (\boldsymbol{\mu} - r_f \mathbf{1})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - r_f \mathbf{1})$.

$\square$

Chapter 3

Factor Model Performances: Empirical Evidence

"*A central problem in finance (and especially portfolio management) has been that of evaluating the performance of portfolios of risky investments.*"

Michael C. Jensen, The Performance of Mutual Funds in the Period 1945-1964, Journal of Finance, 1967.

This chapter will provide a general and empirical overview of the main factor models used by financial investors to evaluate asset returns. Theoretical models are used empirically by financial investors: the real asset returns are assumed to respect the hypothesis and theoretical framework of those models and the discrepancy observed between real asset returns and the returns theoretically expected by factor models are taken as practical investment opportunities for financial investors (any deviance observed between real returns and theoretical returns is considered as an "error" that will be corrected). A number of models commonly used by practitioners will be tested, including: the Capital Asset Pricing Model; factor models based on the Asset Pricing Theory, such as those based on a selection of general external factors, and a Fama and French type of model; and three purely statistical models that are based on Principal Component Analysis, Independent Component Analysis and Cluster Analysis, respectively. The classical methodology introduced by Jensen (1969) will be applied to evaluate the performance of, and compare the aforementioned standard models. The significant fact to note is that the performance of the different models varies greatly over time; thus, investors are unable to exclusively rely on any single one of them.

The different factor models will be evaluated with European stock returns data collected from 2000 to 2010. In contrast to numerous reference papers focusing on portfolio analysis (for instance Fama & French (1998), who test high price to book ratio portfolios against low price to book ratio portfolios, or Jensen (1969), who test mutual funds performance) a single stock approach is adopted here, rather than a portfolio approach. Our aim is not to evaluate precisely some stand alone portfolio strategies, but to compare different factor models using some common criteria, such as the Jensen Alpha (Jensen (1969)) described in Chapter 2, Section 3. Also, daily data for a large number of European stocks are retrieved, allowing better granularity in the data than the two monthly studies cited above. In particular, daily fundamental data are used so that a pure fundamental factor model is built where stock returns are directly explained by fundamental factors.

The chapter is structured as follows: a first section will outline general remarks on factor models and how they are used by financial investors. In Section 2, the data-set used for the analysis and the cleaning process conducted on stock returns will be presented. In Section 3, the CAPM portfolio (as presented in Chapter 2) will be studied, and a description put forward of the Jensen statistics - first used by Jensen (1969) to test the CAPM - that will be applied in this research to test the different factor models considered in this chapter. In addition, several factor model portfolios based on the Asset Pricing Theory will also be studied. In Section 4, an External Factor Model (EFM) will be studied that considers three exogenous factors. In Section 5, a Fundamental Factor Model (FFM) will be presented (largely inspired by the model developed by Fama & French (1992)), that considers three endogenous fundamental factors. Finally, in Sections 6, 7 and 8, three purely statistical factor models will be studied, namely: the Principal Component Analysis (PCA), the Independent Component Analysis (ICA) and a Cluster Analysis (CA). Section 9 will conclude, furthermore incorporating a performance comparison table for the different models tested. All the models presented in this Chapter can be considered as fairly common. However, the analysis of those models and their empirical comparision is original.

3.1 General remarks on factor models

A conceptual discussion on the construction and usage of factor models by financial investors will be outlined. Initially, the two differing purposes of factor models on financial data will be presented, and a description made of the three different types of factor models an investor is able to consider. Lastly, it will be shown, how, generally speaking, an investor can construct a portfolio strategy based on a given factor model.

3.1.1 Factor models objectives

As first stated in the CAPM (see Chapter 2, Section 3), the stock returns can be decomposed into a systematic, and specific part. The aim of a factor model is to explain the systematic risk of stock returns, whereas, the residuals of the model are identified as the specific risks of the considered stock returns.
Factor models are mainly used to achieve the following two objectives:

- **Risk modelling :** given a portfolio of financial assets: an investor wants to understand their exposures with respect to different risk factors. A factor model allows an investor to diversify and control their risk for selected and identified factors. Thus, the idea is to perform a computation to account for the sensitivities of different portfolio components to some identified major factors. A systematic risk model is used to, as far as is possible, reduce a portfolio exposure against identified factors (or sources of risks) a portfolio manager wants to be protected from.

- **Forecasting or Alpha modelling :** a factor model can also be seen as a predictive tool, giving some indications about future financial asset returns. Any discrepancy observed between the theoretical asset return stemming from the factor model equilibrium and the current observed real return is considered as an anomaly that is assumed to be corrected by market strengths (the hypothesis beeing that the theoretical factor model perfectly represent asset returns). The difference between the return observed r_t^i at time t for the stock i, and the expected return stemming from the equilibrium factor model f, denoted by $f(r_t^i)$, is seen as a trading opportunity commonly denoted by "Alpha" and defined as:

$$\alpha_t^i \equiv f(r_t^i) - r_t^i$$

As discussed in the previous chapter, Jensen (1969) first introduced the notion of Alpha applied to the CAPM. Nowadays, the concept of Alpha is widely used in the financial world, and is in fact applied to all types of equilibrium model. In a sense, a model used for forecasting (or Alpha modelling) purposes is more complex than a risk model, as for risk models, the factors are easily identified. The factors depend on the risk aversion of a given investor that wishes not to be exposed to specific risks. In the case of a predictive model, however, the investor aims at exhaustively identifying all the factors that can explain stock returns; which demonstrably, can be very challenging.

In the remainder of this thesis, the manner in which factor models are used for Alpha modelling purposes will be considered.

3.1.2 Types of factor models

In finance, two main types of factor model are usually considered; these are: exogenous factor models, characterised by factors that are identified before the modelling phase; and endogenous factor models, characterised by factors that are computed during the modelling phase. In this chapter particular cases of the following types of models will be studied:

- **Exogenous general factor models :** the factors are chosen as exogenous explanatory variables which are common for all the stocks considered: a multivariate regression is computed to evaluate the sensitivities of the stock returns toward the exogenous variables. The most well known model in that category is certainly the Capital Asset Pricing Model (CAPM), (developed by Sharpe (1964) and presented in Chapter 2), where the unique exogenous factor is the market. An APT model, based on the selection of three external factors will also be presented; namely: the External Factor Model (EFM).

- **Exogenous individual factor models:** in this case the factors are specific to each stock. A cross-sectional regression is computed in order to evaluate the sensitivity of stock returns toward the factor returns. Those factors are usually specific fundamental factors, such as size factors (market capitalisation) and value factors (book to equity ratio); and as suggested in the three factors model tested by Fama & French (1992), that consider the market factor of the CAPM with two additional fundamental factors (size and value factors) to explain US stock returns. In the following section, a Fundamental Factor Model (FFM) based on three fundamental factors will be tested.

- **Pure statistical endogenous factor models:** in this instance, the factors are obtained through the actual modelling process. These models are mainly based on covariance analysis of the stock returns. Among the range of different statistical methods, the most commonly used is the Principal Component Analysis (PCA), which operates under the assumption that the stock returns are Gaussian. More recently still, the Independent Component Analysis (ICA) and also the Cluster Analysis (CA) have been used. The ICA has the advantage that it assumes non-Gaussian returns, which is more realistic in the context of financial markets. The CA offers a more intuitive approach by allowing stocks to be grouped according to a distance criterion that is based upon the correlation of returns. Those three statistical models (PCA, ICA and CA) will be tested in the following section.

The main difficulty presented by the first two types of models lies in the identification of the exogenous factors, as such: if an important explanatory factor is overlooked, the model will be weak in forecasting stock returns; Yet, these models offer a way to analyse and diversify identified risk in a portfolio. The difficulty associated with the third type of model lies in the interpretation of the statistical factors, and specifically, their replication in order to hedge a portfolio against them. However, these last models are potentially able to better explain stock return variations; therefore, offering a more accurate forecasting tool (provided enough factors are used).

In this chapter, nine different factor models will be tested. The models form a wide selection of methods that are relevant to applications in financial investment. The models tested will consist of: the six models outlined above (CAPM, EFM, FFM, PCA, ICA and CA); three benchmark portfolios, consisting of the naive equally weighted portfolio (EW), (that gives equal weight to all risky assets, and which is a good and stable proxy for the Market portfolio see DeMiguel *et al.* (2007)); the Minimum Variance portfolio (MN) and the Mean Variance portfolio (MV) described in Chapter 2.

3.2 Presentation of the data

The following section will describe the data-set used throughout this PhD thesis; first, describing the raw data; second, the cleaning process implemented in this research and applied to the data; and finally, discussing the results of the computation and cleaning processes.

3.2.1 The raw data

The data-set utilised in this study will consist of daily European stock returns computed from closing prices over the period 3rd of January 2000 to 26th of May 2010. Note that, contrary to classical studies, this study deals with daily-data, and not monthly or yearly data. In addition, the time-period considered is a key feature for some of the selected models; and is particularly true for the exogenous factor models, where the choice of relevant factors is obviously conditioned by the historical context and may vary over time.

Four main European index components are considered:

- **The global composite European index, Eurostoxx 600 denoted as SXXP**: is a broad based capitalisation-weighted European stock index. The base value of the index was 100 as of December 31, 1991; it contains 600 assets.

- **The large caps UK index, Ftse 100 denoted as UKX**: is a capitalisation-weighted index and is limited to the 100 most highly capitalised companies traded on the London Stock Exchange. This index was developed with a base level of 1000 as of January 3, 1984; it contains 102 assets, 97 of which are in the Eurostoxx 600.

- **The main French index, Cac 40 denoted as CAC**: is a narrow-based, modified capitalisation-weighted index, consisting of 40 companies listed on the Paris Stock Exchange. The index was developed with a base level of 1,000 as of December 31, 1987; it contains 40 assets, 38 of which are in the Eurostoxx 600.

- **The main German index, Dax 30 denoted DAX** : is a total-return index, consisting of a selection of 30 German stocks traded on the Frankfurt Stock Exchange. The index has a base value of 1,000 as of December 31, 1987; it contains 30 assets, all of which are in the Eurostoxx 600.

The Eurostoxx 600 contains six-hundred elements: all but seven of the stocks of the three other indexes[1].
Closing prices by local currency are collected from January 3rd 2000 to May 26th 2010. All European markets close approximately at the same: GMT time: 4.30pm[2]. Closing-prices are considered to be the official prices that are printed by the exchanges at the end of each trading-day. In most cases, closing-prices represent the last traded price for each stock (or the closing auction price). $T = 2172$

[1]Note: the constituents of the indexes are considered as of the 26th of May 2010. Importantly, there is a survival bias in the set of stocks considered (it can be argued that the constituents considered in the indexes at the end of the period have probably performed better than the constituents of those same indexes at the beginning of the period). However, as the aim of this study is to compare factor models, rather than to accurately estimate the performance of a given index, the said bias should not significantly affect the subsequent analysis.

[2]Note: within the Eurostoxx 600 the components do not trade in the same currency, or by the same exchanges; there are six different currencies and 21 different exchanges.

corresponds to the number of "working business days" (non-holiday weekdays) within the period considered.

3.2.2 The cleaning process

Therefore, a $T \times N$ matrix of close prices is given. However, these prices need to be cleaned before any computation can be performed, as several problems can occur:

- **Missing values:** Firstly, the bank-holidays may differ between the different exchanges where the stocks considered are traded. In fact, the four indexes selected for this analysis incorporate stocks traded on 21 separate exchange markets. Therefore, all missing data resulting from a market being closed, has been replicated with the previous available value; as is common practice, subsequently defaulting all non-trading days to zero-return days.

- **Abnormal prices:** Abnormal prices are trimmed out. Prices are considered to be abnormal if an "abnormal jump" is observed in the closing price series. Jumps may occur if, for instance, a wrong report is sent from the exchange, or a corporate action has been incorrectly reported. If s_t^i denotes the closing price of the stock i at time t, an "abnormal jump" will be detected if:

$$s_{t+1}^i > s_t^i(1+e) \text{ or } s_{t+1}^i < s_t^i(1-e)$$

  where e stands for the maximum jump allowed in percentage form. In practice this is usually $e = 30\%$, as it corresponds to an extreme case (for most European markets, continuous trading is suspended by the exchange if the stock price jumps by more than 10%. The average daily volatility of the stocks considered over the period specified for this study is $1.94\% \ll 30\%$). In such a case, all abnormal prices will be replaced by the last available closing price.

The stock returns are computed from the observed closing prices. In fact, it is more convenient to consider returns rather than prices, because as opposed to prices, returns are more likely to be stationary processes, and in addition, the problems with price scaling when comparing differing stocks held in differing currencies can be avoided.

3.2.3 The returns

The following section offers a more formal description of the computation and cleaning of the stock returns; in addition, some descriptive statistics will be provided, and basic normality tests run on the return time-series.

3.2.3.1 Computation and cleaning process

Let $(\mathbf{r^i})_{1\leq i\leq N}$ denote the $T-1$-dimensional vectors of the geometric returns for the stock i; thus:

$$\forall i \in [1, N],\, \forall t \in [2, T],\, r_t^i = \log(\frac{s_t^i}{s_{t-1}^i}) \approx \frac{s_t^i}{s_{t-1}^i} - 1$$

when $\frac{s_t^i}{s_{t-1}^i} - 1$ is close enough to zero.
$\mathbf{M}$ denotes the $(T-1)\times N$ matrix of stocks returns: $\mathbf{M} = \{(r_t^i)_{2\leq t\leq T}^{1\leq i\leq N}\}$.
As some abnormal prices may have escaped the basic cleaning process, and considering that outliers have the potential to compromise the analysis, it is important to undertake additional procedures that enable abnormal returns to be detected and withdrawn by the analyst. Thus, the next stage of the procedure is the trimming of returns in order to remove extreme data from the observations. A Winsorisation procedure is used to detect major outliers. To perform this procedure, all return values beyond three standard deviation of the empirical mean of each return series are cut off (it is found in this study that only 1.70% of the returns are abnormal).

Note that the outliers are understood as outliers for modelling purposes. Indeed, when building a model to represent asset returns, the investor needs to prevent outliers or extreme events from corrupting the estimation of a given model. However, to be strictly consistant, the investor should only remove those outliers for modelling purposes and re-introduce them when evaluating the performance of the model on real data (an extreme return may be legitimate if an extreme event led the asset return up or down - like for instance bankruptcy rumor, merger rumor...). In this PhD thesis only one clean data set is considered both for estimating and evaluating asset return models, as the purpose is more to compare models than to precisely evaluate them. However, in practice, one should always consider two data sets: one for modelling and one for back testing.

Figure (3.1) displays the empirical returns of the four indexes considered in this study.

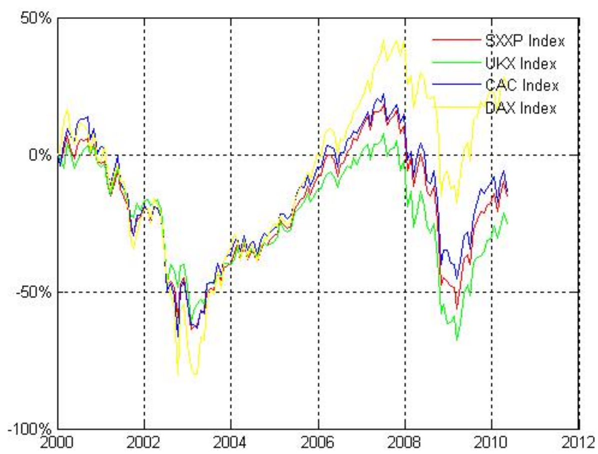


Figure 3.1: Cumulative Indexes Returns in %

3.2.3.2 Some basic statistics

The returns distribution will be briefly studied, and in particular, simple normality tests performed. It is commonly assumed that stock returns are normally distributed (as in the Black and Scholes option pricing theory). However, as has been shown in numerous empirical studies, this assumption does not, in fact, hold true (see for instance Longin & Solnik (2001)); furthermore, it is confirmed by the non-normal distribution seen in the sample considered by this study. To give a more precise description of stock returns, and in addition, to test the normality of the returns distribution, the following statistics displayed in Table (3.1) will be computed for the empirical distribution of each stock contained within each of the four indexes:

- **Mean, Standard Deviation, Kurtosis and Skewness:**[1] The empirical cross-sectional median is computed for all four values for the four indexes considered (i.e. computed daily on the series of the index constituent returns).

- **The percentage of p-value smaller than 5% relative to the Jarque-Bera normality test:** the percentage of a given index components return distribution for which the null hypothesis of normality has been rejected is computed. The statistic for the Jarque-Bera normality test is given by:

  $$n[\frac{\mu_3^2}{6\mu_2^3}+\frac{1}{24}(\frac{\mu_4}{\mu_2^2-3})^2]$$

  where μ_i denotes the ith moment of the returns empirical distribution, n denotes the number of observations ($n=(T-1)*N$).

  Under the null assumption H_0 of normality for the considered stock returns distribution, the test statistic follows a chi-square distribution with two degrees of freedom.

- **The percentage of p-value smaller than 5%, relative to the Kolmogorov-Smirnov normality test:** the percentage of a given index components return distributions for which the null hypothesis of normality is rejected is computed. The Kolmogorov-Smirnov test compares the empirically observed cumulative distribution F_X of a series of observations denoted X, to a normal cumulative distribution F with same mean and variance than the empirical mean and variance of the observed time series X. The test statistic is given by:

  $\mathrm{D}_X \equiv \sup_x(|F_X(x)-F(x)|)$

[1] note: the Kurtosis figure is expressed as an excess, compared to a Kurtosis that is equal to 3 for a normal distribution.

under the Kolmogorov-Smirnov test null hypothesis that the sample comes from the distribution $F(x)$, $\sqrt{D_X}$ converges in distribution toward $\sup_x(B(F(x)))$, where $B(.)$ is the Brownian Bridge distribution.

By testing the same hypothesis on several series (the same test are conducted simultaneously on the return distributions of all the different components of a given index), a simultaneous testing problem is confronted; thus, the first order risk (here chosen to be 5%) must be adjusted in order to prevent bias. The Bonferroni adjustment consists in dividing the single test first order risk by the number of tests per index (i.e. the number of constituents per index; for instance, 40 for the CAC40). Thus, the Bonferroni adjustment is made procedurally for all the subsequent statistics.

	SXXP Index	UKX Index	CAC Index	DAX Index
Median mean (Bps)	2.98	2.64	1.54	3.09
Median std (Bps)	220.05	221.28	230.71	226.43
Median Kurtosis	7.35	7.34	5.82	6.50
Median Skewness	0.25	0.26	0.25	0.30
% of J-B test p-value $\leq$ 5%	99.83	99.03	97.50	96.67
% of K-S test p-value $\leq$ 5%	99.83	99.03	97.50	96.67

Table 3.1: Normality Statistics

At this point, the following observations can be made:

- As the market was generally up for the period covering January 2000-April 2010, the mean return is positive for all four indexes, which accounts for the slight positivity of the Skewness of the distribution.

- As the Kurtosis is around 6-7, the return distributions can be said to show "fat tails" (Kurtosis of a normal distribution is equal to $3 << 6$); confirming previous observations made by Longin & Solnik (2001).

- The Jarque-Bera tests reject normality for almost all the distributions in each market; the distributions are clearly abnormal, and this seems to be mainly due to more extreme values.

Now that the data-set has been introduced, the empirical modelling of the different types of factor-models can be undertaken.

3.3 The Capital Asset Pricing Model (CAPM)

The CAPM, although widely criticised, remains a central tool for assessing asset allocation problems. In particular, investors often want to achieve Beta neutrality (i.e. they compute the CAPM market Beta for each asset, and they construct a Beta neutral portfolio). More precisely, an asset allocation $\boldsymbol{\phi}$ is said to be market- or Beta-neutral if:

$$\sum_{i=1}^{N} \beta^{i,M} \phi^i = 0$$

where $\beta^{i,M}$ is the CAPM Beta, or the sensitivity of the i stock returns against the market returns (as specified in Chapter 2). In the following section, the CAPM portfolio will be tested. In a first sub section, some notations of the CAPM model, introduced in Chapter 2, will be outlined. More specifically, this section will cover how the CAPM Betas are estimated on empirical time series. The discussion will then focus on a major issue of the CAPM, which is the instability of the Betas. Finally, the CAPM model will be evaluated using the different statistical measures introduced by Jensen (1969).

3.3.1 Estimation of the CAPM model

In this section, the value of the CAPM Beta estimates will be explicated, providing some basic statistics for the "goodness of fit" for the CAPM.

3.3.1.1 Betas Estimation

To recall, the CAPM equation is given by:

$$\forall i \in [1, N], t \in [1, T], \quad r_t^i = r_f + \beta^i (r_t^M - r_f) + \epsilon_t^i \tag{3.1}$$

where ϵ_t^i represents the residual of the model for the stock i.

The stock-market is typically represented by an index; such as the four indexes presented in Section 3.2. The financial risk of any asset i is decomposed into two components: a systematic risk (or market risk) that is modelled through the Beta $(\beta^i(r_t^M - r_f))$, and a specific risk (ϵ_t^i). The Beta is estimated to be the coefficient of regression, when the vector of returns of a given financial asset is regressed against the vector of returns of the market-index. More formally:

$$\beta^i \equiv \frac{\mathbb{COV}(\mathbf{r^i}, \mathbf{r^M})}{\mathbb{VAR}(\mathbf{r^M})} = \rho(\mathbf{r^i}, \mathbf{r^M})\frac{\sigma(\mathbf{r^i})}{\sigma(\mathbf{r^M})}$$

where ρ stands for the correlation operator between two variables, and σ stands for the standard deviation of a given variable. $\mathbf{r^M}$ corresponds to the vector of observed returns of the index between time 2 and T, and $\mathbf{r^i}$ corresponds to the vector of observed returns of the ith component of the index considered between time 2 and T. A histogram of the estimated Betas for the Eurostoxx 600 components is displayed in Figure (3.2).

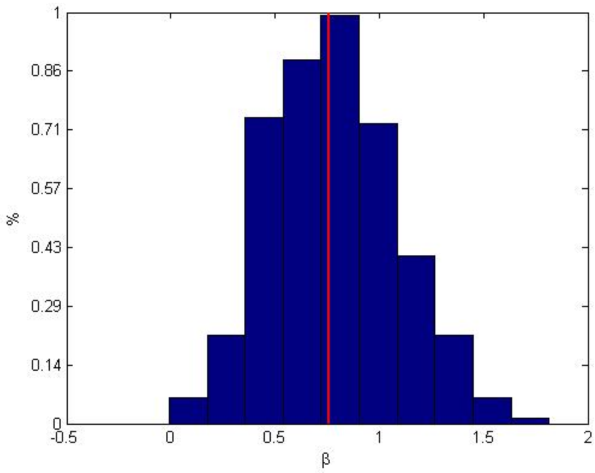


Figure 3.2: Market Betas for the Eurostoxx 600 components

The assumption of constant Betas over the time-window $[1, T]$ is not verified, as will be seen in Section 3.3.2, where the stability of rolling Beta time-series is tested.

3.3.1.2 Goodness of fit of the CAPM

As can be seen in the histogram representation of the R^2 relative to each of the CAPM regressions in Figure (3.3), the average adjusted CAPM R^2 is below 20%.

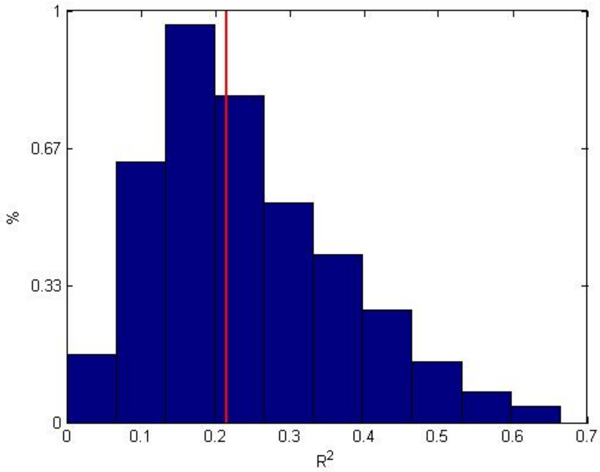


Figure 3.3: CAPM R^2

The CAPM fails to fully explain stock returns. The reason for this is probably due to the fact that important factors had been disregarded in order to explain stock return variability; thus, highlighting the fact that the market return alone is not sufficient to explain stock returns. This is confirmed by closer inspection of the CAPM residuals. The CAPM residuals appear not to be either normally distributed, nor independent. In Figure (3.4) a quantile-quantile plot is displayed, showing the CAPM residuals distribution compared against a normal distribution with same mean and variance. The empirical distributions of the residuals display heavy tails.
With a low R^2 and non-normal residuals, the CAPM appears weak. In particular, the aforementioned assumption made on constant Betas cannot be held, which will be demonstrated in the next section.

3.3.2 Betas instability

A static model, such as the basic CAPM, can be a good predictive tool, on the condition that the parameters remain stable over time. This is the main reason

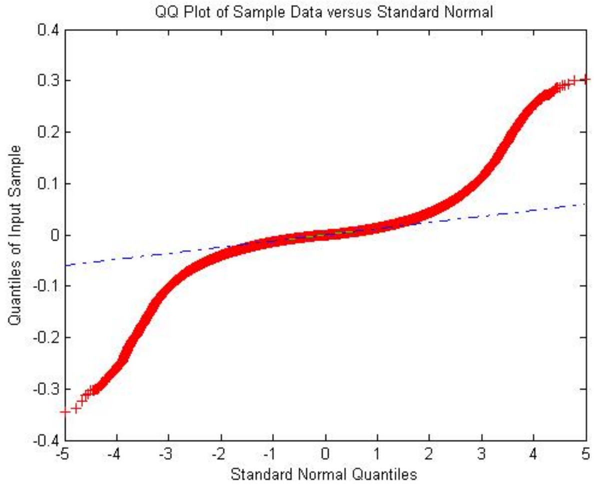


Figure 3.4: CAPM Residuals QQ Plot

a deeper analysis on the CAPM Betas has been conducted.

3.3.2.1 Betas time series

Although generally held to cover several periods, the CAPM is a one-period static model; and as such, the underlying assumption is that the CAPM Beta is constant over time. The CAPM works on the assumption that the returns distribution mean and standard deviation are constant and common to all investors (homogeneous anticipations), and all investors share the same time-horizon. Jensen (1969) relaxes the later assumption by introducing a multi-period horizon, and he proves that the Betas are constant whatever the length of the time horizon set. Merton (1973) extends the CAPM model into a multi-period model, where the conditions under which the single period CAPM can be extended directly to an inter-temporal model are given: if the investment opportunity set stays constant over time and the investor preferences are not state dependent then the inter temporal portfolio maximisation can be treated as if the investor had a single period utility function [1]. On US market data, Fama & French (1992) expose the predictive weakness of the CAPM. The Betas do not seem to satisfy the assumption that they are constant over the prediction period (i.e. it is difficult to use the

[1]This was later developed and extended in the theory of stochastic control through the Hamilton-Jacobi-Bellman equation.

Betas estimated on past periods to estimate future stock returns). In his paper testing the CAPM empirical performances, Jensen also points out the instability of the Betas as he computes Betas estimates on two consecutive non-overlapping samples and scatter plots the two series of Betas estimated (the resulting plot being rather more scatter than a straight line).
The empirical evidence put forward by Jenson is confirmed in the data-set studied here. The confirmation can be made by the fact that when a rolling window regression procedure is performed over 100 days (the Betas are computed on a non-overlapping sub-sample of 100 observations), the resulting Beta time-series are not stable.In the following subsection, stationary tests will be performed to confirm the evidence that the CAPM Beta time-series are non-stationary.

3.3.2.2 Stationary tests

Three different stationary tests that have been computed are displayed in Table (3.2).

% of t-test p-value $\leq$ 5%	95.21
% of D-F test p-value $\leq$ 5%	97.36
% of P-P test p-value $\leq$ 5%	70.79

Table 3.2: Betas Stationarity Statistics

- ***The T test:*** compares the Beta time series values computed over a rolling window of 100 days, with the Beta that has been computed over the whole data-set. The null hypothesis assumes the Beta time-series is equal to the Beta has been computed over the whole data set. The statistics for the test are $\frac{\beta_t - \overline{\beta}}{\sigma(\beta_t)}$, when (β_t) is the Beta rolling window time-series, and $\overline{\beta}$ is the Beta that has been computed over the whole period.

- ***The Augmented Dickey Fuller Unit Root Test (D-F Test):*** An augmented Dickey-Fuller test is a test for a unit root in a time series sample (see Dickey & Fuller (1979)). This test is a version of the Dickey-Fuller test which takes into account more auto regressive lags than the original Dickey-Fuller test. The Dickey-Fuller test tests whether the slope b is equal to one in the econometric equation:

$$y_t - y_{t-1} = a + (b-1)y_{t-1} + \sum_{k=1}^{K} c_k(y_{t-k} - y_{t-k-1}) + \epsilon_t$$

where t is an integer greater than zero indexing time and K is the number of autoregressive lags considered. b is estimated by least square. The test statistic $T(\widehat{b}-1)$ has a known distribution (T is the sample size), which is different from a classical t-test statistics. The null hypothesis is $b = 1$, which means that the time series is not stationary.

- ***The Phillips-Perron Unit Root Test (P-P test):*** as the ADF, it tests for the null hypothesis of a unit root. It is more robust than the ADF test to general forms of heteroscedasticity of the residuals, and it is not necessary to specify an autoregressive lag length as is needed in the case of the ADF test (the parameter K). For more details, see Phillips & Perron (1988).

Although when considering t-tests one cannot reject the hypothesis that the Beta time series-mean is equal to the Beta computed over the whole period, the hypothesis of stationarity is rejected for almost all the Beta time-series. indeed, the null hypothesis of non-stationarity cannot be rejected for more than 90% of the ADF, and almost 80% of the PP tests computed. In the data sample that has been presented, the Beta series has significantly evolved over time.
The instability of the Betas highlights the need for a more complex setting than the one used by the initial CAPM. In particular, to explain stock returns, a number of extra factors need to be considered in addition to the market.

In this chapter, the methodology developed by Jensen (1969) to study time-series of fund performance is used extensively to compare the different factor models that are the subject of this research. In the next section, the Jensen statistics will be studied in detail, and applied to the CAPM model that has been computed using the core data-set considered in this PhD thesis.

3.3.3 The Jensen statistics

The measures Jensen proposes to test whether the CAPM holds on empirical data are computed. Jensen uses those statistics to evaluate the performance of US fund managers. The idea is to deduce the Alpha for the aggregated portfolios for each fund manager. Jensen finds that the Alphas are not significant, and comes to the conclusion that fund managers have no specific stock selection expertise. The approach presented here is different, as it looks at single stock Alphas for all the market index constituents. The Alphas are used to compare factor models, rather than to evaluate the performance of a given portfolio. The Alpha can be seen as a performance criterion. There is a trade off to make here, as the Alpha represents a trading opportunity, as long as the investor trusts their model: the smaller the Alphas, the better the model is on average. However, if the Alpha is too big, it is a signal for the investor that the model is missing information for the purposes of stock return explanation.
More precisely, the following Jensen statistics are considered for the purposes of testing the CAPM. The statistics value are displayed in Table (3.3).

	SXXP Index	UKX Index	CAC Index	DAX Index
Mean β	0.80	0.86	0.90	0.84
Mean Index/Stock corr (%)	47.12	53.99	59.23	58.24
Mean Residuals/Index corr (%)	0.00	−0.00	0.00	0.00
Mean Residuals/Time corr (%)	−0.43	−0.46	−0.46	−0.36
Mean Residuals autocorr (%)	−1.16	−1.50	0.06	0.09
Mean R^2(%)	41.04	34.40	22.33	27.41
Mean α(%)	76.09	68.10	60.84	60.82

Table 3.3: CAPM Statistics

- ***The Beta:*** The average arithmetic Beta is smaller than one, which is caused by the fact that bigger stocks in the index have smaller Betas than smaller stocks; and in addition, they have bigger weight in the index than smaller stocks (the Beta of the index itself is of course one).

- ***The Stock/Index Correlation:*** corresponds to the average correlation between stock returns and the market index. The CAPM assumes there exists a linear stable relationship between stock returns and the market returns. A market correlation of around 50% is observed (slightly less for the Eurostoxx 600, which is probably due to a size effect: with an increasing number of constituents, the average correlation of the single stocks returns and the index returns will decrease). The index variation explains only about half the variation of the single stock returns. The CAPM does not fully explain stock returns.

- ***The Residuals/Index Correlation:*** corresponds to the average correlation between residuals and the market index. The residuals have on average a null correlation with the index. This fits with the CAPM assumption that the residuals are independent of the explanatory factor of the CAPM regression (i.e. the market - here assimilated to the index return).

- ***The Residuals First Order Autocorrelation:*** for the autocorrelation of a time series with T observations to be significantly different from zero, it must be above $\frac{1}{\sqrt{T}} \simeq 2.15\%$ in absolute value. It can be seen that this is not the case for all of the four markets considered: the residuals are not significantly auto correlated. Indeed, this also fits with the CAPM assumption of IID residuals; the CAPM is not state dependent, so the error series should not be autocorrelated.

- ***The Residuals-Time Correlation :*** the residuals are not significantly correlated with time, as, for the four markets considered, the estimated correlation is always smaller than $\frac{1}{\sqrt{T}} \simeq 2.15\%$ in absolute value. If the residuals were time dependant, it would interfere with the CAPM assumption. The CAPM assumes stable state conditions, so the error series should be uncorrelated with time.

- ***The Alpha:*** Finally and more importantly, the Alpha is computed, as described by Jensen[1]. The Alpha corresponds to the average difference

[1] Jensen (1969) did not actually used the term Alpha to denote the excess returns with respect to the CAPM expected return, but he used the notation δ^*.

between the observed stock return and its CAPM predicted value. It is a synthetic measure of model performance. The Alpha is scaled by the average stock return, so the Alpha for the risky asset i, is expressed as a percentage of the average stock return for the stock i:

$$\alpha^i \equiv \frac{\mu^{\mathbf{r^i}} - \beta^i \mu^{\mathbf{r^M}}}{\mu^{\mathbf{r^i}}}.$$

The average Alpha is around 70%, which means that the CAPM fails to explain around 70% of the stock returns. That is why an investigation of a more elaborate model is needed to explain stock returns.

Although the market return is a significant variable to explain stock returns, as the Betas are statistically different from zeros, the CAPM is not a sufficient model to explain and predict stock returns. Indeed, it lacks goodness of fit (with a $R^2 << 50\%$ for almost all regressions). In addition, the CAPM is not stable, as the Betas seem to evolve significantly over time. In the following, different Asset Pricing Theory (APT) models are tested, considering a greater number of factors to explain stock returns.

3.4 An Exogenous Factor Model (EFM)

In this Section, a classical Exogenous Factor Model is presented. The stock returns are decomposed into the systematic returns explained by exogenous factor returns and an idiosyncratic portion that corresponds to the residuals of the factor model. In addition to the market factor, three exogenous factors that are common for all the stocks will be considered: an oil index factor, an interest rate factor and a volatility factor. First, a detailed description is made of the three exogenous factors considered, and intuition about the expected effects of the different factors on stock returns is given. Then, a more formal description of EFM model is made; and finally, the Jensen statistics associated with this model are presented.

3.4.1 The exogenous factors

The choice of the three factors considered here is mainly motivated by the historical context of the considered time period: more precisely, due to the situation in the Middle East, oil prices have been rising dramatically, which has had a significant influence on stock prices. Also, after the 11th September 2001, and the more recent subprime crisis, the stock market has shown an increase in correlation, and investors have been increasingly aware of the volatility factor in explaining stock returns[1]. Finally, the period considered has experienced a decrease in interest rates, primarily because central banks employ "stimulus packages", such as quantitative easing monetary policies to curb the effects of major financial shocks.

3.4.1.1 Expected sensitivities toward the exogenous factors

For economic and rational reasons, the following sensitivities of stock returns for oil, interest rates and volatility factors are expected:

- ***Factor CO1***. The Crude Brent future closing price return as it is quoted on the London Stock Exchange is chosen; therefore, it is synchronous with the European markets. When oil prices rise, stock prices tend to drop (but not for oil companies): oil is often a raw material used by industrial companies, an increase in oil prices is often a burden for an economy, specially in the US, where the domestic market is highly dependent on oil supply.

- ***Factor GDBR10***. The Euro-Bund yield to maturity quoted on the German Exchange is chosen, which corresponds to the implied rate of return for a bond maturity of 10 years.[2] When interest rates rise, stock prices tend to drop, and vice versa. Several reasons could be mentioned: first, the discounted value of future stock dividends drops; therefore, the value of the stock is negatively affected. On the other hand, the fixed income

[1] when the volatility of stock returns tend to increase, reflecting a shock in financial markets, the correlation of stock returns tends to increase as well, this phenomenon is often referred to as the "skew effect".

[2] In the context of the Euro, continental European interest rates are closely correlated one another, and up until the subprime crisis, the UK interest rates are also very correlated to the German interest rates.

investment offers higher return, and investors tend to shift their investment slightly, from equity to fixed income assets. Note that when a bond yield increases, it means the interest rates are going down, and vice versa.

- ***Factor VDAX***. The Virtex index closing price return quoted on the German Exchange is used, which is a volatility index based on the German index DAX components implied volatilities (volatilities obtained when reverting the Black and Scholes formula for At The Money Put and Call option prices). In a bear market, price returns tend to be more volatile than in a bullish market. It is the so called "volatility skew" effect. When the market goes down, stock prices tend to be more correlated; therefore, increasing the overall volatility of a stock portfolio (for a recent analysis of this effect, see Longin & Solnik (2001)). The implied volatility of stock options is a good indicator of anticipated future stock volatility.

3.4.1.2 Exogenous factor returns and index returns

Before presenting the model, the statistical relationship between index returns and the three exogenous factor returns described above will be analysed. First, an analysis will be made of the correlation matrix of index and factor returns; then basic regressions of index returns will be run on factor returns to study the strength of the statistical relationship.
Figure (3.5) displays the three factor returns over the considered period; and Table (3.4) displays the correlation matrix of the factors and index returns.

	VDAX Index	GDBR10 Comdty	CO1 Comdty	SXXP Index	UKX Index	CAC Index	DAX Index
VDAX Index	100.00	−31.00	−13.63	−71.53	−62.70	−69.70	−70.42
GDBR10 Comdty	−31.00	100.00	14.47	42.50	39.42	40.97	38.37
CO1 Comdty	−13.63	14.47	100.00	20.23	22.60	18.07	14.32
SXXP Index	−71.53	42.50	20.23	100.00	93.56	96.07	89.25
UKX Index	−62.70	39.42	22.60	93.56	100.00	85.43	77.30
CAC Index	−69.70	40.97	18.07	96.07	85.43	100.00	87.81
DAX Index	−70.42	38.37	14.32	89.25	77.30	87.81	100.00

Table 3.4: Index Factor Correlation (%)

It can be noticed that:

- ***Factor VDAX:*** All the four index returns have a strong negative correlation with the volatility index returns. The data here, confirms the "volatility

Figure 3.5: Exogenous Cumulative Factors Returns in %

skew effect" that had been expected: the higher the volatility return, the lower the index returns.

- ***Factor GDBR10:*** Higher interest rates tend to make bond investments more attractive, which is to the detriment of stock investment. When the yield of a bond increases, the price drops, because the bond coupon gets smaller relative to the expected rate of returns.

- ***Factor CO1:*** The oil index effect is somewhat curious. As mentionned above, when oil prices increase, the expected effect is a negative return for equity prices. However, a small positive effect is noticed. Over the period considered, the oil price has risen dramatically, see Figure (3.5). In addition, over the same period, the index prices have mostly increased, see Figure (3.1).

- The three exogenous factors are not greatly correlated one another; thus, making good candidates for a multivariate model.

A pre-analysis consists of a study of how the stock returns are correlated with the return of the three factors that have been pre-selected. To study the sensitivities of the index returns towards the different factor returns, four different regressions have been run on the exogenous factors; thus, one for each of the four indexes used in this study. The results are presented in Table (3.5). For each regression, the estimated coefficients for the intercept and the three factors VDAX, GDBR10 and CO1 Betas are presented. In addition, a 95% confidence interval is given for each of the coefficient estimates.
The intercept is never significantly different from 0, as the confidence interval contains 0 for all the four regressions. The Oil factor (CO1) seems to be also non-significant. However, the Volatility (VDAX) and interest Rate (GDBR10) factors are significant. This tends to show, that for all the four indexes considered, if the volatility factor increases, the stock returns decrease; and if the bund yield increases, the stock returns increase also - which is consistent with the remarks made in the factors/stock returns correlation analysis. The average R^2 is around 50% for all four index regressions, which is higher than for the CAPM model

SXXP Index: R2=56.43			
	Beta	lower bound	upper bound
Intercept	0.00	−0.00	0.00
VDAX Index	−0.18	−0.19	−0.17
GDBR10 Comdty	0.07	0.06	0.08
CO1 Comdty	0.05	0.03	0.06
UKX Index: R2=45.17			
	Beta	lower bound	upper bound
Intercept	0.00	−0.00	0.00
VDAX Index	−0.17	−0.17	−0.16
GDBR10 Comdty	0.07	0.06	0.08
CO1 Comdty	0.07	0.06	0.09
CAC Index: R2=53.15			
	Beta	lower bound	upper bound
Intercept	0.00	−0.00	0.00
VDAX Index	−0.21	−0.22	−0.20
GDBR10 Comdty	0.08	0.07	0.09
CO1 Comdty	0.04	0.03	0.06
DAX Index: R2=52.69			
	Beta	lower bound	upper bound
Intercept	0.00	−0.00	0.00
VDAX Index	−0.23	−0.23	−0.22
GDBR10 Comdty	0.07	0.06	0.08
CO1 Comdty	0.02	0.00	0.04

Table 3.5: Index/Factors Regressions

(see Table (3.3)). Therefore, the addition of three exogenous factors improves the explanative power of the CAPM model.

3.4.2 The EFM model

In this section, a simple approach is used to explain stock returns by considering four different exogenous factors. As the factors considered are not necessarily independent the model can be viewed as an extension of the APT model developed by Ross (1976) and briefly presented in Chapter 2. The model estimation is performed as if the residuals were un-correlated with the factors. The following

section will briefly present the model and discuss the results.
For each stock i, the following regression is run:

$$r_t^i = a^i + \beta^{i,M} r_t^M + \beta^{i,VDAX} r_t^{VDAX} + \beta^{i,GDBR10} r_t^{GDBR10} + \beta^{i,CO1} r_t^{CO1} + \epsilon_t^i \quad (3.2)$$

where a^i stands for the constant regression coefficient, r_t^i stands for the stock return i at time t, r_t^M corresponds to the return of the index to which the stock considered belongs to, r_t^{VDAX} represents the Volatility factor return, r_t^{GDBR10} represents the Interest Rate factor return, r_t^{CO1} represents the Oil factor return and ϵ_t^i represents the residual of the model for the stock i.

As in the previous section dedicated to the CAPM, this model is tested on the four considered markets (SPXX, CAC, DAX and UKX).

3.4.3 Results

The same methodology as the one presented for the CAPM is used to test the model. The next section will first discuss the Beta instability of the EFM, and then, give the Jensen statistics associated with the model.

3.4.3.1 Betas instability

As for the CAPM, a major problem of the EFM is the instability of the Betas over time. If a relationship holds for a certain period between stock returns and some factor returns, it may not be stable over time. Table (3.6) summaries the results of the stationary tests computed previously on the CAPM Betas. The EFM Betas seem to be slightly more stable than the CAPM Betas: the EFM t-test rejects less often the null hypothesis of constant Betas mean for the Market Betas. Also the D-F tests and P-P tests reject the null hypothesis of non-stationarity for more than 70% of the cases for the Oil, Interest Rate and Volatility Betas time series. Therefore, the addition of external factors tends to reduce the instability of the model.

	Market Betas
% of t-test p-value $\leq$ 5%	91.07
% of D-F test p-value $\leq$ 5%	37.31
% of P-P test p-value $\leq$ 5%	37.07
	VDAX Index
% of t-test p-value $\leq$ 5%	77.85
% of D-F test p-value $\leq$ 5%	13.75
% of P-P test p-value $\leq$ 5%	13.28
	GDBR10 Comdty
% of t-test p-value $\leq$ 5%	92.73
% of D-F test p-value $\leq$ 5%	11.02
% of P-P test p-value $\leq$ 5%	10.51
	CO1 Comdty
% of t-test p-value $\leq$ 5%	85.12
% of D-F test p-value $\leq$ 5%	15.12
% of P-P test p-value $\leq$ 5%	14.67

Table 3.6: EFM Betas Stationarity Statistics

3.4.3.2 Jensen statistics

In the EFM, the Alpha for the risky asset i is defined as:

$$\alpha^i \equiv \frac{\mu^{r^i} - \beta^{i,M}\mu^{r^M} - \beta^{i,VDAX}\mu^{r^{VDAX}} - \beta^{i,GDBR10}\mu^{r^{CO1}} - \beta^{i,GDBR10}\mu^{r^{CO1}}}{\mu^{r^i}}$$

The higher the Alpha, the less the model can explain stock returns on average. The Alphas of the External Factor Model are slightly smaller than the Alphas of the CAPM. To conclude, it can be said that the External Factor Model improves the plain CAPM. The external factors seem to play an additional role to the Market factor to explain stock returns.

Table (3.7) gives a synthesis of the main statistics obtained from the four models estimated (one for each of the four market considered). The following remarks can be made when comparing them with the CAPM statistics of Table (3.3):

	SXXP Index	UKX Index	CAC Index	DAX Index
Mean $\beta^{M}(\%)$	74.87	80.51	85.29	75.22
Mean $\beta^{VDAX}(\%)$	-1.49	-1.86	-1.48	-3.24
Mean $\beta^{GDBR10}(\%)$	0.61	0.81	0.93	1.50
Mean $\beta^{CO1}(\%)$	1.59	0.97	0.46	1.66
Mean Residuals/Index corr (%)	-0.00	-0.00	-0.00	-0.00
Mean Residuals/Time corr (%)	-0.38	-0.41	-0.43	-0.21
Mean Residuals autocorr (%)	-1.28	-1.76	-0.19	0.02
Mean R^2 (%)	41.38	36.42	23.21	27.79
Mean $\alpha(\%)$	75.31	67.14	60.07	59.38

Table 3.7: EFM Statistics

- The market Betas are lower than in the pure CAPM model for every one of the indexes considered.
- As in the case of the CAPM, the residuals are not correlated with the market, they are not time dependent or auto correlated.
- The average R^2 statistics are higher than for the CAPM for the four indexes considered (The R^2 are above 60% for the four indexes tested, whereas the CAPM R^2 is never greater than 45%). The explanative power of the Exogenous Factor Model is higher than that of the CAPM.
- The average Alpha is consequently lower than for the CAPM, which confirms the fact that the additional three factors improve significantly the CAPM.

Obviously, the more significant factors the investor considers, the better the model will be. The model can be dramatically refined by adding country or sector specific factors; as presented in the Barra model (see Stefek (2002)). An elaborated exogenous factor model, with a larger number of factors, can certainly out perform the CAPM. However, the tractability and implementation have also to be considered when choosing a model. To select the number of factors required to explain stock returns, one must find a trade-off between a good fit (calling for more factors) and over-complexity (calling on the contrary to a smaller number of factors).

In the next section, a different type of factor model is considered, that has fundamental endogenous stock specific factors. Stock by stock regressions are run, and the factor values are specific to each stock.

3.5 A Fundamental Factor Model (FFM)

The following section discusses a number of fundamental factors that explain stock returns. This type of model has first been studied by Fama & French (1993). However, the approach used in this thesis is different because it does not use these factors to build portfolios (based on the ranking of those fundamental factors); they are directly used as the factors of the FFM regression. First, the model is presented, and then empirical estimations will be discussed. The Jensen statistics are used to compare the results of the CAPM model with the EFM model results presented above.

3.5.1 The FFM model

In this section, the FFM is described in greater detail; first, the three fundamental factors considered are described, and then, the FFM equation is presented.

3.5.1.1 The FFM factors

The stock specific factors considered are:

- ***The Price Earning Ratio (PE)***: The relationship between the price of a stock and its earnings per share is calculated as the stock price divided by earnings per share. Earnings per share is calculated on a trailing 12 month basis, where information is available, by adding up the most recent four quarters.

- ***Price to Cash Flaw Ratio (PCF)***: The price to cash flow ratio is the ratio of a stock price divided by the cash flow per share.

- ***Price to Book Ratio (PB)***: The ratio of a stock price divided by the book value per share.

Those three factors have been chosen for their relevance, but also for their tractability: indeed, they constitute clean and available fundamental data, representing reasonably simple financial ratios. Those three factors are also easier to use without adjustment when considering a large portfolio of stocks. With the new IFRS rules, the firms are obliged to provide an Earnings value, which respects strict accountancy rules. This makes the comparison between different European firms easier. The Book ratio and the Cash Flow ratio are straightforward values. As fundamental data is usually dirtier than price related data, a cleaning process is required. Indeed, fundamental data is provided by different brokers, and collected by general data providers such as Reuters or Bloomberg. Because of the variety of sources the data is more likely to be incorrectly reported. First, all missing values are replaced by the last available value. Then the arithmetic returns of the ratios are computed. All returns greater than 100% are defaulted to zero.

3.5.1.2 FFM equation

A cross-sectional regression is used on the stock return space daily. More formally, for each day, the following model is computed:

$$r_t^i = a_t + \beta^{t,PCF} r_t^{i,PCF} + \beta^{t,PE} r_t^{i,PE} + \beta^{t,PB} r_t^{i,PB} + \epsilon_t^i \qquad (3.3)$$

where t stands as an index for day t, a_t stands for the constant regression coefficient, r_t^i stands for the stock return i at time t, $r_t^{i,PCF}$ represents the Price to Cash Flow factor return relative to stock i, $r_t^{i,PE}$ represents the Price to Earning factor return relative to stock i, $r_t^{i,PB}$ represents the Price to Book factor return relative to stock i, and ϵ_t^i represents the residual of the model for the stock i.

3.5.2 Estimation of the FFM model

First, an examination needs to be undertaken of the cross-sectional correlation between the fundamental factor returns and the stock returns, as plotted in Figure (3.6). This figure plots the exponential moving average (half life of one year or 250 observations) of the correlation time-series that increases greatly overtime:

the stock returns get more and more correlated with fundamental factors, up until the collapse of Lehman Brothers in September 2008, that marks the pick of the recent subprime financial crisis. Indeed, it can be noticed that the correlation between the fundamental factors and the stock returns increase steadily, up until the last quarter of 2008: in a period of economic growth, the fundamental financial ratios explain stock returns well. However, at the end of 2008, investors started to loose confidence in the financial ratios due to the fact that many financial statements became doubtful (for instance, the American insurer AIG, the American business bank Lehman Brothers, the national mortgage and loan American companies Freddie Mac and Fanny Mae have been suspected of having manipulated their financial figures). This was especially the case for earning figures, for which valuations became suspicious (the Book Value and Cash Flow Value are more straightforward measures and therefore less subject to caution). That may explain why, since end of 2008, the correlation of the PE ratio with stock returns decreased even more so than the Book ratio or the Cash Flow ratio.

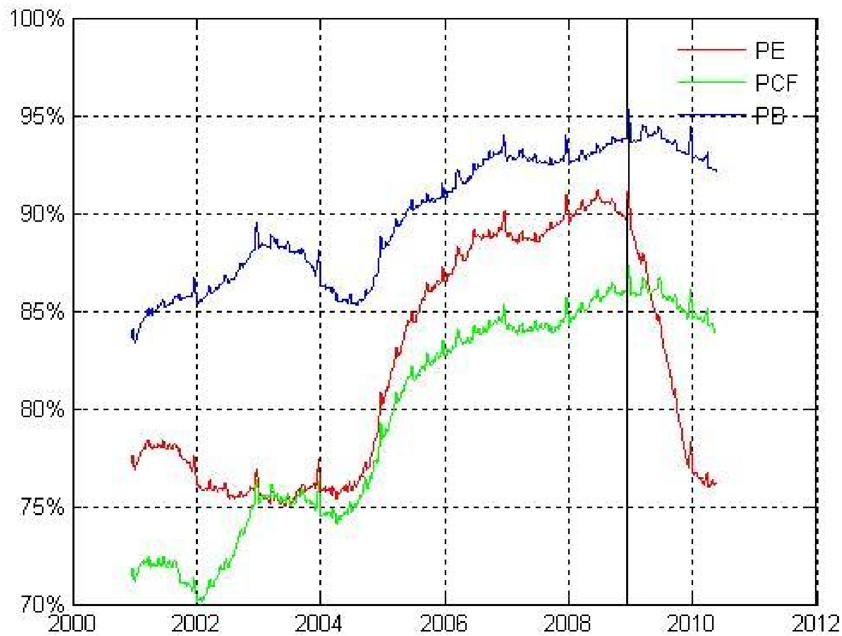


Figure 3.6: Fundamental Factors Returns / Stocks Returns Correlation in %

Table (3.8) displays some basic statistics relative to the FFM. The following remarks can be made:

- The Adjusted R^2 are better than for the CAPM and the EFM. In each of the four markets, the mean R^2 is greater than 90%.
- The Alphas are much smaller than for the CAPM or the EFM (10 to 20% only).

	SXXP Index	UKX Index	CAC Index	DAX Index
Mean $\beta^{PE}(\%)$	11.03	7.46	8.28	15.36
Mean $\beta^{PCF}(\%)$	4.13	7.19	-8.40	4.64
Mean $\beta^{PB}(\%)$	82.12	78.08	87.78	79.01
Mean Residuals autocorr (%)	0.54	-3.43	-3.78	-5.60
Mean R^2 (%)	85.77	85.10	78.97	98.87
Mean $\alpha(\%)$	12.03	8.00	14.58	6.08

Table 3.8: FFM Statistics

This model shows a greater goodness of fit than the two previously studied models. This is due to the fact that much more data is used compared with the other models. Therefore, there is a better granularity of information to explain stock returns. Also many fund managers base their stock picking on the analysis of financial ratios: these measures being widely monitored, they tend to significantly affect stock returns. The Alphas are around 10 to 20%, which represents a dramatic improvement compare to the CAPM or EFM models.

3.5.3 Comments on the FFM

Out of the all the different factor models discussed in this thesis that are based upon external factor selection, the best model appears to be the FFM.
The following sections will focus on statistical models and present more specifically three different statistical models (Principal Component Analysis, Independent Component Analysis and Cluster Analysis), that, conceptually speaking, differ greatly from the precedent models presented above. In those models indeed, the factors are said to be endogenous as they are identified during the

model estimation process; on the contrary to the models previously presented, where exogenous factors are pre-selected before the estimation phase.

3.6 Principal Component Analysis (PCA)

A very common endogenous statistical factor model is based on Principal Component Analysis. The first sub-section provides a brief recap of the PCA principles, and the second sub-section presents the results of the estimations carried out in this research. Here, the selection of the explanative factors is purely data driven: the factors are selected and estimated through the modelling procedure.

3.6.1 The PCA model

In practice, it is impossible for any investor to be able to analyse the entirety of information available, and; therefore near impossible to select all the factors needed to explain stock returns. Well informed investors have access to powerful information tools - such as, Bloomberg and Reuters terminals - which represent very exhaustive data bases of all types of asset prices; such as: Equity and Fixed Income data, as well as a whole set of worldwide news covering economics, companies, and national and international politics. Over the last few years, many data providers have specialised in clean fundamental data (for instance: FactSet and Starmine). If a large amount of such information is relevant in explaining risky asset price processes, the fact remains that it still contains a high level of noise, and for practical reasons, a restrictive set of factors need to be selected. Indeed, to track and monitor the factors in an effective way, a limited number of factors must be considered (computational power is limited to handle too large a number of explanative factors). When an a priori selection of exogenous factors is not believed to be sufficiently exhaustive to explain stock returns, the PCA offers a methodology that can be used to select the most relevant explanatory factors.

3.6.1.1 The PCA principle

In the case of PCA, the model selection is particularly complex on account of it being two-fold:

- First, it is necessary to define a subset of statistically computed explanatory factors for the whole set of information available, through a defined selection procedure.
- Second, it is necessary to identify those explanatory endogenous factors.

The PCA is a statistical method which allows to decompose the variance of a set of variables (here the stock returns) into uncorrelated factors. Note that for the factors to be also independant, it is necessary to add the strong assumption that the returns are Gaussian (uncorrelated Gaussian vectors are also independant, but more generally speaking uncorrelated vectors are not necessarily independant).
In the following, the PCA procedure is detailed.

3.6.1.2 Factors selection and Betas calibration

An outline of the PCA procedure as applied to the matrix of asset returns will be presented. For an in depth explanation of PCA models, refer to Theil (1971). It is assumed that the historical returns matrix $\mathbf{M}$ is of dimension $(T-1)*N$, where T represents the number of observations equally spaced in time (the frequency is one day) and N the number of financial assets. The idea is to project $\mathbf{M}$ in a new base, where the variance between each transformed variable is maximised (i.e. the new variables or principal components are the most distinct possible in terms of variance). It can be shown that this optimal base corresponds to the orthogonal base constituted of the eigenvectors corresponding to the eigen values of the positive defined symmetric covariance matrix of historical excess returns $\mathbf{M'M}$.

Denoted by λ^i for $i = 1, ..., N$, are the N positive eigenvalues of $\mathbf{M'M}$ and $\mathbf{\Lambda}$, the diagonal matrix with diagonal elements the λ^i. Let $\mathbf{P}$ denote the $N * N$ matrix of the eigenvectors of $\mathbf{M'M}$. Hence: $\mathbf{M'M} = \mathbf{P'\Lambda P}$. The first column of $\mathbf{M}$ is the decomposition of the first principal component on the basis of the asset returns, which corresponds to the highest eigenvalue of $\mathbf{M'M}$; the second column corresponds to the decomposition of the second principal component with the second highest eigenvalue of $\mathbf{M'M}$; and so on.

The fraction of the data variance explained by each of the principal components is given by: $\frac{\lambda^i}{\sum_{j=1}^{N}\lambda^j}$. If $\mathbf{F}$ denotes the matrix of factor returns, therefore:

$$\mathbf{F} = \mathbf{PM}$$

and $\mathbf{M} = \mathbf{P'F}$

The explanatory factors (or principal components) are linear combinations of the risky assets. This is in fact the major shortcoming of endogenous factor models, such as the PCA, as, it is difficult to identify explanatory factors as relevant variables (exogenous variables such as economic indicators). However, the advantage of the PCA is that no a priori knowledge is held on the selection of explanatory factors; therefore, it is an interesting model to run if the investor does not have true knowledge on which exogenous variables can explain stock returns.

3.6.2 PCA Estimation and Results

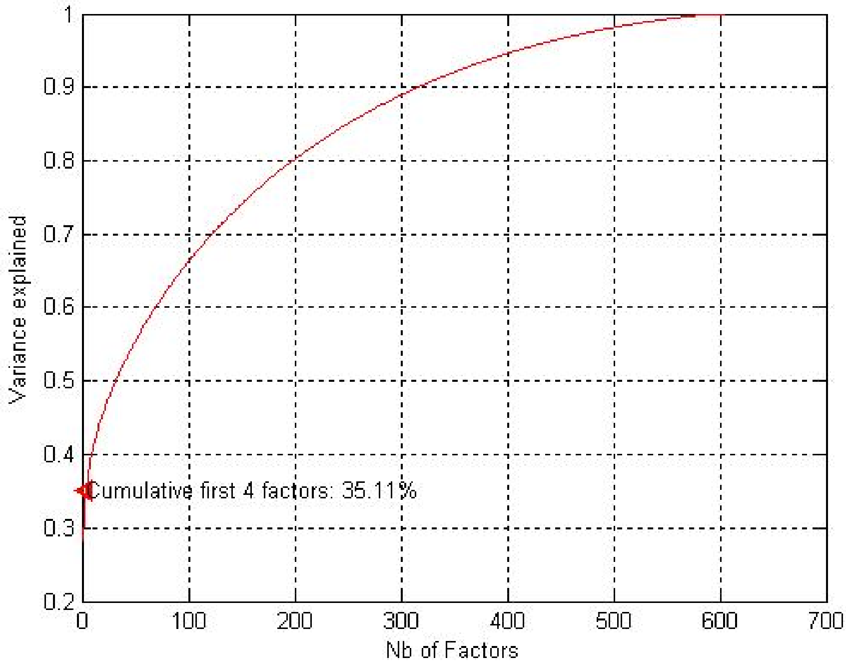
In this study, PCA is performed using the built in "princomp" routine from the statistical toolbox of the statistical software MATLAB 7.5.0. To reduce a market effect bias, the model is run on centred returns (the market return can be approximated as the mean of all asset returns). The stock returns are also normalised by their correlation matrix, in order to prevent highly volatile stocks to biais the procedure.

3.6.2.1 Explanatory power of the PCA

Figure(3.7) displays the cumulative explanatory power of the PCA factors. As can be seen, the first four factors of the PCA explain almost 30% of the total variance of the stock returns. The higher the number of factors selected, the higher the explanatory power.

Here, only the four first factors are selected. There are two main reasons to do so. First, for the exogenous factor models presented in this thesis, a limited

Figure 3.7: PCA Factor Cumulative Explanatory Power



number of factors (four factors for the EFM and three for the FFM models) are utilised. Secondly, the number of factors are selected so that a R^2 above 10% can be obtained. The R^2, assimilated to a measure of model performance, can be expressed as the total percentage of variance explained by the PCA factors.

3.6.2.2 PCA Factors Identification

The PCA factor returns ($\mathbf{F}$) are computed by inverting the loading matrix ($\mathbf{P}$) and multiplying it by the stock return matrix ($\mathbf{M}$). Depending on the number of factors considered, almost any amount of stock return variance can be explained in sample. A big problem, then, is the identification of those factors. As shown in Figure (3.8), the PCA factor returns are very volatile. It is almost impossible to identify them according to their loadings when there are so many stocks in the universe considered ($N = 600$).

A common procedure used to identify such factors is to perform correlation analysis between the statistical endogenous factor returns and a number of external

Figure 3.8: PCA First 4 Factors Returns

factor returns. Table (3.9) displays the correlation matrix of the first four factor returns for the PCA, the SXXP returns (assimilated to the market return), and the three external factors studied previously in Section 3.4 (namely, the volatility factor, the interest rate factor and the oil factor).

	Factor 1	Factor 2	Factor 3	Factor 4
SXXP Index	−93.90	23.70	−1.27	−3.51
VDAX Index	69.99	−14.53	6.43	1.86
GDBR10 Comdty	−42.24	6.12	−0.74	−2.63
CO1 Comdty	−21.82	−12.26	−7.33	11.05

Table 3.9: PCA Factors Correlation

- the first PCA factor is very highly anti-correlated with the SXXP Index (more than 95% of absolute correlation), and positively correlated with the Volatility and Interest Rate factors. This factor could be identified as the

"Market" factor, or the "Size" factor, as commonly found in empirical PCA run on asset returns[1].

- the three other factors display weaker correlations with the other external factors.

Overall, it can be said, that the first factor is the only one to be identified after this first analysis.
The PCA loadings are very volatile over time. The PCA factor loadings have been computed on a rolling window; and as for the Betas of the previous models, the stationary tests reject the assumption of stationarity for the loadings: the PCA factors vary greatly overtime.

3.6.2.3 The Jensen Statistics

The model is described as:

$$\forall i, t, \quad R_t^i = \beta_t^{F1,i} F_t^1 + \beta_t^{F2,i} F_t^2 + \beta_t^{F3,i} F_t^3 + \beta_t^{F4,i} F_t^4 + \epsilon_t^i$$

where $(F_t^j)_{1 \leq j \leq 4}$ are the first four PCA factor returns, and $(\beta_t^{Fj,i})_{1 \leq j \leq 4}$ are the loadings of factor j on stock i, estimated through the PCA procedure. The ϵ_t^i are computed as the difference between the stock returns and the PCA prediction based on the first four factors.
To compare the PCA model to the other models, the statistics that have been computed for the previous exogenous models (CAPM, EFM, FFM) will be studied; see Table (3.10). It can be seen that that the Alphas, that represent the average error between stock returns and the returns computed through the factor model, are very high on average (more than 85% for the four indexes considered). The higher the Alphas, the further away the model prediction is compared to the observed stock returns. In all probability, it is necessary to include many more factors in order to reach a good Alpha value. By increasing the number of factors, the Alphas naturally decrease; as seen in Figure (3.9). To reach an average Alpha of 20%, more than 300 factors need to be considered; however, the more factors

[1] Note: because the PCA aims to explain the total variance of the initial signals, the extracted factors are not signed (in the sense that, the sign of any PCA principal components could be reversed).

added to the PCA model, the more difficult it is to identify them. The number of factors to be selected calls for parsimony. Indeed, the investor needs to make a trade off between the number of factors needed and the Alpha values one wants to achieve.

	SXXP Index	UKX Index	CAC Index	DAX Index
Mean β^{F1}(%)	3.75	9.14	14.82	16.99
Mean β^{F2}(%)	−0.58	−1.35	−0.66	−1.17
Mean β^{F3}(%)	0.28	−1.51	0.71	2.04
Mean β^{F4}(%)	0.02	−0.58	−0.42	−1.66
Mean Residuals autocorr (%)	−2.56	−3.33	−2.04	−2.51
Variance explained(%)	36.84	40.71	47.54	52.10
Mean α(%)	68.68	64.52	64.60	60.64

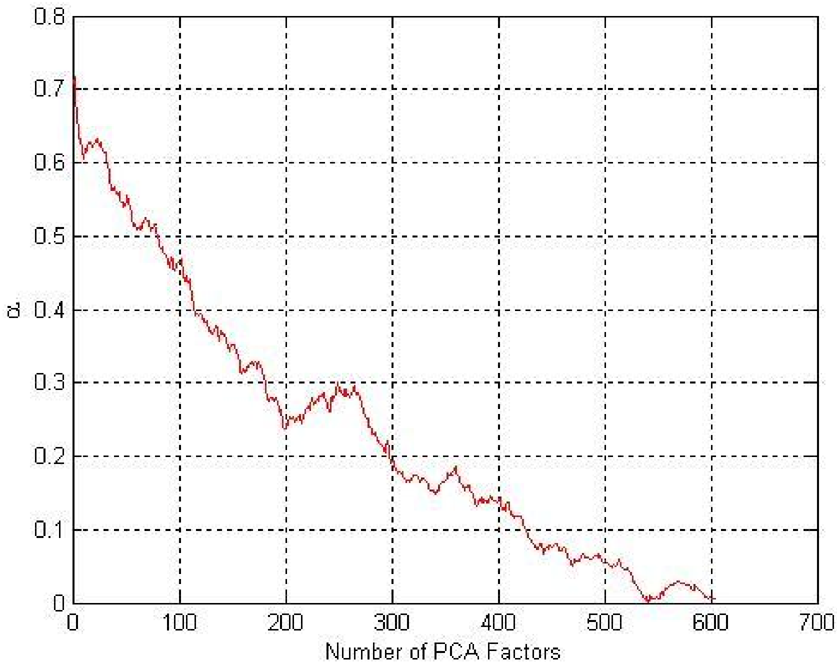
Table 3.10: PCA Statistics

3.6.3 Comment on the PCA model

The PCA model is attractive to investors as it potentially allows the investor to detect "hidden" factors that could explain stock returns. Indeed, it can detect if stock returns are suddenly led by strong factors that may have been overlooked in the first place. The problem is obviously related to the identification of those hidden factors (a correlation analysis with external factors is often used as an identification technique). The PCA can be a good tool for risk management purposes; however, it is weak at transparently explaining stock returns (i.e. with well identified explanatory factors). In addition, the PCA relies on the heavy assumption that stock returns are Gaussian.

Indeed, if the PCA is a straightforward statistical method used to decompose stock returns into independent factors (in the sense of correlation), it can only be valid if used with a strong hypothesis of normality. However, the normality of stock returns has been questioned for quite some time (as has been shown in a basic study, in Section 3.2, the stock return distributions are fat-tailed, and basic normality tests reject the hypothesis of normality). In the next section,

Figure 3.9: Mean Alpha vs. Number of PCA factors selected



an alternative purely statistical technique is presented that has drawn increasing attention in the financial world: the Independent Component Analysis (ICA).

3.7 Independent Component Analysis (ICA)

The ICA is a neuronal network technique, initially developed to filter noise from signals in a context where the signals are considered to be strictly non-Gaussian data. The ICA looks for statistically independent components and uses more information from the data than the PCA which, assuming normality, only considers the first two empirical moments of the multivariate dataset. Several attempts have been made to apply the ICA to financial datasets, and in particular, the study by Back & Weigend (1997) is discussed.

3.7.1 ICA principle

The ICA assumes that the multivariate dataset of historical asset returns M can be decomposed in the following way:

$$\mathbf{M} = \mathbf{SA}$$

Where $\mathbf{S}$ is a $(T-1) * N$ matrix (same dimensions as $\mathbf{M}$) of statistically independent signals (or components) and $\mathbf{A}$ is a $N * N$ mixing matrix, and stands for the loading of the asset returns on the independent components.
The ICA assumes that $\mathbf{M}$ is a multivariate non-Gaussian matrix that can be decomposed into a linear combination of statistically independent variables. In order to define $\mathbf{S}$, an algorithm capable of constructing independent signals from the original dataset $\mathbf{X}$ is needed. Because $\mathbf{S}$ and $\mathbf{A}$ are unknown, the ICA assumes the previous equality holds, and looks for independent components one by one. The strong form of independence states that two variables are independent, if and only if, their joint distribution is the product of their marginal densities (a consequence is that the correlation of the two variables is null; the equivalence holds if the variables considered are Gaussian). As the matrix $\mathbf{S}$ and resultant multivariate distribution are unknown, a statistical definition of independence is required.
The ICA is a stepwise procedure, that, one by one, looks for the non-Gaussian independent components $\mathbf{s_1}, ..., \mathbf{s_N}$ of $\mathbf{S}$; thus:

$$\mathbf{S} = \mathbf{XW}$$

Therefore, the aim of the ICA is to define, by successional process, the columns of $\mathbf{W}$: $(\mathbf{w_1}, ..., \mathbf{w_N})$. It can then be deduced: $A \equiv \mathbf{W}^{-1}$. The idea is to find $\mathbf{w_1}$, insofar that it maximises the non-Gaussianity of $\mathbf{Mw_1} \equiv \mathbf{s_1}$ (therefore, the importance of the non-Gaussian hypothesis, is that it stands as the first selection criterion to define the independent components). Then, $\mathbf{w_2}$ is defined insofar as it maximises the non-Gaussianity of $\mathbf{Mw_2} \equiv \mathbf{s_2}$, and, such that $\mathbf{s_1}$ and $\mathbf{s_2}$ are statistically independent (to make sure $\mathbf{s_2}$ is defined differently from $\mathbf{s_1}$). Then, the algorithm is reiterated until the Nth signal is defined.

3.7.2 ICA Fast Algorithm

Any ICA algorithm needs to define statistical measures for non-Gaussianity and independence. The "FastICA algorithm" introduced below, focuses on negentropy and mutual information.

3.7.2.1 Negentropy: a measure of non-Gaussianity

The FastICA algorithm developed by Hyvrinen & Oja (2000) is used for empirical testing. Statistical tools are required to evaluate non-Gaussianity, and statistical independence. The FastICA algorithm uses negentropy as a measure of non-Gaussianity. The negentropy exploit's the fact that a Gaussian variable has the largest entropy among all variables of equal variance. Therefore the negentropy of a variable y, is defined as the difference between the entropy of the Gaussian variable $\mathbf{y_g}$, defined as the Gaussian variable with same mean and variance as the variable $\mathbf{y}$, and the entropy of $\mathbf{y}$. The empirical entropy measure $H(.)$, is defined as:

$$H(\mathbf{y}) = -\sum_i \mathbb{P}(\mathbf{y} = a_i) \log(P(\mathbf{y} = a_i))$$

Where (a_i) defines the set of possible values for $\mathbf{y}$. Therefore, the empirical negentropy measure J of $\mathbf{y}$ is defined as:

$$J(\mathbf{y}) = H(\mathbf{y_g}) - H(\mathbf{y})$$

3.7.2.2 Mutual information

The mutual information takes into account the whole dependence structure of a dataset, unlike the PCA that only considers covariances. The Mutual information I of a set of variables $(\mathbf{y_1}, ..., \mathbf{y_n})$, is defined as:

$$I(\mathbf{y_1}, ..., \mathbf{y_n}) = \sum_{i=1}^{n} H(\mathbf{y_i}) - H(\mathbf{y_1}, ..., \mathbf{y_n})$$

where $H(\mathbf{y_1}, ..., \mathbf{y_n})$ is defined as the mutual entropy of the variables $((\mathrm{y}_1, ..., \mathbf{y_n}))$:

$$H(\mathbf{y_1}, ..., \mathbf{y_n}) = -\sum_{i_1} ... \sum_{i_n} \mathbb{P}(\mathbf{y_1} = a_{i_1}, ..., \mathbf{y_n} = a_{i_n})) \log(P(\mathbf{y_1} = a_{i_1}, ..., \mathbf{y_n} = a_{i_n}))$$

In actual fact, it is proven that the minimisation of Mutual Information (and therefore maximisation of statistical independence) is equivalent to the maximisation of negentropy, and; therefore, non-Gaussianity.

FastICA uses efficient empirical measures to estimate mutual entropy (see Hyvrinen & Oja (2000) for details). To simplify the algorithmic process, the initial dataset $\mathbf{M}$ is centred (the mean of each variable is subtracted) and whitened (the transformed variables are uncorrelated with a variance of one).

It is important to note, that, unlike the PCA, where the principal components are sorted by their eigenvalues, the independent components have the same variance, and are randomly built; therefore, there is no unique way to sort them. A number of different procedures have been suggested to determine the order of the principal components. As Back & Weigend (1997) suggest, the individual components can be ranked according to their L_∞-norm, (i.e. the maximum coefficient for each independent component).

3.7.3 Estimation of the model

In this study, the ICA is performed using the MATLAB 7.5.0 "FastICA" routine developed by the Laboratory of Computer and Information Science (CIS) at the Helsinki University of Technology[1].

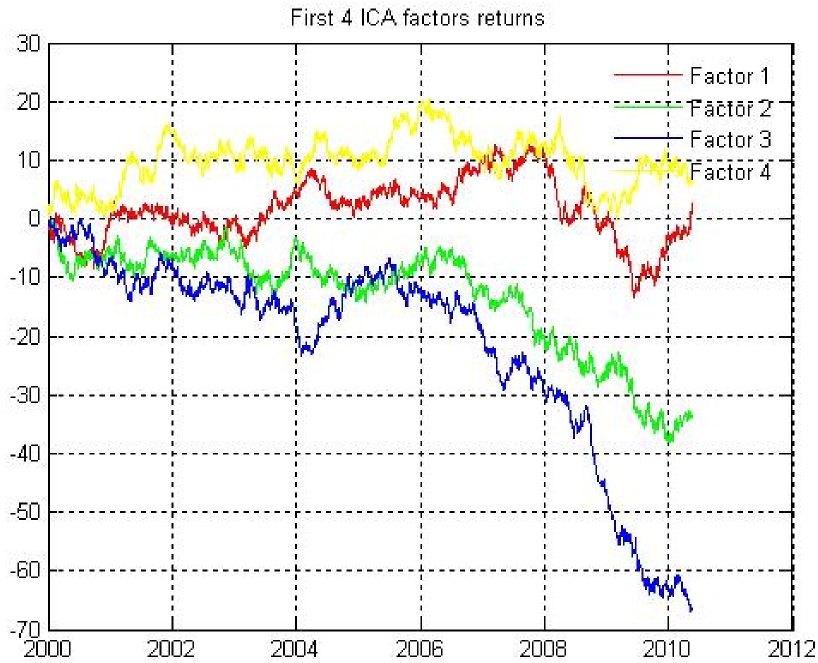
The algorithm is run, and the four components with the highest L_∞-norm are selected. Figure (3.10) plots those four factor returns:

For the PCA, the principal components can be identified by a computation of their returns correlation with the returns of exogenous factors. A similar procedure can be used to identify the ICA factors. Table (3.11) displays those correlations, and the correlation of the first four ICA factors with the PCA factors. As for the PCA, what matters is the amplitude of the correlation; rather, than its sign.

It can be seen that:

[1] http://www.cis.hut.fi/projects/ica/fastica/

Figure 3.10: ICA First 4 Factors Returns



- The first ICA factor is correlated with the first PCA factor and can also be considered as a "size" factor (expressing the fact that all the asset returns move in the same direction). This is corroborated by the fact that the first ICA factor is also much correlated with the SXXP Index (which can be assimilated to the market factor).
- It is difficult to identify any of the ICA factors with one of the external factors (volatility, oil or interest rate).
- The second ICA factor is highly correlated with the third PCA factor; however, it is difficult to identify it through the exogenous factors.
- The third and fourth ICA factors are also difficult to identify through this primary correlation analysis.

Table (3.12) shows the same statistics used for all the precedent models.

	ICA Factor 1	ICA Factor 2	ICA Factor 3	ICA Factor 4
SXXP Index	12.71	13.82	3.86	1.72
VDAX Index	−4.06	−8.90	4.26	−0.18
GDBR10 Comdty	3.30	2.95	2.01	0.70
CO1 Comdty	4.59	6.19	0.17	−2.23
PCA Factor 1	−14.65	−13.96	−1.56	−6.67
PCA Factor 2	4.71	5.24	0.60	8.11
PCA Factor 3	0.80	−9.61	−2.29	−2.00
PCA Factor 4	−0.26	−0.14	10.85	−0.55

Table 3.11: ICA Factors Correlation

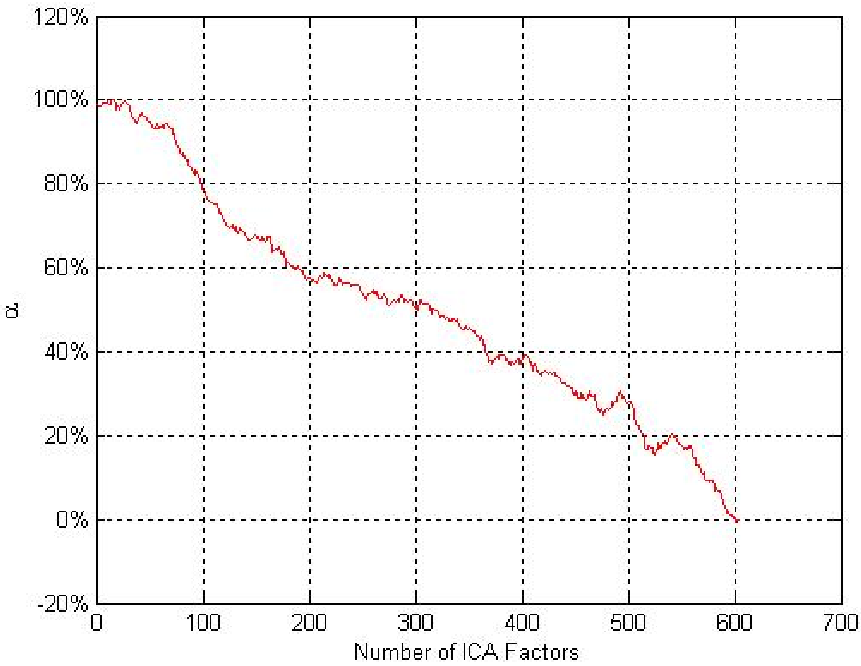
	SXXP Index	UKX Index	CAC Index	DAX Index
Mean $\beta^{F1}(\%)$	−0.02	0.01	−0.00	−0.01
Mean $\beta^{F2}(\%)$	−0.01	0.20	0.19	0.13
Mean $\beta^{F3}(\%)$	0.00	0.14	0.27	0.15
Mean $\beta^{F4}(\%)$	0.08	0.04	−0.15	0.18
Mean Residuals autocorr (%)	−2.66	−1.43	−0.66	−2.05
Pct infinity norm ICA	2.22	24.84	37.42	11.44
Mean $\alpha(\%)$	99.95	96.39	93.29	97.93

Table 3.12: ICA Statistics

The mean Alpha is very high for all the indexes considered. The percentage of infinity norm explained is not very significant, as all the principal components explain the same variance. As can be seen in Figure (3.11), at least 200 principal components are needed to reduce the mean Alpha to 60%.
As is the case for the PCA, the ICA components are not easy to identify. They can be useful for the detection of hidden factors in asset returns among exogenous factors already selected by the investor. However, it has been demonstrated that the first four ICA factors are difficult assimilate to any of the three external factors that have been considered previously (volatility, oil or interest rate).

The ICA and the PCA can be useful and interesting additional tools to explain stock returns; however, as stand alone models, they lack traceability. The next section presents a statistical model that uses some prior information and can

Figure 3.11: ICA Alphas vs. Number of Independent Components



therefore be more traceable than the PCA or ICA techniques.

3.8 Clusters Analysis (CA)

In this section, another purely statistical technique is presented, where the estimation algorithm was specifically developped in this research. The technique allows the investor to group the stocks into clusters. Furthermore, it can be a mix between a pure statistical model and an exogenous model, because prior clusters can be chosen (based, for instance, on industry sectorisation) to run the cluster algorithm. Below, the CA principle and empirical results are presented.

3.8.1 Cluster Analysis principle

The idea is that the returns of a stock behave similarly to the returns of its "brothers": stock returns belonging to the same sector tend to display similar patterns. The reversion assumption is that if the returns of a stock shift away from its brothers's (i.e. the stocks within a common cluster), they tend to revert back to the overall cluster returns behaviour.

3.8.2 Methodology

The matrix of stock returns $\mathbf{M}$ is considered. Initially, C prior original clusters are considered. Each stock is assigned to one of those C predefined clusters. The return matrix $\mathbf{M}$ is adjusted by decaying the returns with respect to time. Indeed, greater relative weight to more recent returns and less relative weight to older returns is desired (to calibrate the decay, a half-life of 20 days, or approximately, one month is considered). Therefore, work is undertaken on the adjusted matrix $\tilde{\mathbf{M}}$, defined as:

$$\tilde{\mathbf{M}} = (\exp^{-\lambda(T-t)r_t^i})_{1 \leq i \leq N, 1 \leq t \leq T}$$

where $\lambda \equiv \frac{\log 2}{20}$[1]. Then, the cluster centroid returns are computed for each of the original clusters. The centroid returns series is defined as the average of the cluster members returns. The distance of each stock is computed to each of the different centroids, and stocks are reassigned to the closest centroid. The procedure is repeated until the membership clustering remains invariant; thus, at each iteration the following steps are proceeded:

- The centroid returns series for each of the clusters is computed,
- The distance of all the stocks to each centroid is computed,
- Each stock is assigned to its closest centroid; therefore, redefining the clusters membership.

If the resulting clustering is invariant when compared to the initial clustering, the algorithm is stopped. In order to ensure that the algorithm does eventually stop, and to exclude outliers, a maximum distance to the centroid is set. The centroid variance is defined as the mean distance of all the cluster members to the cluster centroid. At a given step, the stocks of a given cluster that have a distance to the

[1] For the half life to be equal to 20, the weight of the 20th observation is required be $\frac{1}{2}$; therefore, it is required that: $\exp^{-20\lambda} = \frac{1}{2}$.

centroid greater than ten times the mean distance to the centroid of this specific cluster are excluded.
As a natural measure of distance, the dispersion of a stock return against the centroid returns is considered. If $\mathbf{r^c} \equiv (\tilde{r_t^c})_{1\leq t\leq T}$ denotes the centroid decayed returns of the cluster c, the distance of a stock i toward this centroid can be defined as:

$$\delta(i,c) \equiv 1 - \frac{\mathbf{r^c}'\mathbf{r^i}}{\sqrt{\mathbf{r^c}'\mathbf{r^c}}\sqrt{\mathbf{r^i}'\mathbf{r^i}}}$$

The assumption made in a cluster reversion model, is that each cluster member returns must revert to the centroid cluster returns. The expected return of a stock should therefore be its centroid return. The Alpha is defined as the discrepancy between each stock return and its cluster centroid returns.

3.8.3 Estimation of the model

A clustering algorithm, has been developed in MATLAB 7.5.0 for the purpose of this analysis, and allows clusters to be created from a prior clustering defined as sector membership. The Bloomberg industry sectorisation (ten industry sectors, therefore $C = 10$) is used to initialise the clusters, see Table (3.13):

	Number of stocks	Sector Dispersion (%)
Basic Materials	50	41
Communications	52	51
Consumer, Cyclical	60	52
Consumer, Non-cyclical	106	47
Diversified	10	93
Energy	38	49
Financial	137	46
Industrial	101	55
Technology	24	42
Utilities	28	44

Table 3.13: Sectors

The clusters seem to remain very similar to the original sectorisation, as the correlation between a sector and a corresponding cluster is always above 85%.

The table below displays the mean dispersion for each cluster. As can be seen, the dispersion of the clusters is much lower than the dispersion of the initial sectors.

	Number of members	Cluster Dispersion (%)
Cluster 1	30	15
Cluster 2	68	15
Cluster 3	90	18
Cluster 4	79	28
Cluster 5	94	19
Cluster 6	37	15
Cluster 7	73	13
Cluster 8	73	19
Cluster 9	20	23
Cluster 10	41	18

Table 3.14: Clusters Dispersion

A clustering distance measure is used in order to compare the clustering solution with the initial sectorisation. The clustering distance between two clustering solutions, with respective membership 1 and 2, is computed as follows: For each couple of assets i and j ($1 \leq N$, $1 \leq j \leq N$ and $i \neq j$) the value 0 is affected if i and j belong to a same cluster in either membership 1 or 2, or, if i and j belong to a different cluster in either membership 1 or 2. Otherwise, the value 0 is affected to the couple (i, j). The resulting sum of all the asset couple values is divided by the sum found in the case where the two clustering solutions are the same (i.e. membership 1 and 2 are the same), i.e. $\frac{N(N-1)}{2}$.
It is clear to see that the clusters are more equally constituted than the sectors (whereas, the smallest sector has only 10 members, and the largest, 137, the smallest and largest cluster has 20 and 94 members respectively). The dispersion of the clusters ranges between 13% and 28%, whereas, the dispersion of the initial sectors ranges from 42% to 93%. As mentioned above, the assumption made in a cluster analysis, is that member returns of a cluster revert to the centroid returns. Therefore, the cluster is not exactly a factor model, as factors and Betas are not extracted from the procedure. To compute the Alphas of the cluster model, a different cluster analysis is run on a rolling window of 100 days:

- For each day $t > 100$, the cluster analysis is computed for the returns series between date $t - window$ and $t - 1$.
- The return for date t is computed for the centroid of each sector c, denoted as r_t^c.
- For each member i of the cluster c, the Alpha is computed as: $\alpha_t^i = r_t^i - r_t^c$.

Table (3.15) displays the mean Alphas and clustering distances for the sector clustering of each of the four indexes studied:

	SXXP Index	UKX Index	CAC Index	DAX Index
Mean $\alpha(\%)$	28.83	-60.01	-54.24	-89.65
Mean Clustering Distance(%)	20.46	11.64	15.23	18.63

Table 3.15: CA Statistics

The average Alphas (between 30 and 40%) are preferable to the PCA or ICA average Alphas. This is probably due to the fact that prior knowledge of industry sectorisation has been used to initialise the search for reliable clusters. In fact, the amount by which the clusters differ on account of initial sectorisation is between 13% to 20%, which essentially means that the prior clustering has a significant impact on the final clustering obtained through the algorithm. The CA Alphas are also preferable to those of the CAPM or the EFM. On average, only the FFM Alphas are smaller than the CA Alphas; which is most probably due to the fact that the FFM requires a larger quantity of data than the other models (a value, per factor per stock, rather than just factor returns).
It is commonly held knowledge within the finance industry, that stock returns tend to behave similarly within a given sector. In the case of the CA, prior knowledge allows to use more efficiently pure statistical approaches.

3.9 Conclusions

This chapter has presented a number of models that are commonly used in finance to explain stock returns. For practical reasons, the choice has been made to stick to simple, tractable models, as the main purpose is to provide a broad overview of factor models, and a simple methodology utilising straightforward metrics (similar to the statistics used by Jensen (1969)) to help investors discriminate between different factor models).

Figure (3.12) plots the theoretical returns of the six different factor models studied in this chapter (CAPM, EFM, FFM, PCA, ICA and CA); a benchmark portfolio (the Equally Weighted portfolio EW, as considered in DeMiguel *et al.* (2007)); and the classical portfolios presented in Chapter 2 (the minimum variance portfolio MN and the mean variance portfolio MV). The theoretical returns are computed below:

$$r_t^F = \sum_{i=1}^{N} \alpha_t^{i,F} r_t^i$$

where $\alpha_t^{i,F}$ is the Alpha computed for the Factor Model F.

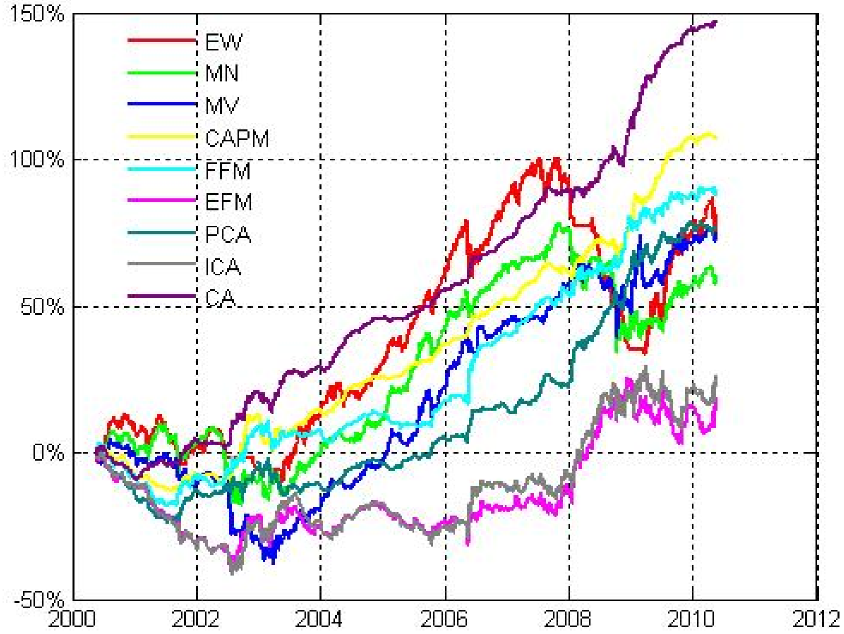
The performance of the Factor Models varies greatly over time. None of the models can be considered to perform the best over the whole time-horizon considered. This can be seen in Table (6.1), where details are given of the annual Sharpe ratios for the different models concerned (the three best Sharpe ratios per year are indicated in bold).

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
EW	**1.54**	−0.58	−1.13	**1.76**	2.45	1.57	0.42	−3.69	**1.98**	−0.44
MN	**0.75**	**0.00**	−0.77	0.75	**2.82**	2.49	1.52	−1.51	0.98	−0.04
MV	**0.01**	−1.28	−1.30	**0.87**	2.46	2.10	1.57	−0.42	1.28	**0.23**
CAPM	−1.88	−0.01	**2.29**	0.18	**3.89**	**4.70**	**2.35**	1.45	**3.51**	−0.57
FFM	−2.07	**0.17**	**2.03**	−0.13	1.21	**2.63**	**2.08**	**1.82**	0.82	−0.28
EFM	−2.15	−1.88	0.84	−0.32	0.18	0.36	0.55	**1.75**	−0.45	0.10
PCA	−3.55	−0.09	1.06	−0.84	2.36	2.26	1.52	**3.93**	1.40	−0.26
ICA	−2.11	−1.97	0.63	−0.05	−0.12	1.15	0.28	1.32	0.06	**0.52**
CA	−2.47	**2.18**	**1.90**	**1.07**	**4.36**	**4.42**	**3.85**	2.02	**4.10**	**1.45**

Table 3.16: Sharpe per Strategy per Period no Transaction Costs

The factor models that have been considered, are only useful if an investor possesses the foresight to be able to choose the most appropriate factors relative to

Figure 3.12: Factor Models Cumulative Returns (%)



the investment period (the factors that are required to construct a useful model may vary greatly over time). In this chapter, an overview has been given of the main types of factor models used in contemporary financial industries, which have been tested on the same daily dataset of European stock returns. In addition, the statistics introduced in an early CAPM testing paper by Jensen have been drawn upon. The intention of this research, rather than to find the best model, has been to provide a methodology that can be applied to compare the models considered. All the models tested, are of course, simplified models that can be refined further (e.g. introducing dynamic components with nonlinear methods - rather, than the linear regressions used - and increasing the number of factors). It is important to note that: because no single model appears to be able to constantly outperform any of the others, it is not possible to select one single model. Instead, one has to consider different models; which raises the fundamental questions on model risk and optimal and dynamic model mixtures, that are addressed in the remainder of this PhD thesis.

Chapter 4

Decision Under Ambiguity: Literature Review

"*But Uncertainty must be taken in a sense radically distinct from the familiar notion of Risk, from which it has never been properly separated. [...] It will appear that a measurable uncertainty, or "risk" proper, as we shall use the term, is so far different from an unmeasurable one that it is not in effect an uncertainty at all. [...] It is this "true" uncertainty, and not risk, as has been argued, which forms the basis of a valid theory of profit and accounts for the divergence between actual and theoretical competition.*"

F.H. Knight, "Risk, Uncertainty, and Profit", Hart, Schaffner & Marx, Houghton Mifflin Co., Boston, MA, 1921.

As presented in the previous chapter, many different models can be considered to represent asset return dynamics. The performance of such models varies greatly over time and investors cannot rely unconditionally on any of those. This chapter will describe the notion of ambiguity in greater detail. In the first section, the theoretical framework behind ambiguity will be discussed by reference to key literature. In particular, the fundamental distinction between risk and uncertainty will be explained, and examples discussed to demonstrate how decision makers adapt their decisions according to aversion toward both risk and ambiguity. In the second section, a brief overview of key approaches developed to solve decision problems under ambiguity will be outlined, with a focus on models in literature that engage problems concerning asset allocation under ambiguity. Finally, the last section will refer to recent studies and analysis that demonstrate the weaknesses of those contemporary models to perform well on real data. Therefore, this chapter will argue that financial investors require new approaches to solve the asset allocation problem under ambiguity.

4.1 The concept of Ambiguity

When under ambiguity, decision makers are prevented from forming beliefs with confidence due to a lack of reliable information. In this section, the theoretical fundaments of ambiguity will be outlined. Knight (1921) is the first to formally describe ambiguity. Later, Ellsberg (1961) illustrated the Knightian uncertainty with the famous Ellsberg Paradox. More recently, Kahneman and Tversky (1979) have formalised the impact of ambiguity in the decision making process.

4.1.1 The Knight Uncertainty

Knight (1921) is the first to formally discuss the intuition behind ambiguity. According to Knight, there is a significant difference between risk (an agent is uncertain about the precise outcome of a gamble, despite certainty concerning the distribution measure determining the set of possible outcomes) and ambiguity or model uncertainty (an agent is uncertain of the distribution measure). Ambiguity corresponds therefore to a non-measurable randomness against which decision makers show some aversion. Knight explains how an economic agent distinguishes between the estimates of their outputs and the degree of confidence they have in their estimates: "The action which follows upon an opinion depends as much upon the amount of confidence in that opinion as it does upon the favourableness of the opinion itself"[1]. Although not formally modelled, Knight pinpoints the significance of opinions of certainty concerning the future outcome of decision

[1] Knight (1921)

variables. It was some forty years later that Ellsberg clearly illustrated the concept of ambiguity through the Ellsberg Paradox; which will be discussed in the next section in reference to the classical settings developed by Savage (1954).

4.1.2 Savage Subjective Expected Utility (1954)

In a classical unambiguous setting, the decision maker is assumed to be certain about the distribution of $\mathbb{P} \equiv (p_1, .., p_n)$ upon the set of possible outcomes $(x_1, ..x_n)$. In that case, the Von Neumann-Morgenstern Expected Utility (EU) paradigm states that the decision maker intends to optimise the following program:

$$\max \mathbb{E}_{\mathbb{P}}[u(x, \lambda)] \tag{4.1}$$

where x represents the random outcome of the gamble played by the decision maker, u is a classical utility function and λ represents the agent's risk aversion. In practice, it is impossible for a decision maker to comprehend the true probability $\mathbb{P}$. Savage (1954) proposes a method to account for the subjective estimations made by decision makers to model the true distribution $\mathbb{P} \equiv (p_1, ..., p_n)$ of the outcomes $(x_1, ..., x_n)$. Savage introduces the concept of Subjective Expected Utility (SEU), to account for the fact that decision makers do not know the true probability $\mathbb{P}$ and may work with several priors $\mathbb{Q}$ to model the distribution of outcomes. In the SEU settings, decision makers are probabilistically sophisticated and form subjective beliefs based upon on a set of prior probability measures $\mathcal{Q}$ on $\Omega \equiv (\omega_1, ..., \omega_n)$. In this context, what differentiates investors is their risk aversion λ and a distribution measure π based upon a set of priors $\mathcal{Q}$. Thus, the decision maker maximizes a linear average over the different priors:

$$\max \sum_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_{\mathbb{Q}}[u(x, \lambda)]\pi(\mathbb{Q}) \tag{4.2}$$

In this setting, the decision maker is uncertain about the true probability $\mathbb{P}$, however, there is no ambiguity about the set of priors and their probability to occur. The decision maker assumes the distribution π used for the set of priors is unambiguous. The priors are uncertain but unambiguous; therefore the decision maker is not exposed to ambiguity.

For example: assume that a decision maker has two priors $\mathbb{Q}^1$ and $\mathbb{Q}^2$ to model the true distribution $\mathbb{P}$. Assuming that the decision maker's subjective weights (i.e. the probability that either $\mathbb{Q}^1$ or $\mathbb{Q}^2$ is the true model) are ϵ and $1-\epsilon$, and the control variable of this problem is ϕ (i.e. the agent decision determines the variable ϕ that affects the random outcome x^ϕ); the Problem (5.2) becomes[1]:

$$\max \epsilon \mathbb{E}_{\mathbb{Q}^1}[u(x^\phi, \lambda)] + (1-\epsilon)\mathbb{E}_{\mathbb{Q}^2}[u(x^\phi, \lambda)] \tag{4.3}$$

In virtue of the expectancy operator linearity, the obtained solution ϕ^* of Problem (4.3) is averagely weighted to the solutions ϕ^1 and ϕ^2 of Problem (5.1) applied to $\mathbb{Q}^1$ and $\mathbb{Q}^2$; obtaining:

$$\phi^* = \epsilon\phi^1 + (1-\epsilon)\phi^2$$

4.1.3 The Ellsberg Paradox

Ellsberg (1961) describes Knight's theory with simple gamble examples. Ellsberg demonstrates that decision makers show a more averse behaviour when betting on events represented by probabilities that are subjective and ambiguous. Therefore, he states that an ambiguity premium is added implicitly by decision makers to the risk premium which balances their risk aversion. In addition, Ellsberg shows that under ambiguity, decision makers violate some of the axioms of decision under uncertainty developed by Savage (1954).
More formally, a gamble is defined as a contract that yields given outcomes x_i with probability p_i, where $\sum_i p_i = 1$. Therefore, considering three gambles **a**, **b** and **c**, and taking $\succ$ to be the order relation on the set of gambles; the four Savage axioms are defined as:

[1]Note that for the equality to hold, some regularity conditions must be respected for the function u, specifically continuity.

(P1) **Complete ordering of actions:** if $\mathbf{a} \succ \mathbf{b}$ and $\mathbf{b} \succ \mathbf{c}$ then $\mathbf{a} \succ \mathbf{c}$. If gamble **a** is preferred to gamble **b**, and gamble **b** is preferred to gamble **c**, then gamble **a** must be preferred to gamble **c** also.

(P2) **Sure thing principle:** if $\mathbf{a} \succ \mathbf{b}$ then $\mathbf{a}+\mathbf{c} \succ \mathbf{b}+\mathbf{c}$. If gamble **a** is preferred to gamble **b**, then any combination of gamble **a** and another gamble **c** must be preferred to gamble **b** combined with the same gamble **c**.

(P3) **Independence of probabilities and payoffs:** if $\mathbf{a} \succ \mathbf{b}$ then $\mathbf{a} + \mathbf{w} \succ \mathbf{b}+\mathbf{w}$, when **w** stands as a certain payoff. If gamble **a** is preferred to gamble **b**, then gamble **a** combined with the sure payoff **w** must be preferred to gamble **b** combined with the same sure payoff.

(P4) **Admissibility (or rejection of dominated actions):** if $\mathbf{a} \succ \mathbf{b}$, then gamble **a** cannot be preferred to gamble **b** by the decision maker.

Even before Savage exposes his axioms, Allais (1953) finds that decision makers tend to attribute excessive weight to outcomes that are considered certain, violating (P3). Indeed, even if the expected payoff of a gamble is larger than a given fixed payoff, decision makers tend to prefer the sure payoff (provided gamble values are not sufficiently large to induce decision maker preference). Ellsberg illustrates how decision makers violate many of the Savage axioms with a simple example recalled below:

Gain per a single draw	**a**	**b**	**c**		**d**	**e**	**f**
Red	100	0	0		0	100	100
Yellow	0	100	0		100	0	100
Black	0	0	100		100	100	0

Table 4.1: Ellsenberg's single-urn Paradox, Single Draw

To explain the example, the urn contains 30 red balls, and 60 additional balls that are in unknown proportions of black and yellow. One ball is randomly drawn from the urn, and the decision maker can gamble on the outcomes of the game. Table (4.1) displays the different gains a decision maker stands to make depending on the states of the draw:

- For gamble **a**, the decision maker receives 100 if a red ball is drawn and 0 otherwise.
- For gamble **b**, the decision maker receives 100 if a yellow ball is drawn and 0 otherwise.
- For gamble **c**, the decision maker receives 100 if a black ball is drawn and 0 otherwise.
- For gamble **d**, the decision maker receives 100 if either a yellow or a black ball is drawn and 0 otherwise.
- For gamble **e**, the decision maker receives 100 if either a red or a black ball is drawn and 0 otherwise.
- For gamble **f**, the decision maker receives 100 if either a yellow or a red ball is drawn and 0 otherwise.

Typically, **a** is preferred to **b** and **c** (where **b** and **c** are chosen indifferently) - it is noted $\mathbf{a} \succ \mathbf{b} \sim \mathbf{c}$, and **d** is preferred to **e** and **f** (where **e** and **f** are chosen indifferently)- noting $\mathbf{d} \succ \mathbf{e} \sim \mathbf{f}$. If a decision maker can only pick one colour to obtain a positive gain (case **a**, **b** or **c**), they pick the colour the proportion of which is known, however if a decision maker can pick two of the three colours to obtain a positive gain (case **d**, **e** and **f**), they pick the colours in unknown proportions (i.e. the couple Yellow/Black). This breaches the Savage axiom (P2) that the utility of an act is an additively separable function of the consequences it yields in different states of the world:
$\mathbf{a} \succ \mathbf{b}$ and $\mathbf{b} \sim \mathbf{c} \;\not\Rightarrow \mathbf{e}(= \mathbf{a} \cup \mathbf{c}) \succ \mathbf{d}(= \mathbf{b} \cup \mathbf{c})$ or $\mathbf{f}(= \mathbf{a} \cup \mathbf{b}) \sim \mathbf{d}(= \mathbf{a} \cup \mathbf{c})$
Such preferences are inconsistent with choices made based on rational probabilities. Indeed, if it is assumed that there were a probability measure $\mathbb{P}$ underlying these choices, then $\mathbf{a} \succ \mathbf{b}$ implies $\mathbb{P}(Red) > \mathbb{P}(Black)$, while $\mathbf{d} \succ \mathbf{f}$ implies $\mathbb{P}(Black) + \mathbb{P}(Yellow) > \mathbb{P}(Red) + \mathbb{P}(Yellow)$[1], hence revealing a contradiction. This paradox describes the attitude of ambiguity aversion displayed by decision makers: **a** is preferred to **b** and **c** because **b** and **c** represent ambiguous states. The

[1]Where $\mathbb{P}(Red)$, $\mathbb{P}(Yellow)$, $\mathbb{P}(Black)$ denote respectively the probability that a red, yellow or black ball is drawn.

state **d** is preferred to the state **e** or **f** because it is not ambiguous: the probability to draw a black or a yellow ball is equal to $\frac{60}{90}$, whereas the probability to draw either a red or a black ball or a red or a yellow ball is unknown. The Ellsberg paradox highlights the aversion to ambiguity: decision makers prefer objective probabilities (the proportion of red balls in the urn is known) in comparison to subjective probabilities where it becomes necessary to infer from incomplete information (the proportion of combined black and yellow balls is known, but the exact proportion of black and/or yellow balls is not).
In this case: "The Bayesian or Savage approach give wrong predictions [...]" (see Ellsberg (1961)). Ellsberg demonstrates that decision makers take into account a degree of ambiguity aversion within their decision-making processes.

4.1.4 Kahneman and Tversky Prospect Theory (1979)

Following the work by Ellsberg and Savage, Kahneman & Tversky (1979) develop a modelling framework (or prospect theory) for decision making under ambiguity (for an analysis of prospect theory see also the paper by Wu & Gonzalez (1999)): the central principle is that for decisions under ambiguity, the weighting of a decision attached to a given event by a decision maker differs from the probability assigned to the event. More specifically, Kahneman and Tversky elaborate the following two-stage model under ambiguity, encompassing:

- the modelling of subjectively judged probabilities assigned by the investor to the different events;

- the modelling of decision probabilities (more conservative than the judged probability) that stand as a transformed version of the subjective probabilities through a function ψ to account for the decision maker's aversion to ambiguity.

Kahneman and Tversky introduce two "scales" to adapt the expected utility framework to decision under ambiguity: a weighting function π and a value function v:

- π associates a decision weight $\pi(p_i)$ to each probability p_i.

- v assigns to each outcome x_i a subjective value $v(x_i)$.

Typically, the value function v can be assimilated to an S-shaped function (concave for gains and convex for losses), and the weighting function π can be formalized as a nonlinear increasing function that over-estimates the weight of small probabilities and underestimates larger probabilities.
In 2002, Daniel Kahneman was awarded the Nobel Prize for his contribution to behavioural economics with his work on Prospect Theory, which has been considered a milestone in research into modelling decisions under ambiguity. The remainder of this chapter will present and discuss the contrasting models proposed in the literature for decision under ambiguity, focusing on applications in asset allocation problems.

4.2 Decision under Ambiguity

In this section, we introduce the concepts of Max-Min Utility and Robust Control that were the first proposed methods to account for ambiguity in decision making problems.

4.2.1 Gilboa and Schmeidler Max-Min Expected Utility (1989)

Gilboa & Schmeidler (1989) are the first to take into account aversion towards ambiguity, by using a max-min criterion for decision making under non-unique prior in the specific framework of risk measures. Instead of maximizing an expected utility, the agent takes a pessimistic view, minimising the maximum expected utility over the set of priors considered to model $\mathbb{P}$. By applying this methodology, the decision maker opts for the most conservative, worst-case scenario (i.e., the prior that leads to the minimum optimal expected utility). Gilboa and Schmeidler re-specify the decision problem under ambiguous priors as:

$$\min_{\mathbb{Q}\in\mathcal{Q}} \max_{\phi} \mathbb{E}_{\mathbb{Q}}[u(x^{\phi}, \lambda)] \tag{4.4}$$

The main drawback of this approach is that it effectively only considers the worst-case scenario, i.e. the prior under which the maximized expected utility of the gamble outcome is the lowest. Robust control theory addresses this issue and provides a less conservative methodology for decision making under ambiguity.

4.2.2 Hansen and Sargent Robust Control (2001)

Hansen and Sargent elaborate on the max-min expected utility of Gilboa and Schmeidler considering robust control theory. Their idea is to refine the max-min principle by adding a penalty function α to the decision problem (4.4). The robust preference approach criterion penalises the different investor models $\mathbb{Q} \in Q$[1] with respect to their difference to a reference model $\mathbb{P}$, defined as:

$$\max \min_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{\mathbb{Q}}[u(w, \lambda)] + \theta\alpha(\mathbb{Q})$$

An example for the penalty function is the relative entropy of each prior $\mathbb{Q}$ with respect to $\mathbb{P}$:

$$\alpha(\mathbb{Q}) = \mathbb{E}_{\mathbb{Q}}[\log(\frac{d\mathbb{Q}}{d\mathbb{P}})]$$

The robustness parameter θ is interpreted as an implicit Lagrange multiplier on the specification error $\alpha(\mathbb{Q}) < \epsilon$, where the investor can set the tolerance ϵ (i.e. how far away a model can be from the reference model $\mathbb{P}$).

Note that the same principle is used in financial mathematics for convex risk measures recently introduced (see Föllmer & Schied (2002)) to assess and manage risky financial positions, when the true probability $\mathbb{P}$ is unknown. The idea is to create an acceptable investment position by determining the minimal amount of capital (or capital requirement) $\rho(x)$, that need to be added to the risky portfolio x in order to account for ambiguity aversion. The first form of risk measure is the

[1]We assume that $\mathcal{Q} \in \{\mathbb{Q} \in \mathcal{M}, \mathbb{Q} \prec\prec \mathbb{P}\}$, so that it is possible to define an entropy measure between $\mathbb{Q}$ and $\mathbb{P}$ for all $\mathbb{Q} \in \mathcal{Q}$.

coherent risk measure as introduced by Artzner *et al.* (1999); the idea of which is to compute the worst expected loss under the different priors considered:

$$\rho(x) = \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{\mathbb{Q}}(-x) \tag{4.6}$$

Föllmer & Schied (2002) extend the risk measure concept with a penalty function as in Hansen & Sargent (2001):

$$\rho(x) = \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{\mathbb{Q}}(-x) - \alpha(\mathbb{Q}) \tag{4.7}$$

However, the risk measure solely takes into account ambiguity aversion.
In the remainder of this Chapter, focus is given more precisely on decision making problems in finance. The ambiguity aversion is treated similarly in a context of portfolio optimisation, where risk aversion is also considered. The next section will focus specifically on the portfolio allocation problem under ambiguity.

4.3 Portfolio Allocation and Model Risk

Focusing on decision under ambiguity in problems of asset allocation, this section will build upon the examples specified in Chapter 2. To recall the problem considered: the investor wants to allocate the initial wealth x_0 among the different assets $N+1$ available in the market (N risky assets and one risk-free asset). In this case, x^ϕ represents the value of the investor's portfolio at a future time horizon, and the control variable ϕ represents the investor's strategy to initially allocate wealth among the assets. More precisely, ϕ is a vector of weights assigning a positive value for each proportion of wealth allocated to a given asset that is bought or a corresponding negative value if this asset is sold. Each element of ϕ belongs to $[-1 : 1]$. Note that for a more in depth literature review of asset allocation under model risk, one can refer to Fabozzi *et al.* (2007).

4.3.1 The classical Markowitz settings

As covered in Chapter 2, the portfolio optimisation problem in the case of Markowitz is as follows:

$$\max_{\phi} \mathbb{E}_{\mathbb{P}}[u(x^{\phi}, \lambda)] \tag{4.8}$$

Subject to some investment constraints, where u is a quadratic utility function parametrized by a risk aversion parameter λ[1].

Originally the problem is solved by treating $\mathbb{P}$ as a known probability. In practical examples where asset returns are assumed to be normally distributed[2], the sample mean $\boldsymbol{\mu}$ and sample covariance matrix $\boldsymbol{\Sigma}$ of observed asset returns are used to estimate the joint distribution of the first two moments of the asset returns. Because the asset return distribution is assumed to be normal, the knowledge of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ encompasses the knowledge of $\mathbb{P}$. The problem (4.8) therefore becomes:

$$\max_{\phi} \mu' x^{\phi} - \lambda \phi' \Sigma \phi \tag{4.9}$$

As demonstrated by contemporary research, the above treatment of the problem results in suboptimal portfolio choices. Indeed, from early on, many authors have challenged the performance of the Markowitz portfolios (see for instance Merton (1973), who points out the instability of the estimation of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ through the sample mean and sample covariance matrix of the asset returns). In practice the investor can only anticipate $\mathbb{P}$, therefore challenging the sustainability of the central assumption that all investors are mutually informed and in agreement about distribution $\mathbb{P}$. Indeed, it is necessary for the investor to infer the true probability $\mathbb{P}$ from historical data. The following section will draw upon the more recent approaches to ambiguity, taking into account investor uncertainty regarding the true probability $\mathbb{P}$.

The following section will outline proposed techniques to improve the robustness of mean and covariance asset returns estimation for model risk in portfolio op-

[1] $u(x) = x - \lambda x^2$

[2] Justifying the quadratic form of the utility function, as the first two moments entirely define the whole distribution of the asset returns.

timisation. The notion of model risk will then be applied to more generalised frameworks.

4.3.2 Learning and Filtering

Historically, model risk was first assimilated with parameterisation uncertainty. Assuming the asset returns belong to a class of parametrized distributions, the problem is to find robust estimators for the parameters. Stochastic filtering and Bayesian statistics provide methods to account for this parameter uncertainty. Assuming markets are complete, the investor learns from the information included in the observed asset returns and can update estimators accordingly. Over time, estimators eventually converge towards the real values of the asset returns distribution parameters.

4.3.2.1 Classical Learning: the Bayesian approach

Basak (2005) considers the case where different investors have different priors concerning the dynamics of financial assets. He uses the stochastic filtering theory to compute market equilibrium when agents disagree on the mean growth rate of asset returns in a context of a dynamic, complete financial market. The agents have heterogeneous beliefs about the dynamic endowment trend (or growth rate) of the risky financial asset (in this setting, $N = 1$). Under the true probability $\mathbb{P}$, the dynamics of the risky asset is:

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_t dw_t$$

Investors consider equivalent probabilities $\mathbb{P}^i$ also equivalent to $\mathbb{P}$. Each investor i considers the following process for the risky asset:

$$\frac{dS_t}{S_t} = \mu_t^i dt + \sigma_t dw_t^i$$

Where μ^i is the anticipated dynamic trend rate for the investor i (σ_t is deduced from the quadratic variations of S_t) and w^i is the investor's innovation process (such that the dynamic anticipated by each investor is coherent with the observed risky financial asset process). Investors update their beliefs in a Bayesian fashion:

$$\mu_t^i = \mathbb{E}_{\mathbb{P}^i}(\mu_t/F_t)$$

Where $(F_t) \equiv \sigma(S_s, 0 \leq s \leq t)$ is the information filtration generated by the observation of the asset prices.
The example of a Gaussian Filtering setting allows Basak to give close formula for the growth rate dynamics of the two agents and a disagreement process. Basak provides solutions for the instance that there is only one risky asset.
He also presents more complex settings adding a risk free asset and considering the case of dividend paying assets, testing scenarios generalised to several investors and several sources of risk. Under heterogeneous beliefs, risk is transferred from a more pessimistic to an optimistic investor. The transfer of risk is proportional to the extent of investor disagreement. The basis for the methodology proposed is a rational learning setting, where the learning model is correctly specified, and the agents have constant beliefs concerning modelled asset prices, despite disagreement.

Brennan & Xia (2001) offer a numerical solution to a similar dynamic portfolio allocation problem. In their problem, the investor does not fully believe in one model and uses a mixture of two different priors to model asset returns: a normal prior and a factor pricing model based normal prior (the factor model used is the CAPM). They find that the investor allocation between the market portfolio and the Fama-French SMB (Small capitalization Minus Big capitalization stocks) and HML (High book/price ratio Minus Low book/price ratio stocks) portfolios changes dramatically if the investor uses a mixture of priors (i.e., the investor is uncertain about the model to use to represent asset returns and therefore considers two different models). The trend process of the asset returns is modelled as follows:

$$\mathbb{E}_{\mathbb{P}}(\mu_t/F_t) = \pi_t \hat{\mu_t^1} + (1 - \pi_t)\hat{\mu_t^2}$$

Where π_t is the weight given to the CAPM model, and $\hat{\mu_t^1}$ is the estimated trend under the CAPM prior, and $\hat{\mu_t^2}$ is the estimated trend under a classical normal prior. Using a normal prior in conjunction with the CAPM model allows the

investor to consider some of the CAPM anomalies as genuine, similar to the example identified by Fama and French (their SMB and HML portfolio returns violate the CAPM model, as a Beta adjusted portfolio of small stocks tends to outperform a Beta adjusted portfolio of big stocks, and a Beta adjusted portfolio of low book to price ratio stocks tends to outperform a Beta adjusted portfolio of high book to price ratio stocks). Brennan and Xia demonstrate that model uncertainty (i.e. uncertainty about the authenticity of the SMB and HML anomaly) can have a major impact on portfolio choices. Empirically they found that if an investor does not fully trust the CAPM, they would reduce investment in the market portfolio (the "Beta" portfolio of the CAPM) and take long positions in the SMB and HML portfolios (outweighing small stocks and low book to price ratios stocks).
Following Brennan & Xia (2001), Cvitanić *et al.* (2006) give a close form solution for a dynamic portfolio choice problem when the investor detects abnormal returns as deviations from an asset pricing model used as a prior (here the dynamic version of the CAPM). The authors consider the market portfolio and N single assets as normal assets. The investor is unaware of the trend process of the assets, dynamically updating priors in a Bayesian fashion. To account for anomalies in the CAPM model, the expected return of an asset j is modelled as:

$$\mathbb{E}_{\mathbb{P}}(\mu_t^j / F_t) = r + \beta^j(\mu_t^M - r) + \alpha^j$$

where μ^M is the market trend process, r is the risk free rate, β^j is the asset j CAPM beta with respect to the market portfolio and α^j accounts for the abnormal or idiosyncratic return of the stock j. The authors use changes in analysts recommendations as an estimate for the Alphas.

However, such models use classical Bayesian updating techniques assuming the parameterisation family to which the distribution $\mathbb{P}$ belongs to is comprehended by the investor. Although the investor is uncertain about the parameters, there is no model ambiguity per se as the Bayesian updating procedure assumes that if the investor has enough observational data, the estimated distribution will ultimately converge towards the true one.

4.3.2.2 Learning Under Ambiguity

In this section, we present the more recent research by Epstein & Schneider (2008), considering an alternative to classical Bayesian updating, when there exists some ambiguity. The authors argue that ambiguity radically transforms results attributable to classical Bayesian updating rules. The agents, when updating beliefs under ambiguous information form attitudes that can be divided into three distinct types:

- Ambiguity Aversion: investor shows ambiguity adverse behaviour, maximising their utility under a worst-case scenario.
- Asymmetric Behaviour: investor displays asymmetric behaviour toward ambiguity. Under ambiguity, bad news affects conditional actions to greater extent than good news.
- Ambiguity Anticipation: investor reduces consumption of particular assets associated with information that is expected to be ambiguous.

The central idea proposed by the theoretical framework developed by Epstein & Schneider (2008), is that in order to model ambiguous information, an investor when confronted with difficulties in the judgement of signals quality, treats those signals as ambiguous. Instead of updating beliefs in standard Bayesian fashion, the investor considers a number of likely outcomes when interpreting the signals. Epstein and Schneider propose the example of the noisy news signal about the dividend s of a given risky asset:

$$s = \theta + \epsilon$$

where θ is the true information and ϵ is an ambiguous noise distributed as $N(0, \sigma_s^2)$, with $\sigma_s^2 \in [\underline{\sigma_s^2}; \overline{\sigma_s^2}]$.

In the pricing theory developed by Epstein and Schneider, the greater influence of bad news on asset returns requires that market participants be compensated for enduring periods of ambiguous news. Epstein and Schneider build a pricing model of financial assets being appraised by their discounted future dividend values where in every period, the agents observe an ambiguous signal of the next period dividend. The stock price today must be equal to the worst-case conditional expectation of the discounted value of all future ambiguous dividends.

Epstein and Schneider conclude that ambiguity lowers the mean return of a portfolio of assets, no matter how many assets constitute the portfolio. On the contrary to pure risk where a diversification phenomenon can take place, the market portfolio does not become less uncertain with an increased number of assets. The news communication process introduces a permanent ambiguity into beliefs about the fundamentals (here the dividends).

The pricing model under ambiguity developed by Epstein & Schneider (2008) is theoretical and deals with a unique risky asset. Models developed by Basak (2005), Brennan & Xia (2001) or Cvitanić *et al.* (2006), are very theoretical as well. More practical models are needed to solve large asset portfolio allocation problems under ambiguity. In the next section, more practical models considering large portfolios of assets are discussed.

Empirically, it has been already demonstrated that the Makrowitz model performs badly (see for instance Merton (1973)). This is due to the fact that the estimation of the mean vector and covariance matrix of a large number of asset returns is often unstable. Also, in the case where the investor considers several models, we have seen that the Gilboa-Schmeilder min-max paradigm forces the investor to consider only the worst case. In order to still take into account the different priors the investor considers, and to overcome the major shortfall of mean and covariance estimation instability, several penalization procedures have been proposed. The different priors are penalized with respect to a defined distance towards a reference model. Unlike the traditional approach, where inputs to the portfolio framework are treated as deterministic (especially the parameters of the parameterized distribution of asset returns), robust portfolio optimization incorporates the notion that inputs have been estimated with errors, and therefore constraint that they should lay in a reasonable interval. The resulting robust portfolio allocation tend to be more stable and less sensitive to small changes in model parameters.

4.3.2.3 The shrinkage approach

The idea behind the shrinkage approach is to use prior knowledge in order to make the portfolio allocation more robust (i.e. less dependant on estimation variations). Black & Litterman (1990) introduced in the classical Markowitz settings the option for the investor to specify some "view" on the asset returns, which effectively boils down to specifying a prior for the asset returns mean vector. More generally speaking, the shrinkage approach proposes to shrink the sample mean toward a prior value, that enhances the robustness of the estimator. In practice, this means estimating the mean vector $\boldsymbol{\mu}$ as a weighted average of the sample mean $\bar{\boldsymbol{\mu}}$ and a prior value. Jorion (1986) uses the Bayes Stein estimator and shrinks the sample mean toward the minimum variance portfolio mean. Pastor (2000) proposes to shrink the unconstrained sample mean toward a mean constrained by a prior model. As an example he uses the CAPM as a prior model (and therefore the constrained mean is defined for each asset i as: $\forall i, \mu^i = \beta^i \mu^M$).

Wang (2005) elaborates on and integrates the shrinkage approach developed by Pastor and max-min optimisation as proposed by Gilboa & Schmeidler (1989). The asset model used is also the CAPM, hence, the mean of the asset returns is estimated as: $\mu = \alpha + \beta\mu^M$. Wang models the prior distribution of α conditional on the covariance matrix Σ as a normal centred distribution with variance proportional to the covariance matrix: $\theta\Sigma, \theta > 0$. Finally, Wang models the investor portfolio decision problem as:

$$\max_{\phi} \min_{\theta} \mathbb{E}[u(x^{\phi}, \theta)]$$

4.3.2.4 The Multiple Prior Approach

Garlappi *et al.* (2009) generalize the approach to multiple priors adding to the min-max optimization a constraint (in the spirit of Hansen & Sargent (2001)) on the parameters in order to relax the worst case scenario settings of Wang. Investors minimize their preferences only among priors that are close enough to the empirical sample estimators. They add a constraint so that they only consider models for which the implied mean is close enough to the empirical mean: in practice, they add to the problem above the following constraint:

$$f(\mu, \bar{\mu}, \bar{\Sigma}) < \epsilon$$

where $\bar{\mu}$ and $\bar{\Sigma}$ are the sample mean and variance and ϵ accounts for the investor model ambiguity aversion (the bigger ϵ, the more averse the investor is). f can be assimilated to a t-test statistics that tests if the model constraint mean is in the neighbourhood of the sample mean.
Note that f plays the role of a confidence interval: only priors for which expected returns are close enough to the empirical mean of asset returns are considered.
Along similar lines, Kogan *et al.* (2002) restrict the set of priors to only those that are close enough to the empirical data according to entropy measurements (the entropy between the empirical data and the priors considered must be smaller than a tolerance level ϵ); the investor solves the following problem:

$$\max_{\phi} \min_{\mathbb{Q} \in \mathcal{Q}_\epsilon} \mathbb{E}_{\mathbb{Q}}[u(X^\phi)]$$

where $\mathcal{Q}_\epsilon = \{\mathbb{Q} \in \mathcal{Q} : \mathbb{E}[\frac{d\mathbb{Q}}{d\hat{\mathbb{P}}} \ln \frac{d\mathbb{Q}}{d\hat{\mathbb{P}}}] < \epsilon\}$ and $\hat{\mathbb{P}}$ stands for the empirical distribution of observed asset returns.
In their recent paper Epstein & Schneider (2007) use a likelihood ratio test to constraint the portfolio optimization problem: only priors close enough to historical data in terms of likelihood ratio are considered.
All those models however are constrained by the choice of relevant reference priors. Some general models have been recently proposed and are presented in the following section.

4.3.3 Generalised framework to model ambiguity in the asset allocation problem

More recently, some authors have considered more generalised models that encompass the different frameworks proposed so far in the literature to account for model ambiguity in the portfolio optimisation problem.

4.3.3.1 A smooth class of preferences that models ambiguity

Maccheroni *et al.* (2006) generalise a smooth class of preferences dealing with ambiguity. They generalise the Gilboa-Schmeidler max-min model to asset allocation problems as they introduce a penalty function α:

$$\max_{\phi} \min_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_{\mathbb{Q}}[u(x^{\phi}, \lambda)] + \alpha(\mathbb{Q}) \tag{4.10}$$

The function α operates as a generalised penalty function that encompasses:

- the entropy penalty criterion used by Hansen *et al.* (2006), where the investor considers a unique prior $\mathbb{Q}$. $\alpha(\mathbb{Q}) \equiv \mathbb{E}_{\mathbb{Q}}(\log \frac{d\mathbb{Q}}{d\mathbb{P}})$ is the relative entropy of $\mathbb{Q}$ with respect to $\mathbb{P}$.

- the multiple prior models (for instance the one studied by Epstein & Schneider (2008)) where the investor considers several priors in a subset $\mathcal{Q}^*$ of the set $\mathcal{Q}$. $\alpha(\mathbb{Q}) \equiv 0$ if $\mathbb{Q} \in \mathcal{Q}^*$ otherwise $\alpha(\mathbb{Q}) \equiv +\infty$.

The bigger is $\alpha(\mathbb{Q})$, the higher the penalisation; and, therefore the greater the ambiguity aversion.

The next section will present a general theoretical method that take into account ambiguity developed by Klibanoff *et al.* (2005). This methodology is a generalisation of convex risk measures as developed by Föllmer & Schied (2002) and adapted to portfolio optimisation. Note that the novel Ambiguity Robust Adjustment methodology proposed in this PhD thesis will be benchmarked by this general model of asset allocation under ambiguity.

4.3.3.2 A Generalised Model

Klibanoff *et al.* (2005) propose a generalised methodology to take into account ambiguity. They extend the theory of coherent risk measure developed by Artzner *et al.* (1999) and elaborate on the model of value function proposed by Maccheroni *et al.* (2006). Klibanoff, Marinacci and Mukerji introduce a smooth function ϕ

characterising the investor ambiguity aversion by the parameter γ. Thus, the portfolio optimisation problem becomes:

$$\max_{\phi} \mathbb{E}_{\pi}\psi[\mathbb{E}_{\mathbb{Q}}[u(x^{\phi}, \lambda)], \gamma] \tag{4.11}$$

Note that the SEU portfolio allocation problem (5.2) is a particular case of (4.11) when the agent is not averse to ambiguity, where ψ is a linear function (subjective probability with an average weighting of the priors in $\mathcal{Q}$) and π represents the subjective distribution of the priors $\mathbb{Q} \in \mathcal{Q}$. $\mathbb{E}_{\pi}\psi[\mathbb{E}_{\mathbb{Q}}[u(X^{\phi}, \lambda)], \gamma]$ can be interpreted as the certainty equivalent of the ambiguous conditional expected utility $\mathbb{E}_{\mathbb{Q}}[u(X^{\phi}, \lambda)]$ under any prior model $\mathbb{Q} \in \mathcal{Q}$. The nonlinearity of ψ accounts for decision maker aversion to ambiguity.

The main drawback of the general methodology proposed by Klibanoff, Marinacci and Munkerji is that it is often impossible to solve without numerical methods, and furthermore, it can be difficult to use empirically (when the dimension of the decision variable is large). Additionally, the ambiguity aversion is not clearly identified: it is assumed the ambiguity aversion is the same against all the different priors. It does not make the distinction between the absolute ambiguity aversion an investor shows for a given prior and the overall ambiguity aversion that the investor displays to the set of priors considered.

4.4 The Poor Performance of Ambiguity Models on Empirical Data.

The problems encountered with the methods above stem from either an over-dependency upon parametric specification of priors and/or rely on too complex a process to be practically implemented. Models that utilise an increase in the number of priors and a refinement of the methods to control estimation errors result in high levels of noise, and ultimately poor realised performance of subsequent optimised portfolios.
In fact, DeMiguel *et al.* (2007) have conducted a thorough study comparing the performance of different portfolios. Monthly equity returns from 1952 to 1999

(provided by the website: Kenneth French website) are used to build and compare the performance of 14 different portfolios (the naive equally weighted portfolio, where the asset weights are all equal to $\frac{1}{N}$; the sample mean-variance portfolio; the Bayesian estimated mean variance portfolio; the Bayes-Stein shrinkage estimated portfolio; the Data and Model Pastor portfolio; the minimum variance portfolio; the CAPM portfolio; the Garlappi-Uppal-Wang multiple prior portfolio; the short sell constrained portfolio, and a number of mixed variations of the above portfolios). They compute the out of sample means and standard deviations of the portfolio returns (denoted μ_k and σ_k), $\forall k \in [1, 14]$. They then use the following performance measures to compare the different portfolios:

- the Out of Sample Sharpe Ratio (mean over standard deviation of out of sample portfolio returns: $Sharpe_k \equiv \frac{\mu_k}{\sigma_k}$)
- the Certainty Equivalent Return of the different portfolio strategies (mean minus risk aversion adjusted variance): $CER_k \equiv \mu_k - \lambda\sigma_k^2$)
- the Turnover of the portfolio (i.e. the amount of shares traded due to portfolio rebalances when the weights are modified)

The main result of the study conducted by DeMiguel *et al.* (2007) is that: all the selected portfolios fail to significantly beat the performance of the simplistic equally weighted portfolio (that is denoted $1/N$, as it affects an equal weight to all the N risky assets considered). Only the minimum variance portfolio significantly out performs the equally weighted portfolio in terms of Sharpe; in terms of CEQ, only the mean-variance portfolio beats the $1/N$ portfolio; and in terms of Turnover, the equally weighted portfolio is by construction the best performer (with a null Turnover), as the allocation remains always the same for all risky assets. Indeed, the greater the uncertainty concerning the set of the constraints in a robust optimisation portfolio problem, the greater the chance that the resulting optimal portfolio will be conservative, and consequently a quantity of the potential portfolio performance will be sacrificed. In a robust portfolio optimisation, the investor trades off optimality against the risk of employing an inaccurate model. It can often prove very costly to over constrain the optimisation problem.

Mounting criticism of ambiguity aversion methodology (see for instance Al-Najjar & Weinstein (2009) who question whether the Ellsberg choices are rational responses to ambiguity, therefore contradicting the ambiguity-aversion postulate and many dependent ambiguity aversion methods) is questioning the theoretical and empirical limitations of accepted methods. Responding to Al-Najjar & Weinstein (2009) critics, Nehring (2009) argues that rational choice under ambiguity aims at robustness rather than an impossible avoidance of ambiguity. However, as Fabozzi *et al.* (2007) point out, robust decisions under ambiguity are often paid for by the poor performance of those models in practice: "By using robust portfolio optimization, investors are likely to trade off the optimality of their portfolio allocation in cases in which nature behaves as they predicted for protection against the risk of inaccurate estimation". Hence this thesis proposes a new methodology that avoids the use of penalization techniques; instead proposing modification of the outputs (i.e. the different asset allocations) of the models considered. The next chapter will present this novel methodology to account for model ambiguity that performs well empirically and outperforms existing models in simplicity of practical application.

Chapter 5

A Robust Alternative Approach to Model Ambiguity

"*Any financial model is by definition a simplified and thus imperfect representation of the economic world and the ways in which agents perform investment, trading or financing decisions under uncertainty.*"

R. Gibson, Model risk, RISK books, 2000:

This chapter constitutes the core contribution of this PhD thesis by proposing a new approach to account for model ambiguity aversion in the portfolio optimization problem. Our aim is to introduce a simple, practical and easily implementable approach to account for model risk in a robust way. Our motivation to propose a simple methodology to account for ambiguity aversion is essentially due to the complexity to solve the allocation problem under the settings proposed by Klibanoff *et al.* (2005) and presented in Chapter 4.
A two-step robust ambiguity methodology is introduced, which offers the advantages of greater tractability and easier implementation when compared with many of the various approaches proposed in the literature, and detailed in Chapter 4. This methodology decomposes ambiguity aversion into both a model specific absolute ambiguity aversion, and a relative ambiguity aversion across the set of

different prior models considered for the asset returns. First, the optimal allocations under each prior are transformed through a generic absolute ambiguity function ψ; then, the adjusted allocations are mixed through an adjustment function π that reflects the relative ambiguity aversion of the investor towards the different models.
The original approach proposed in this thesis is altogether more flexible, easier to compute, and more tractable than the one proposed in the literature. Furthermore, this novel approach is robust in the sense that it is totally independent of the class of priors $\mathcal{Q}$ considered by investors to model financial asset returns, as well as optimisation criteria; and therefore, can be applied to any kind of portfolio optimisation model.

The organisation of the chapter is as follows. Firstly, a general background is outlined for the portfolio optimisation problem under model ambiguity after which, the Ambiguity Robust Adjustment (ARA) methodology will be presented in details. First, the Absolute Ambiguity Robust Adjustment (AARA) function ψ will be introduced. The AARA transforms the optimal weights computed under each prior, considered according to the idiosyncratic ambiguity aversion the investor displays for each given prior. Then, the Relative Ambiguity Robust Adjustment (RARA) adjustment function π will be also introduced. The RARA accounts for the systematic ambiguity aversion of the different priors considered. The function π allows a mix to be made of the individual optimal weights obtained in the precedent phase through the function ψ. In addition, the specific role of the risk free asset in the ARA methodology will be discussed. In a third section, some key properties of the Ambiguity Robust Adjustment (ARA) transformation will be listed. The last two sections give theoretical examples of the original ARA methodology presented in this research: in the fourth section, the ARA methodology will be compared to the methodology developed by Klibanoff *et al.* (2005) (denoted KMM) in their landmark paper, showing that in the specific example given by Klibanoff *et al.* (2005), the ARA methodology is very similar to the KMM methodology. Finally, a theoretical example of greater complexity will be presented to illustrate how the novel ARA methodology can deal with more complex settings, and to exhibit how in the case considered, portfolio allocation is affected by ambiguity aversion.

5.1 Settings

Initially, the same settings as in Chapter 2 are considered: an investor with a given initial wealth x_0 wants to allocate their wealth among the $N+1$ different assets available in the market. $x^{\boldsymbol{\phi}}$ represents the value of the investor's portfolio at a future time-horizon, and the control variable $\boldsymbol{\phi}$ represents their strategy (i.e. how the wealth is allocated amongst the assets). More precisely, $\boldsymbol{\phi}$ is a vector of weights with each component corresponding to the proportion of wealth the investor allocates to a given asset; a negative value translating the fact that a particular asset is sold. Each element of $\boldsymbol{\phi}$ belongs to a domain$D_{\boldsymbol{\phi}}$ that is considered to be $[-1:1]$. Note that there is an investment constraint: $\sum_{i=0}^{N}\phi^i = 1$, in order to translate the idea that 100% of the initial wealth has been invested. It is assumed that the investor considers several different models to represent the dynamic of $x^{\boldsymbol{\phi}}$; and that the investor is ambiguous in regard to those models.
The following section will re-specify the background literature on ambiguity, exposed earlier in Chapter 4 that is of specific interest to the portfolio allocation problems discussed in this chapter.

The standard Markowitz framework (as presented in Chapter 2) does not include model uncertainty for investment decision-making. The optimal portfolio allocation is obtained as the solution of the following optimisation programme:

$$\phi^* \equiv \underset{\phi}{\operatorname{argmax}}\, \mathbb{E}_{\mathbb{P}}[u(x^{\boldsymbol{\phi}}, \lambda)] \tag{5.1}$$

where u is a Von Neumann-Morgenstern utility function characterising the investor's preferences, and parametrized by the risk-aversion parameter λ. In such a setting, $\mathbb{P}$ stands for the only prior (or model for the distribution of the assets returns) the investor has, which is held without ambiguity. Hence, the risk is perfectly quantifiable by the investor through knowledge of the distribution $\mathbb{P}$.

As suggested by the Subjective Expected Utility (SEU) framework developed by Savage (1954) and discussed in Chapter 4, an agent may also consider different models $\mathbb{Q}$ in a finite set of possible models $\mathcal{Q}$. In such a case, the investment problem is modified according to the subjective view $\pi(\mathbb{Q})$ taken by the investor for each model $\mathbb{Q}$. More precisely, $\pi(\mathbb{Q})$ represents the investor subjective likelihood of the model $\mathbb{Q}$ to occur. The investor operates a linear blending of the different models, weighted by their subjective probability $\pi(\mathbb{Q})$ to be the "real" model. Under each model, the investor considers the objective expected utility of their future wealth. Across all priors, the investor considers the subjective expected value of the expected utilities under the different models. The optimal portfolio allocation is then obtained as:

$$\phi^* \equiv \underset{\phi}{\operatorname{argmax}} \sum_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_{\mathbb{Q}}[u(X^{\phi}, \lambda)]\pi(\mathbb{Q}) \tag{5.2}$$

According to this framework however, even if the agent considers several priors, he is neutral towards model uncertainty: there is no ambiguity towards the set of models considered - or their likelihood to occur.

However, as demonstrated by Ellsberg (1961), decision-makers show more adverse behaviour when betting on events for which outcomes are ambiguous (i.e. when there is also some uncertainty regarding the underlying model); rather, than when betting on events for which the outcomes are only risky (i.e. the underlying model is well known). Consider a financial illustration of the Ellsberg Paradox: the risk premium paradox. Investment tends to be placed in local markets, despite the fact that the expected returns are lower in comparison with those that can be made through foreign markets. This is due to investors adding an ambiguity-premium to foreign risky assets (investors prefer investing in assets located in their geographical zone, because they believe they can better apprehend their return distribution).

The SEU framework fails to take into account this additional source of aversion for financial investors. Of the various approaches presented in Chapter 4 that take into account such aversion towards model uncertainty in the investor decision-making process, the most general approach is the Klibanoff *et al.* (2005) model.

This approach considers an increasing, concave transformation function Ψ that characterises investor ambiguity aversion through the parameter γ. Thus, the optimal-weights vector is determined as:

$$\phi^* \equiv \underset{\phi}{\operatorname{argmax}} \sum_{\mathbb{Q} \in \mathcal{Q}} \Psi \left\{ \mathbb{E}_{\mathbb{Q}}[u(x^{\phi}, \lambda)], \gamma \right\} \pi(\mathbb{Q}) \tag{5.3}$$

This theoretical approach can be challenging to implement in practice for a variety of different reasons; including, the difficulty involved in finding the calibration of the various parameters. Indeed, no distinction is made between specific ambiguity aversion for a given model (i.e. How closely does a specific model represent reality?), and general ambiguity aversion for the whole class of models (i.e. How well does the set of all models encompass reality?). Moreover, providing an explicit solution to Programme (5.3) can be extremely difficult, even numerically; and especially in the multi-dimensional case; or, when a number of constraints are added to portfolio allocation. It is important to note that in their paper, Klibanoff *et al.* (2005) only provide a simple numerical example for a portfolio with 3 assets to illustrate their methodology, whereas practitioners often consider portfolios with hundreds of assets. In addition, the Klibanoff, Marinacci and Munkerji approach lacks flexibility: if an investor considers a new model, they have to entirely re-compute the optimization programme to find the new adequate allocation. To overcome the aforementioned limitations, this thesis proposes a robust general framework for decision-making under uncertainty where, rather than aiming towards the optimal solution for a given criterion, the objective is to find a robust solution in cases when a large number of assets is considered and several different priors taken into account.

5.2 An alternative robust approach to model uncertainty: the Ambiguity Robust Adjustment (ARA)

The main idea behind the novel methodology proposed here, is to perform an ambiguity robust adjustment of each allocation obtained for each individual prior

before performing a global adjustment over the class of models. This methodology, hereafter referred to Ambiguity Robust Adjustment (ARA), has the key advantages of being very flexible (any type of priors can be considered, and any priors can be easily added or removed from the set), tractable (the investor ambiguity aversion can be precisely described in monetary units), and (as will be argued in Chapter 6 where an empirical study in conducted) is better suited to practical situations. Indeed, the ARA methodology is easily applied to large dimensional problems, when investors want to allocate their wealth among a large number of assets.
More precisely, the ARA consists of two main steps, which correspond to an adjustment for two different types of ambiguity aversion:

- Absolute ambiguity aversion : This refers to the ambiguity aversion the investor has for a given prior. It operates an adjustment on the preferred allocation given under a prior $\mathbb{Q}$. More specifically, the investor will first solve the optimisation programme for each prior $\mathbb{Q}$ in the set $\mathcal{Q}$, subject to an investment constraint:

$$\max_{\phi} \mathbb{E}_{\mathbb{Q}}[u(X^{\phi})]$$

  assuming that $\mathbb{Q}$ was the only prior model available to the investor. Then, the optimisation outcome $\operatorname{argmax}_{\phi} \mathbb{E}_{\mathbb{Q}}[u(X^{\phi})]$ is distorted by a function ψ to account for the (absolute) level of ambiguity aversion the investor has toward the prior $\mathbb{Q}$:

$$\psi[\operatorname*{argmax}_{\phi} \mathbb{E}_{\mathbb{Q}}[u(X^{\phi})], \gamma^{\mathbb{Q}}]$$

  The distortion function ψ is called Absolute Ambiguity Robust Adjustment (AARA). ψ is common to all priors, however it is parametrized by a coefficient $\gamma^{\mathbb{Q}}$, which depends on the prior $\mathbb{Q}$. The adjustment of the optimisation outcome through ψ answers the question of how much the prior $\mathbb{Q}$ can be trusted for a particular decision problem.

- Relative ambiguity aversion : This refers to the relative ambiguity aversion the investor has for the set of priors $\mathcal{Q}$. More precisely, the investor aggregates the adjusted allocations per prior they have obtained in the first step, using a function π, called Relative Ambiguity Robust Adjustment (RARA).

The asset allocation of the model $\mathbb{Q}$ will be denoted as $\boldsymbol{\phi}^{\mathbb{Q}}$. The ARA portfolio allocation $\boldsymbol{\phi}^{ARA} \equiv (\phi^{ARA,i})_{i\in\{1,\ldots,N\}}$ is then obtained as:

$$\begin{cases} \phi^{ARA,i} \equiv \sum_{\mathbb{Q}\in\mathcal{Q}} \psi\left\{\phi^{\mathbb{Q},i}, \gamma\right\} \pi(\mathbb{Q}),\ i \in \{1, \ldots, N\} \\ \phi^{ARA,0} \equiv 1 - \sum_{i=1}^{N} \phi^{ARA,1} \end{cases}$$

In the section below the characteristics of the functions ψ and π are described, with particular focus on their desired properties.

5.2.1 The Absolute Ambiguity Robust Adjustment (AARA)

The idea behind the AARA is that because the investor has doubts about the weights generated by a given model, they wish to scale-down those weights, in order to reduce the decision weights obtained for ambiguous models, and especially the biggest absolute weights that could entail the biggest risks in their portfolio. The investor treats the absolute ambiguity aversion with the same type of transformation across all the different models (ψ is the same for all the models). What distinguishes the absolute ambiguity aversion transformation across the models is the specific ambiguity aversion parameter $\gamma^{\mathbb{Q}}$ the investor attributes to each model $\mathbb{Q} \in \mathcal{Q}$. As the optimal weights obtained for each model $\phi^{\mathbb{Q}}$ are bounded by 1, $\psi(1, \gamma^{\mathbb{Q}})$ represents the maximum weight the investor will assign to any asset after the AARA transformation. The following notations are used:

$$\begin{cases} \forall\mathbb{Q} \in \mathcal{Q},\ a^{\mathbb{Q}} \equiv \psi(1, \gamma^{\mathbb{Q}}) \\ \text{and } a \equiv \max_{\mathbb{Q}\in\mathcal{Q}} a^{\mathbb{Q}} \end{cases}$$

Therefore, ψ is defined on the set of optimum model dependent weights $[-1; 1] \times \mathcal{Q}$ onto a set $[-a; a]$ of transformed weights. Note that the investor can set the value of $a^{\mathbb{Q}}$ to express how much he is willing to reduce a maximum weight for a given model $\mathbb{Q}$, and deduce from there the value of $\gamma^{\mathbb{Q}}$ depending on the explicit form chosen for ψ.

Definition 5.1 (AARA). *A function ψ:*

$$\begin{cases} \psi : [-1;1] \times \mathfrak{Q} \to [-a;a] \\ (\phi, \mathbb{Q}) :\to \psi(\phi^{\mathbb{Q}}, \gamma^{\mathbb{Q}}) \end{cases}$$

is an Absolute Ambiguity Robust Adjustment (AARA) if it satisfies the following properties of Universality, Monotonicity and Convexity.

The following properties apply for the risky assets only, (i.e. for $i \in \{1, ..., N\}$).

Property 5.1 (**Universality:**). *The function ψ is the same for all the priors.*

The investor distorts the optimal allocation obtained for each prior using the same type of transformation. The absolute ambiguity aversion adjustment can be different across the different priors depending on the absolute ambiguity aversion parameter $\gamma_{\mathbb{Q}}$; which may differ from one model to the other.

Property 5.2 (**Monotonicity:**). *The function ψ preserves the ranking of the individual risky asset allocations obtained for a given prior $\mathbb{Q}$.*

If for a given model $\mathbb{Q}$, and for two risky assets i and j, $\phi_i < \phi_j$ then $\psi(\phi_i, \gamma^{\mathbb{Q}}) \leq \psi(\phi_j, \gamma^{\mathbb{Q}})$. In other words, the investor is consistent in their choices and the transformation ψ preserves their preferences.
The function ψ also satisfies some properties of convexity to express ambiguity aversion:

Property 5.3 (**Convexity:**). *The function ψ is concave on $[0;1]$ and convex on $[-1;0]$.*

More precisely, the function ψ is parametrized by a coefficient of ambiguity aversion $\gamma^{\mathbb{Q}}$, so that the function ψ reduces the absolute largest weights given by the optimised portfolios under each model considered. The bigger the aversion coefficient $\gamma^{\mathbb{Q}}$ the more averse the investor is to large weights inferred by $\mathbb{Q}$. As it applies greater penalisation to the largest positive and negative weights , the function ψ has an S-shape.

For all assets $i \in [1, N]$ and all models $\mathbb{Q} \in \mathfrak{Q}$: $|\psi(\phi^{\mathbb{Q},i})| \leq |\phi^{\mathbb{Q},i}|$.
In absolute terms, the absolute ambiguity adjusted weights are smaller than the optimal weights computed under a given model $\mathbb{Q}$.
Some additional properties can be considered depending on the assumption made concerning investor preferences and trading constraints.

Property 5.4 (**Invariant point:**). *There is no ambiguity aversion for a zero weight:* $\psi(0) = 0$.

If the model $\mathbb{Q}$ assigns no weight on a given asset, the transformation ψ should not modify the "neutrality" of model $\mathbb{Q}$ in respect to this asset.

Property 5.5 (**Symmetry :**). *The function* ψ *is an odd function symmetric around zero.*

In a context where short selling is possible, there is no reason to differentiate the long or short weights of the same magnitude in terms of ambiguity aversion. It can be assumed that a long-short investor has the same aversion to positive or negative weights of the same absolute value: The AARA function penalises the scale of the optimal weights of a given model without discriminating between negative and positive weights. Which translates to:

$$\forall \phi^i \in [-1;1], \psi(-\phi^i) = -\psi(\phi^i)$$

Property 5.6 (**Limit behaviour:**). *The function* ψ *has the following limit values:*

$$\begin{cases} \forall x \in [-1,1], \ \lim\limits_{\gamma \to \infty} \psi(x,\gamma) = 0 \\ \forall x \in [-1,1], \ \lim\limits_{\gamma \to 0} \psi(x,\gamma) = x \end{cases}$$

An investor who is infinitely averse to ambiguity will be prevented from trading as none of the models considered can be trusted. Therefore, all the portfolio weights should be defaulted to zero. However, if the investor is neutral to ambiguity, the function ψ should leave the model-dependent weights invariant.
This thesis uses a similar function to the one applied by Klibanoff *et al.* (2005) to account for ambiguity; the function ψ can be any classical S-Shape function that possesses the useful properties of concavity (convex for negative values), symmetry and monotonicity (similar attributes as for classical utility functions). An example for the function ψ is:

$$\psi(x,\gamma) \equiv \begin{cases} \frac{1-\exp^{-\gamma x}}{\gamma}, 0 \le x \le 1 \\ \\ \frac{\exp^{\gamma x}-1}{\gamma}, -1 \le x \le 0 \end{cases}$$

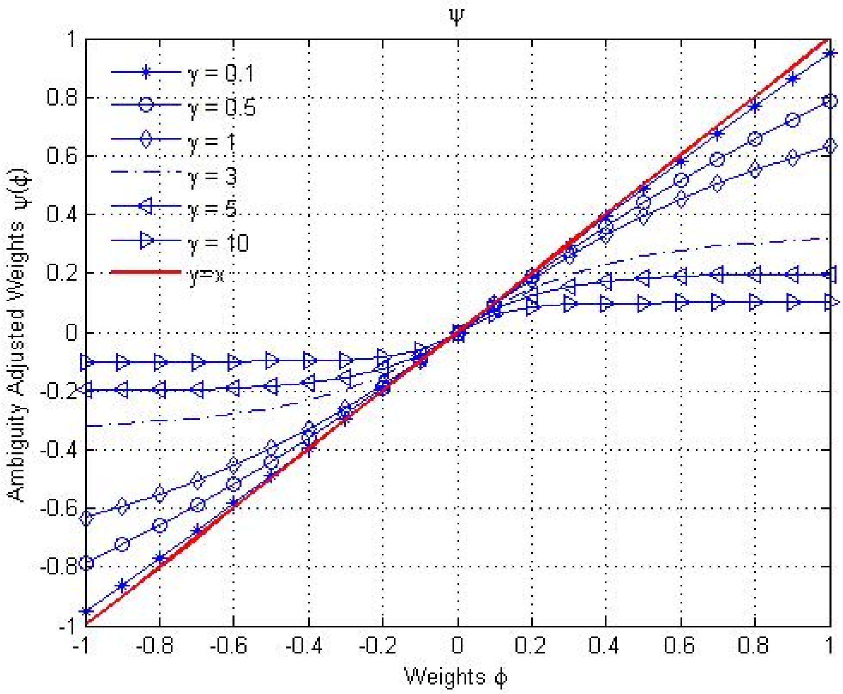


Figure 5.1: ψ for different values of the ambiguity aversion parameter γ

What actually characterises the function ψ is the ambiguity parameter γ that gauges the concavity of the function ψ; and therefore, the ambiguity aversion of the investor.

Note that the transformation by the function ψ does not modify the preferences of the investor, as it is applied on the allocation that maximizes the investor value function. All the four utility preference axioms (completeness, reflexivity, transitivity and continuity) are therefore still respected.

5.2.2 The Relative Ambiguity Robust Adjustment (RARA)

Once the allocations have been computed for each prior $\mathbb{Q}$ and have been independently adjusted for ambiguity aversion through the AARA function ψ, they need to be aggregated across all priors in the set $\mathcal{Q}$. The RARA function takes into account the ambiguity aversion of each prior relative to the whole class of priors $\mathcal{Q}$ and therefore depends on $\mathbb{Q}$ and $\mathcal{Q}$. Such an adjustment is made through a mixture function π. The RARA function $\pi(\mathbb{Q})$ represents the likelihood or degree of confidence the decision maker has for the adjusted result given under the model $\mathbb{Q}$ when all the adjusted results for all the other priors are considered.

$\pi(\mathbb{Q})$ can be seen as a subjective weight given by the decision-maker to the adjusted solution $\psi(\phi^{\mathbb{Q}}, \gamma^{\mathbb{Q}})$ for the model $\mathbb{Q}$. Therefore, $\pi(\mathbb{Q})$ is always non-negative. If the decision-maker categorically mistrusts the prior $\mathbb{Q}$ relative to the other priors, they will simply set the value $\pi(\mathbb{Q})$ to zero. Yet, if on the contrary, the prior $\mathbb{Q}$ relative to the additional priors is fully trusted, then the value of the function π for all the additional priors will be zero. Note that in this case, the weight $\pi(\mathbb{Q})$ is not necessarily one, since the agent may assume a less than full understanding of the situation (i.e. the set of prior $\mathcal{Q}$ does not encompass the true probability $\mathbb{P}$).
More formally the following definition for the RARA function can be given:

Definition 5.2 (RARA)**.** *The function $\pi : \mathcal{Q} \rightarrow [0;1]$ is a Relative Ambiguity Robust Adjustment (RARA) function if:*

$$\forall \mathbb{Q} \in \mathcal{Q}, 0 \leq \pi(\mathbb{Q}) \leq 1$$

and $\sum_{\mathbb{Q}\in\mathcal{Q}} \pi(\mathbb{Q}) \leq 1$

After the transformation through ψ, the weight of any risky assets is defined as:

$$\forall i \in [1; N],\ \phi^{ARA,i} = \sum_{\mathbb{Q}\in\mathcal{Q}} \psi(\phi^{\mathbb{Q},i}, \gamma^{\mathbb{Q}})\pi(\mathbb{Q})$$

and the weight of the risk free asset is defined as:

$$\phi^{ARA,0} = 1 - \sum_{i=1}^{N} \phi^{ARA,i}$$

The major difference with the Subjective Expected Utility framework, is that the sum of the weights $\pi(\mathbb{Q})$ over the set of priors $\mathcal{Q}$ does not necessarily sum to one, since the agents may doubt the existence of a full comprehension of reality. Unlike under the SEU settings, the weights $\pi(\mathbb{Q})$ cannot be assimilated to probabilities. More precisely, if $\sum_{\mathbb{Q}\in\mathcal{Q}} \pi(\mathbb{Q}) < 1$, it means the investor does not hold the belief that a perfect representation of the asset returns distribution with the set of priors $\mathcal{Q}$ can be made.

5.2.3 The role of the risk free asset

Due to the specific nature of the risk free asset, it has no model risk associated with it (its future value is known with certainty). Therefore, it plays a specific role in the ambiguity adjusted optimal asset allocation. It can be assimilated to a refuge value in the following sense: the more the investor is averse to ambiguity, the more they invest in the risk-free asset. Thus, as the "disinvested" part of the wealth from the risky assets is transferred to the risk-free asset, the adjusted weight of the risk-free asset corresponds to the amount of money the investor is reluctant to invest in risky assets due to their aversion towards model risk. After the transformation ψ, the weight of the risk free asset allocation is defined as the residual of wealth not invested in risky assets:

$$\forall \mathbb{Q} \in \mathcal{Q},\ \psi(\phi^{0,\mathbb{Q}}, \gamma^{\mathbb{Q}}) \equiv 1 - \sum_{i=1}^{N} \psi(\phi^{i,\mathbb{Q}}, \gamma^{\mathbb{Q}})$$

And the reserve made because of absolute ambiguity aversion towards the model $\mathbb{Q}$ is therefore defined as:

$$\rho^{\mathbb{Q}} \equiv \psi(\phi^{0,\mathbb{Q}}, \gamma^{\mathbb{Q}}) - \phi^{0,\mathbb{Q}}$$

I.e., it represents the allocation invested in the risk free asset after the AARA transformation minus the allocation initially granted to the risk free asset under the prior $\mathbb{Q}$.

Similarly, after the mixture function π is applied, the final ARA risk free asset allocation is defined as the residual of wealth not invested in risky assets:

$$\phi^{ARA,0} \equiv 1 - \sum_{i=1}^{N} \psi(\phi^i, \gamma^{\mathbb{Q}})\pi(\mathbb{Q})$$

The reserve made because of total ambiguity aversion is therefore defined as:

$$\rho \equiv \phi^{ARA,0} - \sum_{\mathbb{Q}\in\mathcal{Q}} \phi^{0,\mathbb{Q}}\pi(\mathbb{Q})$$

And the reserve made because of relative ambiguity aversion towards the set of models $\mathcal{Q}$ is deduced to be:

$$\rho^{\mathcal{Q}} \equiv \rho - \sum_{\mathbb{Q}\in\mathcal{Q}} \rho^{\mathbb{Q}}\pi(\mathbb{Q})$$

5.2.4 ARA parameterisation

The investor aversion to ambiguity is dynamic; as depending on the period considered, the investors are more or less confident about their models and the overall set of models considered. Therefore, the function π and the ambiguity aversion parameter γ are allowed to adapt dynamically, and expand or contract the total investment size; whether, or not the total ambiguity aversion decreases or increases over time (the ambiguity parameter γ and the function π can be re-parametrized every time a decision is made). As pointed out by Epstein & Schneider (2007), the ambiguity aversion of an investor does not decrease monotonically over time. The novel RARA function allows the investor to dynamically adjust their portfolio weights depending on their beliefs concerning the accuracy of a given prior $\mathbb{Q}$ to model the true distribution $\mathbb{P}$, relatively to the set of priors $\mathcal{Q}$ considered.

The approach differs considerably from the classical Bayesian updating approach (where the investor learns more about the underlying model with any new information flowing in the stock price returns). In a Bayesian framework, investors believe that they have adequate information, and their model ultimately converges toward the true model. Therefore, investor confidence in their model increases gradually and monotonically. Under model ambiguity; however, this is not the case. Investors can become more or less confident over time in a non-monotonic way; thus, investors do not assume that more information can systematically increase confidence about their model.

Many methods could be used in practice to calibrate the adjustment $\pi(\mathbb{Q})$, and the ambiguity aversion parameter $\gamma^{\mathbb{Q}}$ for a given model $\mathbb{Q}$. As an illustration, a simple empirical methodology is proposed that takes into account the relative historical performance of the different models: initially, a number of performance measures (the Sharpe, Sortino, Gain Loss or Win Lose ratios, as described in Chapter 2) are computed for the different models considered, and evaluated over a given time-window. The adjustment π can then be computed as a weighted average of a given performance measure; whereas, the ambiguity aversion parameter γ can be parametrized as the inverse of the particular performance measure chosen for the prior model $\mathbb{Q}$ considered. In Chapter 6, a more in depth description is made of the calibration used in the empirical study testing the performance of the ARA methodology on real data.

5.3 Some definitions relative to the ARA asset allocation

To form comparisons between the different asset allocations for different models, and the impacts of the ambiguity aversion on the different weights assigned to each asset, the following section will present a number of properties and definitions relative to the ARA, as also described in Tobelem & Barrieu (2010a). In addition, a measure is proposed to represent the distance between two different

asset allocations, which is used to compare differing asset allocations in the empirical study developed in Chapter 6. The Ambiguity Robust Adjustment refers to the combined adjustment; firstly, by the AARA of the optimal weights independently computed for each prior; and secondly, by the RARA performed to combine those adjusted weights. The Absolute Ambiguity Adjustment (AAA) and the Relative Ambiguity Adjustment (RAA) measures presented below are relative to the AARA and RARA respectively. Those measures can be used to describe an investor ambiguity aversion toward a specific model and relatively to the whole set of models considered.
For the following definitions, let $\phi^{\mathbb{Q}}$ be the asset allocation conditional on model $\mathbb{Q} \in \mathcal{Q}$ and ϕ^{ARA} the ARA asset allocation.

Definition 5.3. ***Portfolio distance***

Consider two models: $\mathbb{Q}^1$ and $\mathbb{Q}^2$, in the set of priors $\mathcal{Q}$. The distance measure δ between the two models is defined as:

$$\delta(\boldsymbol{\phi}^{\mathbb{Q}^1}, \boldsymbol{\phi}^{\mathbb{Q}^2}) = \sum_{i=0}^{N} |\phi^{\mathbb{Q}^1,i} - \phi^{\mathbb{Q}^2,i}|$$

$\delta(\phi^{\mathbb{Q}_1}, \phi^{\mathbb{Q}_2})$ represents the turnover value to rebalance the investor portfolio from the asset allocation $\phi^{\mathbb{Q}^1}$ to the asset allocation $\phi^{\mathbb{Q}^2}$.

Definition 5.4. ***Absolute Ambiguity Adjustment (AAA)***
The value of the Absolute Ambiguity Adjustment (AAA) of an investor toward the model $\mathbb{Q}$ is defined as:

$$AAA(\mathbb{Q}) \equiv \sum_{i=0}^{N} |\phi^{\mathbb{Q},i} - \psi(\phi^{\mathbb{Q},i}, \gamma^{\mathbb{Q}})|$$

$AAA(\mathbb{Q})$ represents the theoretical turnover to rebalance the investor's asset allocation on the risky assets $i = 1, ..., N$ obtained under the prior $\mathbb{Q}$ to the Absolute Ambiguity Adjusted asset allocation, taking into account the investor absolute aversion against the same prior $\mathbb{Q}$.

Thus, it is deduced that the total value of the investors' Absolute Ambiguity Adjustment toward all their priors is defined as:

$$AAA(\mathcal{Q}) \equiv \sum_{\mathbb{Q}\in\mathcal{Q}} \sum_{i=0}^{N} |\phi^{\mathbb{Q},i} - \psi(\phi^{\mathbb{Q},i}, \gamma^{\mathbb{Q}})| \pi(\mathbb{Q})$$

Definition 5.5. ***Relative Ambiguity Adjustment (RAA)***
The value of the Relative Ambiguity Adjustment (RAA) of an investor is defined as:

$$RAA(\mathcal{Q}) \equiv \sum_{i=0}^{N} |\phi^{ARA,i} - \sum_{\mathbb{Q}\in\mathcal{Q}} \phi^{\mathbb{Q},i}\pi(\mathbb{Q})|$$

Which effectively represents the turnover between the Robust Ambiguity Portfolio and the Subjective Expected Utility Portfolio.

Recall that, in the singular SEU case, π is considered to be a probability and; therefore, $\sum_{\mathbb{Q}\in\mathcal{Q}} \pi\mathbb{Q} = 1$, the SEU portfolio allocation is thus defined as:

$$\phi^{SEU,i} \equiv \sum_{\mathbb{Q}\in\mathcal{Q}} \phi^{\mathbb{Q},i}\pi(\mathbb{Q})$$

Note that this is the assumption made in the remainder of this and the next Chapter (where some empirical tests are run to evaluate the ARA methodology). In Chapter 7, this assumption about π is relaxed.

5.4 Comparison with the Klibanoff, Marinacci, Mukerji model (KMM)

In the present section, the approach of Klibanoff *et al.* (2005) is compared to the novel ARA methodology presented in this research.

5.4.1 Settings

In the article illustrating their methodological basis, Klibanoff, Marinacci and Mukerji provide an example that considers a simple one period, three asset model with a: risk free asset s^0, a risky asset s^1 and an ambiguous asset s^2. In this model, there are two different states of the world ω^1 and ω^2. The investor considers two different priors $\mathbb{Q}^1$ and $\mathbb{Q}^2$ and has equal subjective beliefs for both of them:

$$\pi(\mathbb{Q}^1) = \pi(\mathbb{Q}^2) = \frac{1}{2}$$

The three assets have the same initial value of 1; their terminal values at the horizon time are given as follows:

Table 5.1: KMM example framework

	$\mathbb{Q}^1$		$\mathbb{Q}^2$	
	$\pi(\mathbb{Q}^1) = \frac{1}{2}$		$\pi(\mathbb{Q}^2) = \frac{1}{2}$	
	ω^1	ω^2	ω^1	ω^2
	$\mathbb{Q}^1(\omega^1) = \frac{1}{4}$	$\mathbb{Q}^1(\omega^2) = \frac{3}{4}$	$\mathbb{Q}^2(\omega^1) = \frac{3}{4}$	$\mathbb{Q}^2(\omega^2) = \frac{1}{4}$
s^0	1.15	1.15	1.15	1.15
s^1	3	1	3	1
s^2	2	2	1	1

Note that:

- s^0 is a risk-free asset, as it has the same terminal value across all states and the universe of priors;
- s^1 is a risky asset as it has the same possible terminal values for both models $\mathbb{Q}^1$ and $\mathbb{Q}^2$ but its value depends on the state of the universe (ω^1 and ω^2);
- Finally s^2 is ambiguous as its terminal value depends on the model taken into account but not the state of the universe;

Recall that under the ARA methodology, the solution is defined as:

$$\boldsymbol{\phi}^{ARA} \equiv \sum_{\mathcal{Q}} \psi(\boldsymbol{\phi}^{\mathbb{Q}}, \gamma^{\mathbb{Q}})\pi(\mathbb{Q})$$

In this case, the different models asset allocation $\phi^{\mathbb{Q}}$, where $\mathbb{Q} \in \mathcal{Q}$, are defined as the solutions of the expected utility maximisation:

$$\boldsymbol{\phi}^{\mathbb{Q}} \equiv \underset{\boldsymbol{\phi}}{\operatorname{argmax}}\, \mathbb{E}_{\mathbb{Q}}[u(x^{\boldsymbol{\phi}}, \lambda)]$$

Whereas, the KMM solution is defined as:

$$\boldsymbol{\phi}^{KMM} \equiv \underset{\phi}{\operatorname{argmax}}\, \mathbb{E}_{\pi}\psi[\mathbb{E}_{\mathbb{Q}}[u(x^{\boldsymbol{\phi}}, \lambda)], \gamma]$$

5.4.2 Results

Using the same choice criterion presented by Klibanoff, Marinacci and Mukerji (the same utility function is considered), the weights for the three different assets are obtained for various scenarios on the risk aversion parameter λ and the ambiguity aversion parameter γ for both the KMM and the ARA models. The respective solution weights are plotted in Figure (5.3) and Figure (5.2).

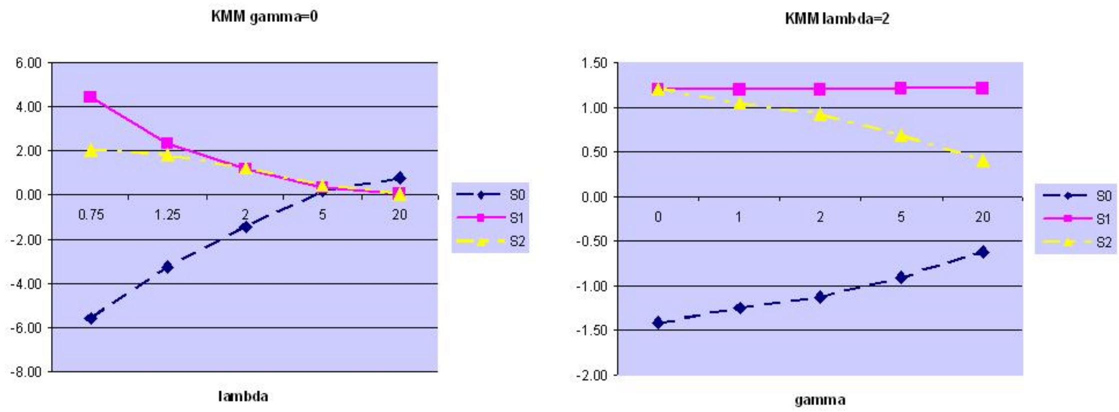


Figure 5.2: KMM weights

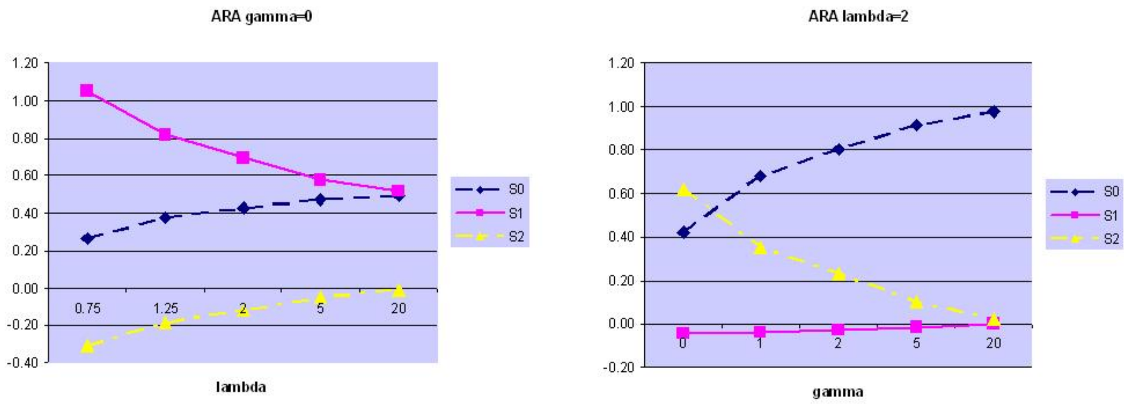


Figure 5.3: ARA weights

Two cases are considered:

- The ambiguity parameter $\gamma = 0$ is fixed, and the risk aversion parameter λ is allowed to vary. As the risk aversion increases (assuming the ambiguity aversion remains null), the proportion of the risky asset decreases, both for the ARA model and the KMM model; whereas, the proportion of the risk free asset increases. However, the ambiguity asset proportion tends to rise in the ARA model; whereas, it also decreases in the KMM model (although it decreases less than the proportion of the risky asset). The ARA model tends to discriminate between the risky and ambiguous assets better than the KMM model. Also, the weights are more extreme in the KMM model than in the ARA model (ranging from -6 to 8; whereas, the ARA weights remain in the [-.4;1.2] range).

- The risk parameter $\lambda = 2$ is fixed and the ambiguity aversion parameter γ is allowed to vary. Both models display the same behaviour: when the ambiguity aversion parameter increases, the ambiguity asset allocation decreases to the profit of the risk free asset while the risky asset allocation remains constant.

The KMM and ARA models adjust the asset allocation in a very similar way, but the ARA model tends to allocate more weight to the risk free asset as the ambiguity aversion increases; although, this depends mainly on the calibration of the risk and ambiguity aversion parameters. To conclude, it can be said that when considering the conditions within the framework given by Klibanoff, Marinacci and Mukerji in their basic example, the ARA and KMM methodologies are indeed, very similar. The great advantage of the ARA methodology is that it can be applied to more complex theoretical settings and to large dimensional empirical problems - as will be demonstrated in the next section.

5.5 A parametrized model application

In this present section, a theoretical example with settings of greater complexity is presented (this example has also been developed in Tobelem & Barrieu (2010b)).

In this example both the set of priors, and the distribution for the risky and ambiguous assets, are continuous. It is shown that the original ARA methodology can be used to explicitly solve the asset allocation problem under ambiguity in a given theoretical case of high complexity, which is simply not possible with the KMM model. The framework of this theoretical example is first outlined, then the resulting ARA weights are formally computed. Finally, a particular focus is given on the asymptotic ARA weights when $\gamma \to 0$ and $\gamma \to \infty$.

5.5.1 Settings

The settings for this specific example will now be described; it is assumed that:

- It is a one period model: at time 0, an investment decision is made and at time T, the terminal payoff of the strategy is observed.
- The set of priors $\mathcal{Q}$ is a countable set of priors $\mathbb{Q}^q$, where $q \in [0, d]$. π defines the distribution upon the different priors, and it is assumed that all the priors are equipotent for the investor:

  $$\forall q \in [0, d], \pi(\mathbb{Q}^q) = \frac{1}{d}$$

  Meaning, all priors have the same likelihood $\frac{1}{d}$. As assumed in the precedent example, three assets s^0, s^1 and s^2 are considered, with an initial value 1. The terminal values s_T^0, s_T^1 and s_T^2 of those assets at time T are defined as:
- $s_T^0 = r_f$. The risk free return is r_f. The asset s^0 displays a constant return r_f whatever prior is considered; thus, it is non-risky and non-ambiguous.
- $s_T^1 = s$, where s follows a normal distribution with mean $q \in [0, d]$ (where $r_f < d$), and standard deviation σ. The risky asset follows the same normal distribution under any prior $\mathbb{Q}^q$:

  $$\forall \mathbb{Q}^q,\ s \hookrightarrow_{\mathbb{Q}^q} N(q, \sigma)$$

  The asset s^1 displays a normally distributed return, whatever prior $\mathbb{Q}^q$ is considered; thus, it is is a non-ambiguous, risky-asset.

- $s_T^2 = q$. The ambiguous asset follows a uniform distribution upon the priors distribution:

$$q \hookrightarrow_\pi U_{[0,d]}$$

  The asset s^2 displays a constant return q depending on the prior $\mathbb{Q}^q$ considered. Under a given prior $\mathbb{Q}^q$, s^2 is risk free; thus, it is an ambiguous, non-risky asset.

In addition, the investor utility function is defined as:

$$u(x, \lambda) = -\exp^{-\lambda x}$$

where λ stands for the investor risk aversion parameter. The investor wants to form an optimal portfolio that maximises future expected wealth utility; where the future wealth x_T^{ϕ} is defined as:

$$x_T^{\phi} = \sum_{i=0}^{2} \phi^i s_T^i$$

where the $\phi^i_{1 \leq i \leq 2}$ denote the different weights of the assets s^i in the investor's portfolio; note that the following is given:

$$x_0^{\phi} = \sum_{i=0}^{2} \phi^i = 1$$

and the ambiguous, risky terminal wealth is defined as:

$$x_T^{\phi} = \phi^0 r_f + \phi^1 s + \phi^2 q$$

In the following subsection, the ARA transformed weights are computed.

5.5.2 The ARA transformation

The optimal weights obtained through an ARA transformation will now be computed. First, the optimal weights under each prior $\mathbb{Q}^q$ will be computed. Two cases can be distinguished:

- Case where $0 \leq q \leq r_f$

  For all the priors $\mathbb{Q}^q$, $q \in [0, r_f]$, the ambiguous asset always provides a lower return than the risk-free asset. Therefore, under a prior $\mathbb{Q}^q$, the investor will only consider investments in the risky asset and the most profitable risk-free asset (i.e. in the present case s^0). In this case, $\phi^2 = 0$, and for simplification it can be denoted that $\phi^1 = \phi$ and $\phi^0 = 1 - \phi$.

  Under a prior $\mathbb{Q}^q$, $q \in [0, r_f]$, the investor wants to optimise the following program:

  $$\max_{\phi} \mathbb{E}_{\mathbb{Q}^q}[u(x_T^{\phi}, \lambda)] = -\max_{\phi} \mathbb{E}_{\mathbb{Q}^q} \left\{ \exp^{-\lambda[\phi s + (1-\phi) r_f]} \right\}$$

  It is deduced that the optimal solution in this case is:

	Optimal weights, $0 \leq q \leq r_f$
$\phi^{\mathbb{Q}^q,0}$	$1 - \frac{d - r_f}{\lambda \sigma^2}$
$\phi^{\mathbb{Q}^q,1}$	$\frac{d - r_f}{\lambda \sigma^2}$
$\phi^{\mathbb{Q}^q,2}$	0

- Case where $r_f < q \leq d$

  For all the priors $\mathbb{Q}^q$, $q \in]r_f, d]$, the ambiguous asset always provides a greater return than the risk-free, ambiguous free asset. As previously argued, under a prior $\mathbb{Q}^q$, the investor will only consider investments in the risky and the most profitable risk-free assets (i.e. in the present case s^2). In this case, $\phi^0 = 0$, and for simplification it can be denoted that $\phi^1 = \phi$ and $\phi^2 = 1 - \phi$. Under a given prior $\mathbb{Q}^q$, $q \in [r_f, d]$, the risk-free return of the ambiguous asset is q and the investor wants to optimise the following programme:

  $$\max_{\phi} \mathbb{E}_{\mathbb{Q}^q}[u(x_T^{\phi}, \lambda)] = -\max_{\phi} \mathbb{E}_{\mathbb{Q}^q} \left\{ \exp^{-\lambda[\phi s + (1-\phi) q]} \right\}$$

  By applying calculus similar to that previously applied, it is found that the optimal solution in this case is:

	Optimal weights, $r_f < q \leq d$
$\phi^{\mathbb{Q}^q,0}$	0
$\phi^{\mathbb{Q}^q,1}$	$\frac{d-q}{\lambda\sigma^2}$
$\phi^{\mathbb{Q}_q,2}$	$1-\frac{d-q}{\lambda\sigma^2}$

It is now necessary to apply the AARA transformation to the optimal weights obtained for each prior $\mathbb{Q}^q$. The following AARA function ψ is considered:

$$\psi(x,\gamma) \equiv \begin{cases} \frac{1-\exp^{-\gamma x}}{\gamma}, 0 \leq x \leq 1 \\ \frac{\exp^{\gamma x}-1}{\gamma}, -1 \leq x \leq 0 \end{cases}$$

To simplify the case-study, it is assumed that the parameter γ remains the same for all the priors considered (this goes along with the fact that the investor applies an homogeneous weight to all the priors: a priori, the investor considers all the priors equally ambiguous). Thus:

$$\forall \mathbb{Q}^q, \gamma^{\mathbb{Q}^q} = \gamma$$

The RARA transformation is also applied across all the priors considered; therefore, the final ARA optimal weights are defined as:

$$\begin{cases} \phi^{ARA,i} \equiv \int_0^d \psi(\phi^{\mathbb{Q}^q,i},\gamma)d\pi(\mathbb{Q}^q), i \in \{1;2\} \\ \phi^{ARA,0} \equiv 1-\phi^{ARA,1}-\phi^{ARA,2} \end{cases}$$

It is assumed that the investor's risk aversion λ is such that $\lambda \geq \frac{d}{\sigma^2}$; and therefore, all the optimal weights under all the priors are defined on the interval $[0,1]$. When $\lambda < \frac{d}{\sigma^2}$, the calculus would be similar.
The optimal weights under the ARA transformation are computed as:

$$\phi^{ARA,i} = \frac{1}{d}\int_0^d \frac{(1-\exp^{-\gamma\phi^{\mathbb{Q}^q,i}})}{\gamma}dq, i \in \{1;2\}$$

More specifically, the optimal weights are given as:

- $\phi^{ARA,0} = 1-\phi^{ARA,1}-\phi^{ARA,2}$
- $\phi^{ARA,1} = \int_0^{r_f}\frac{1}{d}\frac{(1-\exp^{-\gamma(\frac{d-r_f}{\lambda\sigma^2})})}{\gamma}dq + \int_{r_f}^d \frac{1}{d}\frac{(1-\exp^{-\gamma(\frac{d-q}{\lambda\sigma^2})})}{\gamma}dq$

- $\phi^{ARA,2} = \int_{r_f}^{d} \frac{1}{d} \frac{(1-\exp^{-\gamma(1-\frac{d-q}{\lambda\sigma^2})})}{\gamma} dq$

Finally, the ARA weights are defined as:

	ARA weights
$\phi^{ARA,0}(\gamma)$	$1 - \phi^{ARA,1}(\gamma) - \phi^{ARA,2}(\gamma)$
$\phi^{ARA,1}(\gamma)$	$\frac{d-r_f}{d\gamma} + (\frac{r_f}{d} - \frac{\lambda\sigma^2}{d\gamma})\frac{1-\exp^{-\gamma\frac{d-r_f}{\lambda\sigma^2}}}{\gamma}$
$\phi^{ARA,2}(\gamma)$	$\frac{d-r_f}{d\gamma} + \frac{\exp^{-\gamma}\lambda\sigma^2}{d\gamma}\frac{1-\exp^{\gamma\frac{d-r_f}{\lambda\sigma^2}}}{\gamma}$

It has been shown that it is straightforward to compute the ARA weights under theoretical settings of greater complexity. The ARA solutions asymptotic behaviour will now be studied, when $\gamma \to \infty$ and $\gamma \to 0$.

5.5.3 Asymptotic behaviour of the ARA weights

In order to test the consistency of the ARA methodology, the ARA weights are studied at the limit values of the parameter γ, and it is shown that they comply with Property 5.6.

5.5.3.1 ARA weights asymptotic behaviour when $\gamma \to \infty$

It will now be assumed that the aversion to ambiguity of the investor is infinite: the investor does not trust any of their models. Therefore, it is given (see proof in the Appendix):

$$\begin{cases} \lim\limits_{\gamma \to \infty} \phi^{ARA,1}(\gamma) = 0 \\ \lim\limits_{\gamma \to \infty} \phi^{ARA,2}(\gamma) = 0 \\ \lim\limits_{\gamma \to \infty} \phi^{ARA,0}(\gamma) = 1 \end{cases}$$

When the ambiguity aversion extends to infinity, investors invest all their wealth in the risk-free non-ambiguous asset; as they do not trust any prior models, they cannot invest in any of the risky assets.

5.5.3.2 ARA weights asymptotic behaviour when $\gamma \to 0$

The ARA weights will now be computed for a scenario when an investor has no aversion to ambiguity: $\gamma = 0$.

$$\begin{cases} \phi^{ARA,1}(0) = \frac{r_f}{d}\frac{d-r_f}{\lambda\sigma^2} + \frac{d-r_f}{r_f}\frac{d-\frac{d+r_f}{2}}{\lambda\sigma^2} \\ \phi^{ARA,2}(0) = \frac{d-r_f}{d}(1 - \frac{d-\frac{d+r_f}{2}}{\lambda\sigma^2}) \\ \phi^{ARA,0}(0) = \frac{r_f}{d}(1 - \frac{d-r_f}{\lambda\sigma^2}) \end{cases}$$

When the aversion to ambiguity is null, it can be considered in a similar light to the case of the Subjective Expected Utility as seen in Savage (1954); and thus, the optimal weights are equal to the expected weights of the prior conditional weights:

$$\phi^{ARA,i} = \mathbb{E}_\pi(\phi^{\mathbb{Q}^q,i}), i \in \{0,1,2\}$$

This is consistent with the remark made in Chapter 4.

5.5.3.3 ARA weights for different values of the parameter γ

In Figure (5.4), the ARA weights are plotted with a parameter γ ranging from 0.1 to 5:

The aversion to ambiguity affects both the risky asset allocation, and the ambiguous asset allocation. The weights of both assets decrease with an increase in the aversion parameter γ; whereas, the allocation of the risk-free non-ambiguous asset increases. Therefore, an easy to compute, close-form solution can be provided for an asset allocation problem under ambiguity when considered under relatively complex settings (i.e. continuous set of priors and continuous distribution for the risky and ambiguous assets).

5.6 Conclusion

In this chapter, an easy to implement and robust ambiguity methodology has been proposed that allows the investor to adapt their portfolios to their level of ambiguity aversion. The calibration for the Absolute Ambiguity Aversion parameter γ and the Relative Ambiguity Adjustment function π will be the subject of

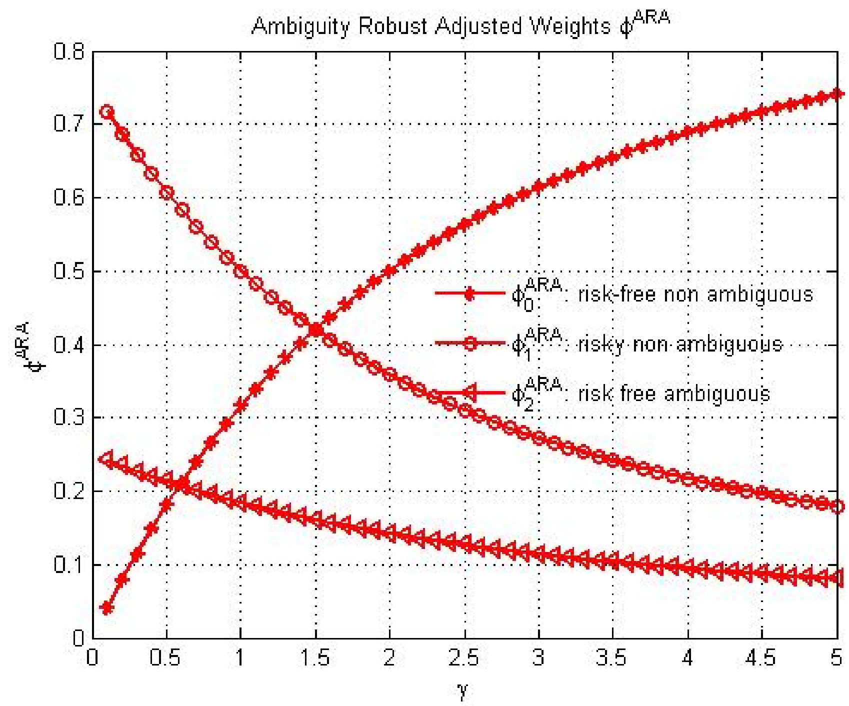


Figure 5.4: ϕ^{ARA}

the following two chapters. The question of mixing appropriately the weights obtained through heterogeneous priors remains a great challenge in many scientific fields. In particular, in Chapter 6, a linear form for the function π is considered. In Chapter 7, an investigation is undertaken of nonlinear forms for the function π.

5.7 Appendix

5.7.1 ARA-KMM weights comparison

The Klibanoff, Marinacci and Mukerji example settings are recalled, and three assets are considered: s^0, s^1 and s^2 respectively as a risk-free, risky and ambiguous asset. In the two different states of the world ω_1 and ω_2 and under the two different priors considered $\mathbb{Q}^1$ and $\mathbb{Q}^2$, the asset values are taken to be the following:

	$\mathbb{Q}^1$		$\mathbb{Q}^2$	
π	$\frac{1}{2}$		$\frac{1}{2}$	
	ω_1	ω_2	ω_1	ω_2
μ	$\frac{1}{4}$	$\frac{3}{4}$	$\frac{3}{4}$	$\frac{1}{4}$
s^0	1.15	1.15	1.15	1.15
s^1	3	1	3	1
s^2	2	2	1	1

The utility function is defined as:

$$u(x) = 1 + \frac{x^{1-\lambda} - 1}{2^{1-\lambda} - 1}, \quad \lambda \neq 1$$

Thus, it is gained:

$$u'(x) = \frac{1-\lambda}{2^{1-\lambda} - 1} x^{-\lambda}, \quad \lambda \neq 1$$

5.7.1.1 Computation of the ARA weights

The respective weights of the assets s^0, s^1 and s^2 are defined by ϕ^0, ϕ^1 and ϕ^2 .

- **Under $\mathbb{Q}^1$**

  The ambiguous asset gives a higher return than the risk free asset under the two states of the world, therefore it is denoted:

$$\phi^0 = 0$$

, $\phi^1 = \phi$ and $\phi^2 = 1 - \phi$

Therefore:

$$V_{\mathbb{Q}^1}(\phi) = \mathbb{E}_{\mathbb{Q}^1}[u(X^\phi)]$$

where X^ϕ is the final wealth of the investor.

$$V_{\mathbb{Q}^1}(\phi) = \frac{1}{4}u(\phi+2) + \frac{3}{4}u(2-\phi)$$

Thus, the following is gained:

$$\frac{\delta V}{\delta \phi} = 0 \Leftrightarrow \frac{1}{4}u'(\phi+2) - \frac{3}{4}u(2-\phi)$$

Finally, it is given:

$$\phi^{\mathbb{Q}^1} = \frac{2(1-\exp^{\frac{\log 3}{\lambda}})}{1+\exp^{\frac{\log 3}{\lambda}}}$$

- **Under** $\mathbb{Q}^2$

  The risk free asset gives a higher return than the ambiguous asset under the two states of the world, therefore it is denoted:

$$\phi^0 = 1-\phi$$

  , $\phi^1 = \phi$ and $\phi^2 = 0$

  It is gained:

$$V_{\mathbb{Q}^2}(\phi) = \mathbb{E}_{\mathbb{Q}^2}[u(X^\phi)]$$

$$V_{\mathbb{Q}^2}(\phi) = \frac{3}{4}u(1.85\phi+1.15) + \frac{1}{4}u(1.15-0.15\phi)$$

  Thus:

$$\frac{\delta V}{\delta \phi} = 0 \Leftrightarrow \frac{5.55}{4}u'(1.85\phi+1.15) - \frac{0.15}{4}u(1.15-0.15\phi)$$

Finally it is given:

$$\phi^{\mathbb{Q}^2} = \frac{1.15(\exp^{\frac{\log 37}{\lambda}} - 1)}{1.85 + 0.15 \exp^{\frac{\log 37}{\lambda}}}$$

5.7.1.2 ARA and KMM weights comparison

The results computed by Klibanoff, Marinacci and Mukerji are compared to the ARA weights. Two cases are considered: firstly, when the ambiguity aversion γ is set to 0 and the risk-aversion varies; and secondly, when the risk aversion λ is set to 2 and the ambiguity aversion varies. The comparative results are displayed in the following tables:

$\gamma = 0$	KMM			ARA		
λ	ϕ^0	ϕ^1	ϕ^2	ϕ^0	ϕ^1	ϕ^2
0.75	-5.5514	4.4715	2.0799	-0.8311	1.0233	0.8078
1.25	-3.2214	2.3605	1.8608	-0.1065	0.4180	0.6886
2	-1.4097	1.1961	1.2136	0.1858	0.1955	0.6187
5	0.1977	0.3576	0.4447	0.3969	0.0554	0.5477
20	0.8211	0.0762	0.1027	0.4767	0.0114	0.5119

Table 5.2: Comparison of ARA and KMM portfolio weights when $\gamma = 0$

$\lambda = 2$	KMM			ARA		
γ	ϕ^0	ϕ^1	ϕ^2	ϕ^0	ϕ^1	ϕ^2
0	-1.4097	1.1961	1.2136	0.1858	0.1955	0.6187
1	-1.2493	1.2017	1.0476	0.5175	0.1276	0.3549
2	-1.1278	1.2052	0.9226	0.6867	0.0843	0.2290
5	-0.9044	1.2102	0.6943	0.8740	0.0262	0.0998
20	-0.6210	1.2139	0.4071	0.9748	0.0002	0.0250

Table 5.3: Comparison of ARA and KMM portfolio weights when $\lambda = 2$

5.7.2 Theoretical illustration

5.7.2.1 Computation of ϕ^{ARA}

Under a given prior $\mathbb{Q}$, the investor wants to optimise the following program ($V(\phi)$ denotes the value function the investor wants to optimise):

$$\max_{\phi}[V(\phi)] \equiv \max_{\phi} \mathbb{E}_{\mathbb{Q}}[u(X_T^{\phi}, \lambda)] = -\max_{\phi} \mathbb{E}_{\mathbb{Q}} \left\{ \exp^{-\lambda[\phi s + (1-\phi) r_f]} \right\}$$

$$V(\phi) = -\exp^{-\lambda[(1-\phi)r_f]} \mathbb{E}_{\mathbb{Q}}[\exp^{-\lambda \phi s}]$$

By a Laplace transform the following is given:

$$\mathbb{E}_{\mathbb{Q}}[\exp^{-\lambda \phi s}] = \exp^{-\lambda(\phi d - \frac{\sigma^2 \phi^2 \lambda}{2})}$$

So that:

$$V(\phi) = -\exp^{-\lambda[(1-\phi)r_f + \phi d - \frac{\sigma^2 \phi^2 \lambda}{2}]}$$

The first order condition becomes:

$$\frac{\partial V(\phi)}{\partial \phi} = -\lambda(d - \phi \sigma^2 \lambda) \exp^{-\lambda[(1-\phi)r_f + \phi d - \frac{\sigma^2 \phi_2^2 \lambda}{2}]}$$

It is deduced:

$$\frac{\partial V(\phi)}{\partial \phi} = 0 \Leftrightarrow \phi = \frac{d - r_f}{\lambda \sigma^2}$$

5.7.2.2 Computation of $\phi^{ARA,1}(0)$

It is recalled that:

$$\phi^1(\gamma) = \frac{d - r_f}{d\gamma} + (\frac{r_f}{d} - \frac{\lambda \sigma^2}{d\gamma}) \frac{1 - \exp^{-\gamma \frac{d - r_f}{\lambda \sigma^2}}}{\gamma}$$

In the neighbourhood of zero, the following limited developments are gained:

$$\frac{1 - \exp^{-\gamma \frac{d - r_f}{\lambda \sigma^2}}}{\gamma} \overset{\gamma \to 0}{=} \frac{d - r_f}{\lambda \sigma^2} - \frac{(d - r_f)^2}{(2\lambda \sigma^2)^2} \gamma + \circ(\gamma)$$

Therefore,

$$\phi^1(\gamma) = \frac{d-r_f}{d\gamma} + (\frac{r_f}{d} - \frac{\lambda\sigma^2}{d\gamma})(\frac{d-r_f}{\lambda\sigma^2} - \frac{(d-r_f)^2}{(2\lambda\sigma^2)^2}\gamma) + \circ(\gamma)$$

The following result is immediately given:

$$\phi^{ARA,1}(0) = \frac{r_f}{d}\frac{d-r_f}{\lambda\sigma^2} + \frac{d-r_f}{r_f}\frac{d-\frac{d+r_f}{2}}{\lambda\sigma^2}$$

This is consistent with the fact that $\psi(\phi, 0) = \phi$. Indeed, it is also given that:

$$\phi^{ARA,1}(0) = \int_0^{r_f} \frac{d-r_f}{\lambda\sigma^2}dq + \int_{r_f}^{d} \frac{d-q}{\lambda\sigma^2}dq = \frac{r_f}{d}\frac{d-r_f}{\lambda\sigma^2} + \frac{d-r_f}{r_f}\frac{d-\frac{d+r_f}{2}}{\lambda\sigma^2}$$

The end result is the classical Savage Expected Utility optimal weights.
In a similar way, the risk free and ambiguous weight when γ tends to 0 can be computed:

5.7.2.3 Computation of $\phi^{ARA,2}(0)$

It is recalled that:

$$\phi^{ARA,2}(\gamma) = \frac{d-r_f}{d\gamma} + \frac{\exp^{-\gamma}\lambda\sigma^2}{d\gamma}\frac{1-\exp^{\gamma\frac{d-r_f}{\lambda\sigma^2}}}{\gamma}$$

In the neighbourhood of zero, the following limited developments are gained:

$$\begin{cases} \exp^{-\gamma} \overset{\gamma\to 0}{=} 1-\gamma+\circ(\gamma) \\ \frac{1-\exp^{\gamma\frac{d-r_f}{\lambda\sigma^2}}}{\gamma} \overset{\gamma\to 0}{=} -\frac{d-r_f}{\lambda\sigma^2} - \frac{(d-r_f)^2}{2(\lambda\sigma^2)^2}\gamma + \circ(\gamma) \end{cases}$$

The following result is immediately given:

$$\phi^{ARA,2}(0) = \frac{d-r_f}{d}(1 - \frac{d-\frac{d+r_f}{2}}{\lambda\sigma^2})$$

This is consistent with the fact that $\psi(\phi, 0) = \phi$. Indeed, it is also given:

$$\phi^{ARA,2}(0) = \int_{r_f}^{d} 1 - \frac{d-q}{\lambda\sigma^2}dq = \frac{d-r_f}{d}(1 - \frac{d-\frac{d+r_f}{2}}{\lambda\sigma^2})$$

Thus, the classical Savage SEU solution is found.

Chapter 6

Evidence from Empirical Study : Outperformance of the ARA Portfolio

"*Look deep into nature, and then you will understand everything better.*"

Attributed to Albert Einstein.

In practice, the portfolio allocation problem often deals with a great number of assets, as investors want to capture the diversification effect. When considering hundreds of assets, it is crucial to be able to use an easy to compute, simple methodology that allows investors to make decisions in a timely fashion. More and more, portfolio strategies involve high-frequency rebalances when the portfolio allocation is revised up to several times a day. The Ambiguity Robust Adjustment proposed in this thesis has the great advantage of being easily applicable to the sorts of large dimension, complex empirical problems faced by financial investors, which is not the case for other theoretical methodologies presently proposed in the literature. In this section, an empirical study is run to evaluate the performance of the Ambiguity Robust Adjusted (ARA) portfolio on real data. The performance of the ARA methodology is compared to simple single strategies and to the Subjective Expected Utility (SEU) portfolio (that does not take into account ambiguity when blending the single portfolio allocations). More precisely, an investor who uses different models is considered, and the performance of the portfolios obtained by a classical SEU blending and the ARA blending (adjusting the allocation with respect to ambiguity aversion, as presented in Chapter 5) of the single models are compared. It is shown that adjustment to ambiguity aversion allows the investor to significantly enhance portfolio returns and reduce the portfolio turnover (and therefore transaction costs). The ARA portfolios that adjust asset allocation with respect to absolute and relative ambiguity aversion consistently beat the SEU portfolios obtained as a simple linear combination of the portfolios built from the different models considered by the investor, in terms of all the performance measures considered. This empirical study, being relatively close to the reality faced by financial investors, practically shows that the methodology proposed can be easily used to mix different models in order to compute the final investment portfolio and enhance investor returns.

This chapter is organised as follows. Firstly, the general background for the empirical study undertaken is provided: the specificities of the dataset used are briefly recalled, some notations given, the description of the portfolios tested and details of the empirical computation of the different portfolio weights provided, as well as the performance measures used to parametrize the ambiguity aversion parameters. In the second section, details of the empirical calibration of the aversion parameter γ and the aversion adjustment π are given, and the performances of the different single portfolios considered, as well as of the combined SEU and ARA portfolios, are displayed.

6.1 Empirical study framework

Using the same dataset, as presented in Chapter 3, a back test on historical European data, when investors make a daily re-balance of their portfolio, is run, by re-estimating the different models over a rolling estimation window, re-calibrating their aversion to ambiguity and re-setting the resulting investment weights every day. The different performance measures presented in Chapter 2 are then used to evaluate the performance of the different strategies.

In this section, the framework of the empirical study is first presented; then, the computation of the portfolio weights for the different models considered is outlined, i.e. the Equally Weighted (EW), the Minimum Variance (MN), the Mean Variance (MV) and the Capital Asset Pricing Model (CAPM) portfolios as presented in Chapter 2, as well as the External Factor Model (EFM), Fundamental Factor Model (FFM), Principal Component Analysis (PCA), Independent Component Analysis (ICA) and Cluster Analysis (CA) portfolios as described in Chapter 3. The computation of the four performance measures introduced at the end of Chapter 2 is detailed and further used to parametrize the absolute ambiguity aversion parameter γ and the relative aversion adjustment π.

6.1.1 Dataset and notations

The same cleaned dataset of European stock returns (based on historical closing prices for the Eurostoxx 600 constituents as of end of May 2010) from January 2000 to May 2010, as presented in Chapter 3 is utilised. First, some specific notations used throughout this chapter are recalled:

- $\mathbf{S} \equiv \{s_t^n\}_{T \times N}$ denotes the matrix of stock prices over the period considered, where $T = 2712$ denotes the number of days and $N = 600$ the number of stocks.
- $\mathbf{M} \equiv r_{t+1}^n = \{\frac{s_{t+1}^n}{s_t^n} - 1\}_{T-1 \times N}$ defines the matrix of asset arithmetic returns.
- $\mu_{\mathbf{t},\mathbf{t+h}}$[1] and $\mathbf{\Sigma}_{\mathbf{t},\mathbf{t+h}}$[2] denote the first two empirical moments of the matrix $\mathbf{M}_{\mathbf{t},\mathbf{t+h}}$ between the dates t and $t+h$ (with $t > 0$ and $h > 0$).
- Finally, the return of a strategy ϕ at date t is defined as: $r_t^\phi \equiv \phi' \mathbf{r}_{\mathbf{t},.}$. By extension, $\mathbf{r}_{\mathbf{t},\mathbf{t+h}}^\phi$ denotes the vector of daily returns of the strategy ϕ between date t and $t+h$.

6.1.2 Portfolios tested

Several models are considered to predict the asset price returns as detailed in Chapters 2 and 3. More precisely, for each model, the portfolio weights are

[1] $\mu_{\mathbf{t},\mathbf{t+h}} \equiv \overline{\mathbf{M}_{t,t+h}} \equiv \frac{1}{t+h-t+1} \mathbf{M}_{\mathbf{t},\mathbf{t+h}}' \mathbf{1}_{\mathbf{t+h-t+1}}$, where $\mathbf{1_t}$ denotes the unit vector of dimension t.

[2] $\mathbf{\Sigma}_{\mathbf{t},\mathbf{t+h}} \equiv \sigma(\mathbf{M}_{\mathbf{t},\mathbf{t+h}}) \equiv \frac{1}{(t+h-t+1)(t+h-t)} \mathbf{M}_{\mathbf{t},\mathbf{t+h}}' \mathbf{M}_{\mathbf{t},\mathbf{t+h}}$

computed using an estimation window, denoted w. Details are given below on how the different weights are computed; this study focuses on how investors deal with their model ambiguity after the different model portfolio weights have been computed.

6.1.2.1 Single strategies considered

In this sub section, the precise empirical asset allocations that correspond to each one of the simple models considered are provided (i.e. the single priors $\mathbb{Q}$ that constitute the set of priors $\mathcal{Q}$.)

- **The Equally Weighted portfolio (EW)**: gives an equal weight to all the risky assets. The EW portfolio asset allocation is defined as:

$$\phi_t^{EW} = \frac{1}{N_t}\mathbf{1}_{N_t}$$

  This portfolio represents the market benchmark. N_t represents the number of assets considered "active" at time t (i.e. the assets for which at least 50 return observations are available over the estimation window $[t-w : t-1]$). $\mathbf{1}$ stands for the N-vector of ones.

The Minimum Variance and the Maximum Sharpe portfolios as presented in Chapter 2 are considered:

- **The Minimum Variance portfolio (MN)**: is the fully invested Markowitz efficient portfolio with minimum variance, obtained when investors minimise the expected variance of their portfolio. The MN portfolio allocation is defined as:

$$\phi_t^{MN} = \frac{\Sigma_{t-w,t-1}^{-1}\mathbf{1}_{Nt}}{\mathbf{1}_{Nt}'\Sigma_{t-w,t-1}^{-1}\mathbf{1}_{N}}$$

  where $\Sigma_{t-w,t-1}$ is the empirical covariance matrix estimated over the window $[t-w, t-1]$.

- **The Mean Variance portfolio (MV)**: is the fully invested, maximum Sharpe, mean-variance Markowitz efficient portfolio, obtained when the investor maximises the empirical quadratic expected utility. Note that a risk aversion equal to 1 is considered, as in DeMiguel *et al.* (2007). The MV allocation is defined as:

$$\phi_t^{MV} = \frac{\Sigma_{t-w,t-1}^{-1}\mu_{t-w,t-1}}{\mathbf{1}'_{N_t}\Sigma_{t-w,t-1}^{-1}\mu_{t-w,t-1}}$$

  where $\Sigma_{t-w,t-1}$ is the empirical covariance matrix estimated over the window $[t-w, t-1]$ and $\mu_{t-w,t-1}$ is the empirical vector of mean returns estimated over the same window.

  Due to the singularity of the covariance matrix $\Sigma_{t-w,t-1}$ (some of the asset returns are almost collinearly dependant), it is not straightforward to obtain a stable value for the inverted matrix $\Sigma_{t-w,t-1}^{-1}$. A Singular Value Decomposition methodology is used to estimate the empirical inverse of the covariance matrix $\Sigma_{t-w,t-1}$, as shown in the Appendix at the end of the chapter.

- **The CAPM portfolio (CAPM)**: the CAPM portfolio is based upon the Jensen Alphas, as presented in Chapter 2. The CAPM Betas are estimated over the estimation window. Considering $r_{t-w,t-1}^M$ as the vector of the Eurostoxx 600 market returns over the period $[t-w, t-1]$, the Beta of the risky asset, i, is, therefore, estimated at time t as:

$$\forall i \in [1, N],\ \beta_t^i \equiv \frac{\mathbb{COV}(r_{t-w,t-1}^i, r_{t-w,t-1}^M)}{\mathbb{VAR}(r_{t-w,t-1}^M)}$$

  The Jensen Alpha is then computed as the difference between the observed return at time t of the asset i and the Beta adjusted market return:

$$\forall i \in [1, N],\ \alpha_t^i \equiv r_t^i - \beta_t^i r_t^M$$

The CAPM weights are then defined as the weighted average Alphas across all the risky assets considered adjusted by the variance of the CAPM residuals[1]:

$$\forall i \in [1, N],\ \phi_t^{CAPM,i} \equiv \frac{\frac{\alpha_t^i}{\mathbb{VAR}(\alpha^i)}}{\sum_{j=1}^{N} \frac{\alpha_t^j}{\mathbb{VAR}(\alpha^j)}}$$

So that $\sum_i \phi_t^{CAPM,i} = 1$.

- In addition, consideration is given to the factor model portfolios presented in Chapter 3: i.e. the EFM, FFM, PCA, ICA and CA portfolios. More formally, as for the CAPM portfolio, the Jensen Alpha of the different factor models is computed as the difference between the observed return at time t of the asset i and the Beta adjusted factor returns:

$$\forall i \in [1, N],\ \alpha_t^{m,i} \equiv r_t^i - \boldsymbol{\beta}_t^{m,i} \mathbf{F}_t^m$$

Where $\boldsymbol{\beta}_t^{m,i}$ denotes the vector of factor loadings of the asset i for the model $m \in \{EFM, FFM, PCA, ICA, CA\}$ as estimated over the window $[t-w, t-1]$, and $\mathbf{F}_t^m$ the factor returns vector of the model m at date t. Then, the model m weights are defined as the weighted average Alphas across all the risky assets considered, and adjusted by the variance of the residuals:

$$\forall i \in [1, N],\ \phi_t^{m,i} \equiv \frac{\frac{\alpha_t^{m,i}}{\mathbb{VAR}(\alpha^i)}}{\sum_{j=1}^{N} \frac{\alpha_t^{m,j}}{\mathbb{VAR}(\alpha^j)}}$$

6.1.2.2 Combined portfolios

In this sub section, the allocations obtained when considering portfolios that combine the allocations obtained by the different strategies mentioned in the previous sub-section, are exposed.

[1] As suggested in Brennan & Xia (2001), the optimal portfolio is found by scaling the Alpha by its empirical variance

- **The Subjective Expected Utility Portfolio (SEU)**: it is recalled that the SEU portfolio is obtained when mixing different priors when the investor is neutral to ambiguity. If q_t defines the vector of probabilities given to the models considered at any date t (the investor is neutral to ambiguity in the sense that the probabilities of each model are supposed known and well defined), thus, the result is:

$$\sum_{\mathbb{Q}\in\mathfrak{Q}} q_t(\mathbb{Q}) = 1$$

Therefore, the SEU portfolio weights are defined as the different weights of the models linearly weighted by q_t.

$$\forall i \in [1, N],\ \phi_t^{SEU,i} \equiv \sum_{\mathbb{Q}\in\mathfrak{Q}} \phi_t^{\mathbb{Q},i} q_t(\mathbb{Q})$$

- **The Ambiguity Robust Portfolio (RA)**: now, if the investor is averse to ambiguity (i.e. the investor does not know for sure the probability of each prior to occur), the optimal ambiguous portfolio weights are defined as the different weights of the models adjusted by the Absolute Robust Ambiguity Adjustment ψ, calibrated by the absolute ambiguity aversion parameters $(\gamma_t^{\mathbb{Q}})_{\mathbb{Q}\in\mathfrak{Q}}$ and weighted by the Relative Ambiguity Robust Adjustment π_t:

$$\forall i \in [1, N],\ \phi_t^{ARA,i} \equiv \sum_{\mathbb{Q}\in\mathfrak{Q}} \psi(\phi_t^{\mathbb{Q},i}, \gamma_t^{\mathbb{Q}})\pi_t(\mathbb{Q})$$

In this empirical study, a linear form for π is considered, such that:

$$\sum_{\mathbb{Q}\in\mathfrak{Q}} \pi_t(\mathbb{Q}) = 1$$

and in particular, no differentiation is made between the Relative Ambiguity Robust Adjustment, and the classical Subjective Expected Utility probability weights for each model. Therefore, in this particular empirical study, the ARA and SEU allocations only differ because of the Absolute Ambiguity Robust Adjustment through ψ.
In Chapter 7, more complex forms for π are considered.

6.1.3 Performance measures

Once the portfolio weights have been computed for the different models, the returns of the different strategies are computed and evaluated through the four performance measures, details of which are presented in Chapter 2 (Sharpe, Sortino, Win Lose and Gain Loss ratios). For a given strategy ϕ in the period $[t, t+h]$, the performance measure values are estimated as:

- **Sharpe ratio**:

$$Sharpe^{\phi}_{t,t+h} \equiv \frac{\mu^{\mathbf{r}^{\phi}_{\mathbf{t,t+h}}}}{\sigma^{\mathbf{r}^{\phi}_{\mathbf{t,t+h}}}}$$

  where $\mu^{\mathbf{r}^{\phi}_{\mathbf{t,t+h}}}$ stands for the strategy ϕ return between date t and $t+h$ and $\sigma^{\mathbf{r}^{\phi}_{\mathbf{t,t+h}}}$ stands for its standard deviation.

- **Sortino ratio**:

$$Sortino^{\phi}_{t,t+h} \equiv \frac{\mu^{\mathbf{r}^{\phi}_{\mathbf{t,t+h}}}}{\sigma_n^{\mathbf{r}^{\phi}_{\mathbf{t,t+h}}}}$$

  where σ_n stands for the standard deviation of the negative components of a series.

- **Gain Loss ratio**:

$$GainLoss^{\phi}_{t,t+h} \equiv \frac{\sum_{i=1}^{N} r^{\phi,i}_{t,t+h} 1_{r^{\phi,i}_{t,t+h}>0}}{\sum_{i=1}^{N} r^{\phi,i}_{t,t+h} 1_{r^{\phi,i}_{t,t+h}>0} - \sum_{i=1}^{N} r^{\phi,i}_{t,t+h} 1_{r^{\phi,i}_{t,t+h}<0}}$$

  where $\sum_{i=1}^{N} r^{\phi,i}_{t,t+h} 1_{r^{\phi,i}_{t,t+h}>0}$ stands for the sum of positive returns, and $\sum_{i=1}^{N} r^{\phi,i}_{t,t+h} 1_{r^{\phi,i}_{t,t+h}<0}$ stands for the sum of negative returns of the strategy ϕ between the dates t and $t+h$.

- **Winner Loser ratio**:

$$WinLose^{\phi}_{t,t+h} \equiv \frac{\sum_{i=1}^{N} 1_{r^{\phi,i}_{t,t+h}>0}}{\sum_{i=1}^{N} 1_{r^{\phi,i}_{t,t+h}<0} + \sum_{i=1}^{N} 1_{r^{\phi,i}_{t,t+h}>0}}$$

where $\sum_{i=1}^{N} 1_{r^{\phi,i}_{t,t+h}>0}$ stands for the number of positive returns of the strategy ϕ between the dates t and $t+h$, and $\sum_{i=1}^{N} 1_{r^{\phi,i}_{t,t+h}<0}$ stands for the number of negative returns.

All the measures presented above will be used independently to calibrate the ambiguity parameter γ and the ambiguity adjustment function π; the following two measures will be used to compare the different portfolio performances, as in the comparative study of portfolio performances by DeMiguel *et al.* (2007).

- **The Certainty equivalent return**: which corresponds to the equivalent risk-free return of the strategy return:

$$CER_{t,t+h} \equiv \mu^{\mathbf{r}^{\phi}_{\mathbf{t,t+h}}} - \lambda(\sigma^{\mathbf{r}^{\phi}_{\mathbf{t,t+h}}})^2$$

  As in DeMiguel *et al.* (2007), it is assumed that $\lambda = 1$ in the empirical study.

- **The Turnover**: corresponds to the change in portfolio weights from one period to the other; i.e. to the absolute sum of the trades needed to rebalance the portfolio weights from one period to the next. The investor aims at reducing the turnover as trading implies costs (exchange fees, price impact...):

$$T/O_{t,t+h} \equiv \sum_{i=1}^{N} |\phi^i_{t+h} - \phi^i_t| = \delta(\phi_{\mathbf{t+h}}, \phi_{\mathbf{t}})$$

For the Turnover and the Certainty Equivalent ratios, an average daily value is given.
In the following section, the performances of the single portfolio strategies are presented. A more precise outline will then be made of how the absolute ambiguity parameter, γ, and the relative ambiguity parameter, π, are calibrated in practice. Finally, the performances of the classical SEU portfolio and the ARA portfolio are presented and compared. It is shown, in this empirical study, that taking into account ambiguity in the portfolio selection problem, does make a difference in practice, and allows the investor to achieve better performances.

6.2 Calibration and empirical portfolio performances

In this section, the performances of the nine single strategies presented in Section 6.1 are analysed and compared, and the performances of the SEU and ARA portfolios are computed and displayed. First, the performances of the single strategies are presented, initially without taking into account transaction costs and then with the addition of a 3 basis points transaction cost, with respect to the turnover generated by the given daily strategy rebalances. It is found that the performances of the single strategies, post-transaction costs, are very poor and unstable over time. Then, the means by which the absolute ambiguity parameter, γ, and the relative ambiguity parameter, π, are parametrized according to the different four performance measures considered (Sharpe, Sortino, Win Lose or Gain Loss ratios) are established. The focus will then be on the analysis of the linearly blended strategy (SEU) and the ambiguity averse strategy (ARA) performances post-transaction costs, as - to evaluate real performances - it is necessary to take into account those costs. It is found that the SEU strategies improve the single strategies and are more stable over time. Finally, the performances of the ARA portfolios are exposed. It is found that the ARA portfolios consistently beat the SEU portfolios, providing the investor an enhanced and stable performance across the long period considered (from January 2000 to May 2010).

6.2.1 Single portfolios performances

It is assumed that the risk-free rate is negligible, as the portfolios are rebalanced every day[1]. Figures (6.1) and (6.2) plot the cumulative return of the models (EW, MN, MV, CAPM, FFM, EFM, PCA, ICA and CA) over the period January 2000 to May 2010; and Tables (6.1) and (6.2) display the performance statistics of the different strategies across the whole period considered, without and with transaction costs respectively (the worse measures are in red and the best measures are in blue).

[1]For an annual Libor rate of 0.75% as of end of May 2010, the daily rate represents around 0.3 basis point.

	EW	MN	MV	CAPM	FFM	EFM	PCA	ICA	CA
μ(%)	72.27	57.31	72.43	106.68	87.58	16.61	77.12	25.09	147.22
$\overline{\mu}$(Bps)	2.77	2.19	2.77	4.08	3.35	0.64	2.95	0.96	5.64
σ(%)	13.69	12.93	12.41	6.30	8.14	12.06	7.15	12.57	6.30
max(μ)(Bps)	656.54	806.84	766.76	377.46	543.72	647.09	508.51	644.48	365.03
min(μ)(Bps)	−504.58	−707.95	−938.54	−376.17	−362.11	−554.71	−349.11	−504.68	−290.25
Sharpe	0.51	0.42	0.56	1.62	1.03	**0.13**	1.03	0.19	**2.24**
Sortino	0.57	0.52	0.68	2.14	1.39	**0.18**	1.36	0.26	**3.02**
GainLoss(%)	52.52	52.11	52.89	58.09	55.03	**50.65**	55.04	50.95	**61.01**
WinLose(%)	56.47	54.94	52.03	56.40	53.32	50.81	54.40	**50.76**	**58.99**
CER(Bps)	2.39	1.86	2.47	4.00	3.22	**0.34**	2.85	0.64	**5.56**
T/O(%)	**19.33**	114.29	108.03	143.49	150.36	148.89	148.60	148.88	**150.53**

Table 6.1: Strategies Performances No Transaction Costs

It is also assumed that the transaction costs (exchange fees, slippage and so on) account for 3 basis points of the daily portfolio turnover. As a reference, DeMiguel *et al.* (2007) assumed a 0.5 basis point transaction cost per monthly transaction. Note that the 3 basis points transaction cost assumption is very optimistic [1]. On a daily basis, the real cost reaches probably more than this. In the empirical example presented, a transaction cost of 5 basis points (more realistic) kills all the single strategies, and especially the CA, which has the highest turnover. Due to the fundamental empirical importance of transaction costs, a constant 3 basis points transaction cost is assumed for the remainder of the empirical study.

	EW	MN	MV	CAPM	FFM	EFM	PCA	ICA	CA
μ(%)	60.83	−40.36	−23.04	−3.06	−29.52	−97.29	−39.33	−90.69	32.07
$\overline{\mu}$(Bps)	2.33	−1.55	−0.88	−0.12	−1.13	−3.72	−1.51	−3.47	1.23
σ(%)	13.69	12.91	12.50	6.30	8.14	12.06	7.14	12.57	6.29
max(μ)(Bps)	656.03	805.77	766.45	373.50	539.07	643.67	503.37	640.99	361.20
min(μ)(Bps)	−505.07	−710.86	−948.33	−380.61	−367.13	−560.71	−355.11	−510.23	−295.45
Sharpe	0.43	−0.30	−0.18	−0.05	−0.35	**−0.77**	−0.53	−0.69	**0.49**
Sortino	0.48	−0.37	−0.22	−0.06	−0.48	**−1.07**	−0.71	−0.95	**0.68**
GainLoss(%)	52.13	48.51	49.09	49.77	48.31	**46.23**	47.42	46.59	**52.43**
WinLose(%)	**55.35**	50.98	48.66	47.90	46.69	46.11	47.40	**45.75**	50.47
CER(Bps)	**1.95**	−1.88	−1.19	−0.20	−1.26	**−4.02**	−1.61	−3.79	1.15
T/O(%)	**19.33**	114.29	108.03	143.49	150.36	148.89	148.60	148.88	**150.53**

Table 6.2: Strategies Performances 3 bps Transaction Costs

It is recalled that μ stands for the total strategy return over the whole period in percentage. $\overline{\mu}$ stands for the average daily return in basis points and σ is the standard deviation of the strategy returns.

Without transaction costs, the CA portfolio has the best overall performance, and the EFM portfolio, the worse, in terms of all performance measures. However, when transaction costs are considered, all the single strategies are greatly

[1]Note that this transaction cost assumption is chosen; otherwise most of the single strategies would not give any positive returns and, therefore, the empirical study would not mean much.

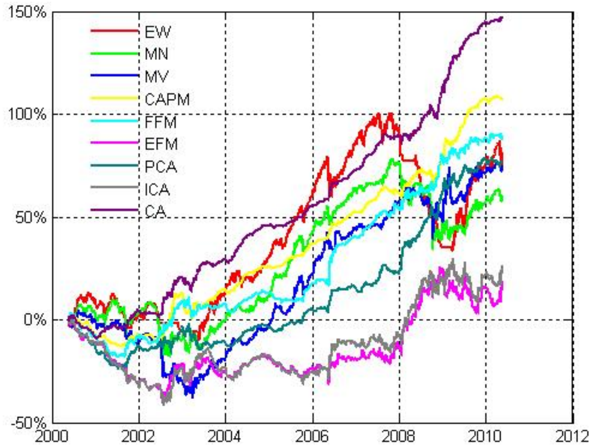


Figure 6.1: Cumulative Strategies Returns (%) without transaction costs

Figure 6.2: Cumulative Strategies Returns (%) with 3 basis points transaction costs

penalised due to their high turnover (to the exception of the EW strategy, that by construction has a very low turnover) and especially the CA strategy that has the highest turnover (above 150%). Indeed, the amount of shares traded to rebalance the different portfolios every day, entails transaction costs: the higher the turnover, the higher the costs, and; therefore, the more penalised the strategy returns.

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
EW	**1.54**	−0.58	−1.13	**1.76**	2.45	1.57	0.42	−3.69	**1.98**	−0.44
MN	**0.75**	**0.00**	−0.77	0.75	**2.82**	2.49	1.52	−1.51	0.98	−0.04
MV	**0.01**	−1.28	−1.30	**0.87**	2.46	2.10	1.57	−0.42	1.28	**0.23**
CAPM	−1.88	−0.01	**2.29**	0.18	**3.89**	**4.70**	**2.35**	1.45	**3.51**	−0.57
FFM	−2.07	**0.17**	**2.03**	−0.13	1.21	**2.63**	**2.08**	**1.82**	0.82	−0.28
EFM	−2.15	−1.88	0.84	−0.32	0.18	0.36	0.55	**1.75**	−0.45	0.10
PCA	−3.55	−0.09	1.06	−0.84	2.36	2.26	1.52	**3.93**	1.40	−0.26
ICA	−2.11	−1.97	0.63	−0.05	−0.12	1.15	0.28	1.32	0.06	**0.52**
CA	−2.47	**2.18**	**1.90**	**1.07**	**4.36**	**4.42**	**3.85**	2.02	**4.10**	**1.45**

Table 6.3: Sharpe per Strategy per Period no Transaction Costs

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
EW	**1.41**	**−0.71**	−1.25	**1.68**	**2.34**	**1.50**	**0.35**	−3.75	**1.94**	**−0.50**
MN	**−0.28**	**−0.62**	−1.25	−0.75	**1.76**	1.36	0.34	−1.88	0.38	−0.98
MV	**−1.02**	−2.32	−1.90	**0.20**	**1.62**	0.82	**0.37**	−1.06	0.74	−0.79
CAPM	−3.65	−1.88	**1.04**	−1.28	0.08	**1.46**	−0.39	0.41	**2.19**	−3.41
FFM	−3.70	−1.38	**0.60**	−1.42	−1.68	1.30	0.28	**0.92**	−0.22	−1.89
EFM	−3.45	−3.01	−0.20	−1.33	−1.65	−0.65	−0.37	**1.17**	−1.12	**−0.52**
PCA	−5.49	−1.84	−0.13	−2.13	−1.08	−0.22	−1.16	**2.72**	0.39	−2.43
ICA	−3.39	−3.18	−0.37	−1.08	−1.90	0.19	−0.64	0.75	−0.54	**−0.12**
CA	−4.50	**0.25**	**0.55**	**−0.43**	0.49	**1.53**	**1.21**	0.91	**2.69**	−1.03

Table 6.4: Sharpe per strategy per periods 3 bps Transaction Costs

Most of the single strategies are, also, highly unstable overtime. In Tables (6.3) and (6.4)[1], the annual Sharpe ratio has been computed for the different strategies. It can be seen, that, depending on the years, the best strategies differ (the three

[1]The three best Sharpe measures per period are highlighted in bold.

best yearly strategies are never the same from year 2001 to year 2010, for raw performances as well as for transaction cost adjusted ones). Therefore, it is crucial for investors to diversify their investments among different strategies, as has been outlined in the previous chapter.

As the performance of the different models varies greatly overtime, how can investors achieve the best mix for their asset allocation? It has to be kept in mind, that, ex ante, the investors do not know which of the different single strategies will perform best. Therefore, investors need to consider the different models in order to define the preferred asset allocation. In the following section, the performance measures of two different portfolios are presented, taking into account the whole set of prior models: the classical SEU portfolio that linearly blends the models and is neutral to model ambiguity; and the ARA portfolio approach developed in this research, that takes into account model ambiguity expressed through the investor's ambiguity aversion.

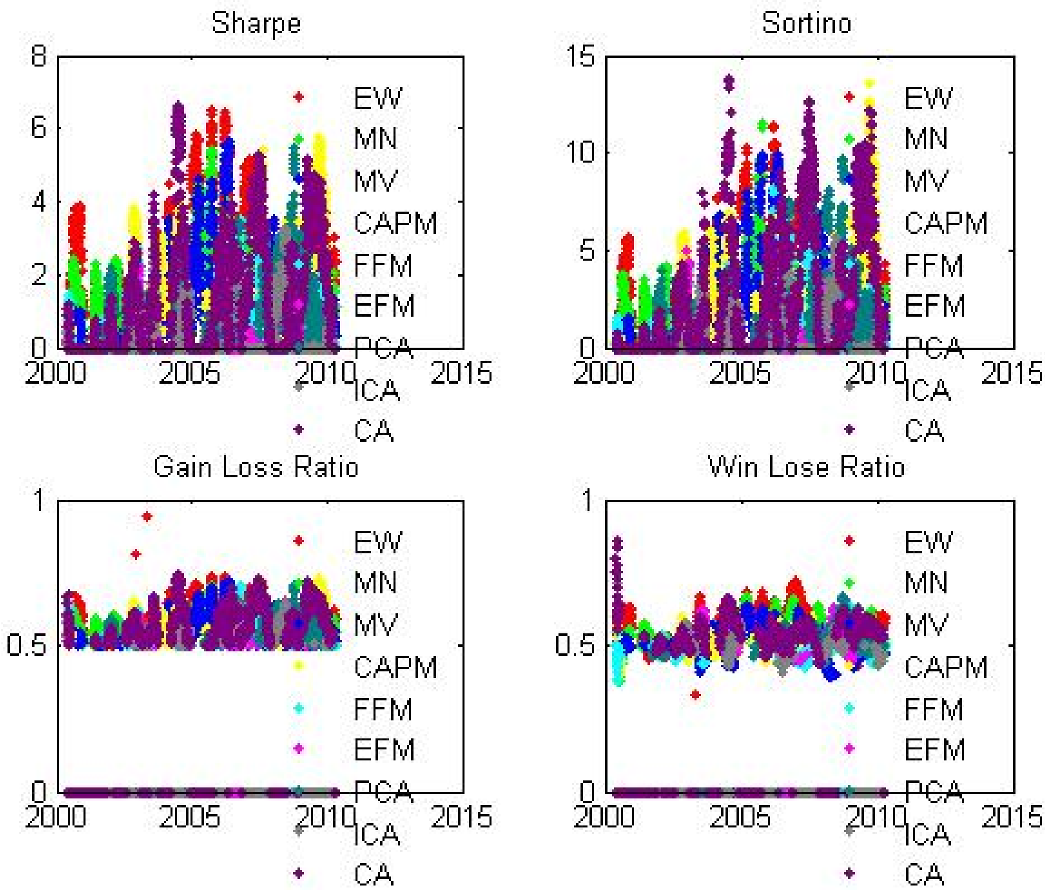
6.2.2 Calibration of the absolute ambiguity parameter γ and the relative ambiguity adjustment π

To compute the SEU and ARA asset allocations, the ambiguity parameters, γ and π, presented in Chapter 5, need to be calibrated. Thus, a methodology that links the ambiguity aversion parameters with the performance measures of the single portfolios considered is proposed.

Figure (6.3) plots the different values of the performance measures over time.
It can be seen that the performance of the different strategies is very volatile, as also shown in the year on year performances of the single strategies in Table (6.4). Also, it should be noted that the investor is not supposed to favour one particular strategy and can not know, ex ante, which one of those strategies will perform well in the future. This, therefore, implies that the adjustment of each models should be dynamic.
More precisely, to calibrate γ and π, an estimation window of $w \equiv 100$ business days (equivalent to approximately five months)[1] is used to estimate the portfolio

[1] A minimum number of observations is needed to reliably estimate model parameters, 100 data points are sufficient to obtain good asymptotic properties of the model parameters esti-

Figure 6.3: Performance measures



weights: Typically, at each date t, the performance measures as presented above are computed based upon daily returns computed over the window $[t-w; t-1]$. Note that the Sharpe and Sortino values are annualised (their daily value is multiplied by $\sqrt{250}$, as there are around 250 business days per calendar year), and floored to zero (all negative Sharpe and Sortino ratios are defaulted to zero). The Win Lose and Gain Loss ratios are floored to 50%; all measures below this threshold are also defaulted to zero. Indeed, if the Sharpe or Sortino ratios are negative over the past window considered, the corresponding strategies have had a particularly bad performance, and, therefore, it is assumed that the investor will not invest in this strategy for the following period. Similarly, if the Gain Loss or Win Lose ratios are under 50%, the corresponding strategy has lost more than it gained over the considered period, and it is assumed that the investor will not invest in this strategy for the following period, either. $PM_t^{\mathbb{Q}}$ denotes a given performance measure for the model $\mathbb{Q}$ over the window $t-w, t-1$, where PM stands for one of the performance measures (Sharpe, Sortino, GainLoss, WinLose) described above. Note that the CER and T/O measures are not used for parameterisation, but only for performance comparison as in DeMiguel *et al.* (2007). More formally, the following parameterisation for $\gamma_t^{\mathbb{Q}}$ and $\pi_t(\mathbb{Q})$ is considered:

- The absolute ambiguity parameter $\gamma_t^{\mathbb{Q}}$ is estimated as the negative inverse absolute past performance measure of the portfolio computed from the model $\mathbb{Q}$ (the worst the past performance of the portfolio, the bigger the absolute ambiguity aversion):

$$\gamma_t^{\mathbb{Q}} \equiv -\frac{1}{PM_t^{\mathbb{Q}}} \tag{6.1}$$

  The absolute ambiguity parameter $\gamma^{\mathbb{Q}}$ is absolute in the sense that it is specific to each model $\mathbb{Q}$, and does not depend on the performance of the other models in the set of priors $\mathcal{Q}$.

mates.

- The relative ambiguity parameter $\pi^{\mathbb{Q}}$ (that effectively depends on the whole set of models $\mathcal{Q}$ considered) is estimated as the relative performance measure of the model $\mathbb{Q}$ in the class $\mathcal{Q}$ of models considered:

$$\pi(\mathbb{Q}) \equiv \frac{PM^{\mathbb{Q}}}{\sum_{\mathbb{P}\in\mathcal{Q}} PM^{\mathbb{P}}} \tag{6.2}$$

The measure π is relative, as it takes into account the performances of all the different models; whereas γ is absolute, as it solely considers the performance of a given model. Both the relative ambiguity adjustment π and the absolute ambiguity parameter γ are proportional to the performance measure considered: the higher the performance measure, the higher the parameter. Also, if $PM_t^{\mathbb{Q}} = 0$ (i.e. as expressed above, if the Sharpe or Sortino is negative or if the Gain Loss or Win Lose ratio is below 0%) then $\pi_t^{\mathbb{Q}} = 0$ and $\gamma_t^{\mathbb{Q}} = 0$. Indeed, at worse the investor will not invest in a strategy at all, and all the weights are defaulted to zero (if either the absolute aversion parameter or the relative ambiguity parameter is null, the subsequent weights of the SEU or ARA portfolio are defaulted to zero).

Note that it is for simplicity that the choice has been made to parametrize π and γ similarly; what matters is that the absolute and relative ambiguity aversion are positively correlated with the performance measure considered. Other ways to parametrize γ or π that do not depend on the past performance of single strategies could also be considered. However, empirically, this simple calibration method makes sense and can be used for the empirical study carried out in this research.

6.2.3 The SEU portfolio performance

DeMiguel *et al.* (2007) use the out of sample Sharpe ratio as well as the CER and Turnover to compare 14 different optimised allocations of portfolios (rather than the allocation of individual stocks that have been proposed here) and the equally weighted portfolio monthly (rather than daily) performances. They find that none of the classical optimised portfolios outperform the basic EW portfolio, significantly, in terms of Sharpe, CER or Turnover measures. Similar results are

found and it is shown that the SEU portfolio outperforms all the single strategy allocations (EW included).
In Table (6.5), the performances and statistics are computed for four SEU portfolios, where the probability of each model is defined as the weighted average of one of the four performance measures: Sharpe ratio, Sortino ratio, Win Lose or Gain loss ratios. Figure (6.4) displays the returns of the different SEU strategies, with a transaction cost of 3 basis points. All the SEU portfolios outperform the single strategies in terms of CER (SEU strategies have higher CER: 1.66 to 1.95 against -3.79 to 1.95) and Turnover (SEU strategies have lower Turnover: 115% on average against 140% on average for single strategies, with the exception of the naive market representative EW strategy).

	sharpe	sortino	gainloss	winlose
$\mu(\%)$	55.32	51.28	47.80	49.10
$\overline{\mu}$(Bps)	2.12	1.96	1.83	1.88
$\sigma(\%)$	9.30	9.04	9.26	9.31
$\max(\mu)$(Bps)	656.03	656.03	656.03	656.03
$\min(\mu)$(Bps)	-408.09	-408.09	-408.09	-408.09
Sharpe	0.57	0.54	0.49	0.50
Sortino	0.69	0.66	0.60	0.61
GainLoss(%)	52.73	52.62	52.39	52.44
WinLose(%)	53.58	53.48	52.96	53.12
CER(Bps)	1.94	1.80	1.66	1.71
T/O(%)	113.38	115.09	117.08	116.84

Table 6.5: SEU Strategies Performances 3 bps Transaction Costs

Where the four different columns correspond to the calibration of the parameters γ and π with respect to respectively the Sharpe, Sortino, Gain Loss and Win Lose ratios.

Figure 6.4: Subjective Expected Utility Strategies Cumulative Returns

Mixing the different models based on their past performance measures enhance investor performance, especially because it allows the investor to reduce the turnover, and; therefore, the total transaction costs. But also, it allows the investors to smooth their performances over time, as can be seen in the SEU strategies annual Sharpe values displayed in Table (6.6): the different SEU strategies display a positive Sharpe for 6 years on average, whereas single strategies display a positive Sharpe in only one to five years (expect for the CA and EW strategies).

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
sharpe	−0.03	−1.00	**0.41**	−0.89	**2.06**	**1.66**	**0.93**	**0.39**	**1.29**	−1.37
sortino	**0.03**	−1.00	**0.41**	−0.91	**2.05**	**1.70**	**0.90**	**0.31**	**1.30**	−1.43
gainloss	−0.45	−0.89	**0.69**	−0.95	**1.98**	**1.50**	**0.61**	**0.52**	**0.99**	−1.35
winlose	−0.43	−0.87	**0.72**	−0.97	**1.99**	**1.42**	**0.56**	**0.59**	**0.98**	−1.29

Table 6.6: Sharpe per SEU strategy per periods

However, further improvement can be made to the performance of the investor portfolio with the Ambiguity Robust Adjustment approach presented here, which is demonstrated in the following section.

6.2.4 The ARA portfolio performance

In this subsection, the ARA portfolio performances are presented. An emphasis is made of the fact that taking account of ambiguity affects, empirically, affects the performance of the portfolio positively: the portfolio performance is better and the portfolio returns are more stable overtime. Figure (6.5) plots the performance of the four different ARA portfolios (where the absolute ambiguity parameter γ and the relative ambiguity adjustment π are estimated through the four different performance measures considered: Sharpe, Sortino, Gain Loss and Win Lose ratios).

Figure 6.5: Ambiguity Robust Adjusted Strategies Cumulative Returns

In Table (6.7), the statistics of the various ARA portfolios have been computed. The ARA strategies beat all the single strategies with transaction costs in terms of Sharpe, Sortino and CER, and they have a reduced Turnover (except for the EW strategy, which by construction has a very low turnover).

More importantly, the ARA strategies also outperform the SEU strategies. In Table (6.7), the Relative Ambiguity Adjustment (RAA) measure, as presented in Chapter 5, is computed, which represents the turnover between the SEU portfolio and the ARA portfolio, and gives the amount of wealth invested in the risk free asset (or buffer) due to the investor Absolute Ambiguity aversion[1]. The RAA

[1]As $\sum \pi = 1$. In the example used here, the ARA solely differs from the SEU portfolio by

	sharpe	sortino	gainloss	winlose
$\mu(\%)$	63.34	57.65	55.12	58.30
$\overline{\mu}$(Bps)	2.43	2.21	2.11	2.23
$\sigma(\%)$	9.23	8.75	9.69	9.91
max(μ)(Bps)	656.03	656.03	656.03	656.03
min(μ)(Bps)	-408.09	-408.09	-408.09	-430.06
Sharpe	0.66	0.63	0.54	0.56
Sortino	0.81	0.79	0.66	0.69
GainLoss(%)	53.20	53.10	52.69	52.77
WinLose(%)	52.91	52.77	52.91	52.83
CER(Bps)	2.25	2.05	1.92	2.04
T/O(%)	99.84	104.60	108.21	107.42
AAA	0.70	0.62	0.91	0.92
RAA	0.59	0.53	0.73	0.73

Table 6.7: ARA Strategies Performances 3 bps Transaction Costs

is around 60%: the ARA and SEU portfolios effectively differ a lot. These allocation differences allow the ARA portfolios to achieve a much lower turnover than the SEU portfolios (around 100% against 115%). The different ARA portfolios outperform all the SEU portfolios, in terms of CER and Turnover: all ARA strategies CER are above, or close to, 2, whereas all SEU strategies CER are well below 2.

As shown, specifically, in the summary Table (6.8), where a comparison of the two SEU and ARA Sharpe strategies is displayed, the ARA strategy beats the SEU strategy in terms of all the performance measures considered (but for the Win Lose ratio). In particular, the Sharpe and Sortino are improved by 15% and 17% respectively. Also, in comparison to the SEU strategy, the ARA strategy improves the two benchmark measures greatly (the CER is improved by almost 16% and the Turnover is reduced to under 100%, 12% less than in the SEU case). Note that similar results are found for the other performance measures: Sortino, Gain Loss or Win Lose ratios.

The annual Sharpe is also computed for each of the ARA strategies in Table (6.9). It can be seen that the performance of the ARA strategies is more stable over

the Absolute Ambiguity Robust Adjustment: the ARA weights are shrunk by the function ψ, parametrized by the absolute ambiguity parameter γ.

time than the single strategies, as only the CA and EW strategies have a positive Sharpe over more than six years, whereas it is the case for all SEU or ARA strategies.

	ARA	SEU	Diff(%)
μ(%)	63.34	55.32	14.51
$\overline{\mu}$(Bps)	2.43	2.12	14.51
σ(%)	9.23	9.30	−0.71
max(μ)(Bps)	656.03	656.03	0.00
min(μ)(Bps)	−408.09	−408.09	0.00
Sharpe	0.66	0.57	**15.32**
Sortino	0.81	0.69	**17.51**
GainLoss(%)	53.20	52.73	0.88
WinLose(%)	52.91	53.58	−1.24
CER(Bps)	2.25	1.94	**15.92**
T/O(%)	99.84	113.38	**−11.94**

Table 6.8: SEU and RA Sharpe Strategies Comparison

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
sharpe	**0.53**	−1.52	**0.03**	−0.97	**1.94**	**1.59**	**0.21**	**0.68**	**1.48**	−1.19
sortino	**0.51**	−1.54	−0.03	−0.99	**1.98**	**1.73**	**0.26**	**0.53**	**1.60**	−1.32
gainloss	**0.29**	−1.30	**1.07**	−0.69	**1.73**	**1.11**	−0.12	**0.63**	**1.36**	−1.25
winlose	**0.36**	−1.26	**1.15**	−0.80	**1.75**	**0.98**	−0.03	**0.68**	**1.34**	−1.11

Table 6.9: Sharpe per ARA strategy per periods

6.3 Conclusion

In this empirical study, it has been shown that investors, facing a choice between different models to explain stock returns, enhance their portfolio returns with combined strategies, such as the SEU or ARA strategies. Moreover, the ARA ambiguity averse portfolio improves the SEU portfolio, significantly, in terms of CER and Turnover, but also in terms of Sharpe, Sortino, Gain Loss and Win Lose ratios. With the ARA methodology, an improved alternative to the SEU strategy is presented, that is easy to implement in practice; robust, in terms of performances, and also very flexible. It has been shown that taking into account ambiguity gives better results than a basic blending of models in an empirical study conducted over a very long period of time, encompassing two major financial crisis (2001 and 2008). Note that in this empirical example, a very simple linear function, π, is used in order to easily compare the SEU and ARA portfolios. In the following chapter, a more elaborate version of the relative ambiguity adjustment, π, is presented by introducing some nonlinearity.

6.4 Appendix: Estimation of the empirical inverse of the covariance matrix

In practice, the estimation of the covariance matrix is a difficult task (see Ledoit & Wolf (2004)), as the covariance matrix can be close to singular. Indeed, the eigenvalues of the covariance matrix can be very small, and, therefore, the inversion of the covariance matrix can become problematic. Actually, in order to estimate the covariance matrix $\mathbf{\Sigma}$ of the excess returns, , a Singular Value Decomposition procedure is adopted (the time indices are dropped to simplify notations in this section). More precisely, the matrix , and is found, such that:

$$\Sigma = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^T$$

Where $\boldsymbol{D}$ is a diagonal matrix with elements $(\lambda_i)_{1\leq i\leq N}$ and $\boldsymbol{U}^T = \boldsymbol{U}^{-1}$ and $\boldsymbol{V}^T = \boldsymbol{V}^{-1}$. $\mathbf{\Sigma}^{-1}$ is estimated to be:

$$\mathbf{\Sigma}^{-1} = \boldsymbol{V}\boldsymbol{D}^{*}\boldsymbol{U}^T$$

Where $\boldsymbol{D}^*$ is a diagonal matrix with elements $(\lambda_i^*)_{1\leq i\leq N}$, such that:

$$\lambda_i^* = \begin{cases} \frac{1}{\lambda_i} & \text{if } \lambda^i > \epsilon \\ 0 & \text{otherwise.} \end{cases}$$

where ϵ is the singularity threshold, which is set to 0.1 in this study, to keep the percentage of null eigen values below 10%.

Chapter 7

Nonlinear Relative Ambiguity Adjustment

"*I know nothing except the fact of my ignorance.*"

Socrates, cited in Diogenes Laertius, Lives of Eminent Philosophers.

In this Chapter, the novel Ambiguity Robust Adjustment methodology presented in Chapter 5is enhanced by considering a none linear form for Relative Ambiguity Robust Adjustment π. As has been shown in the previous chapter, the Ambiguity Robust Adjustment (ARA) can allow an investor to enhance their portfolio allocation performance on real empirical data. However, the linear form considered for the Relative Ambiguity Robust Adjustment, π, does not take into account nonlinear effects that can be produced by ambiguity adjustment. For instance, aggregated asset allocations could be capped, and penalised to a greater degree if models widely disagree; a cash buffer can be set aside if overall ambiguity aversion is high. The impact of the nonlinearity of the function, π, on the performance of the ARA portfolio will be tested on real data.

It is recalled that the function π represents the relative adjustment for a given model among the class of models considered by the investor, and accounts for the blending of the different models taken into account by the investor. Once the different model outputs have been computed and adjusted for Absolute Ambiguity Robust Aversion through the function ψ, the different adjusted solutions must

be combined to compute the final portfolio allocation. A linear form for π, along with the fact that $\sum_{\mathbb{Q}\in\mathfrak{Q}} \pi_{\mathbb{Q}} = 1$, relates to the Subjected Expected Utility (SEU) framework: the different models are given independently a fixed weight[1], that is applied for the blending allocation of all the assets.
On contrary to the prior chapters, it is no longer assumed, in this chapter, that π is linear. More theoretically, the ARA weights are defined as:

$$\forall i \in [1, N],\ \phi^{ARA} \equiv \pi[\psi(\phi^{\mathbb{Q}_1,i}, \gamma^{\mathbb{Q}_1}), ..., \psi(\phi^{\mathbb{Q}_p,i}, \gamma^{\mathbb{Q}_p})]$$

In this context, the function π is more flexible, and can account for a number of effects that cannot be expressed through a linear function

- Weight dispersion: first, the function π can operate a nonlinear blending of the different models, or class of models, amplifying the weights of assets when several models agree, and further reducing the weights of assets for which the models disagree.
- Precautionary principle: also, the final asset weights can be capped in order to express the fact that an investor, averse to ambiguity, is always reluctant to invest fully in any risky asset.
- Global ambiguity aversion: finally, the function π can partly control the buffer the investor decides to invest in a cash reserve, due to aversion to ambiguity.

The calibration of the function π is more an art than an exact science; quantitative, as well as more qualitative, methods can be used to estimate π. In the remainder of this Chapter, two different methods to calibrate π are presented: a non-parametric quantitative method, stemming from the neuronal network area, and a more ad hoc, qualitative methodology, where the function π is constructed with respect to some desired nonlinear properties.

The chapter is organised as follows: in the first section, a statistical non-parametric method, that can be used to estimate π, is presented: the Support Vector Machines (SVM). The SVM methodology is explained, more theoretically, and then

[1]In this case it corresponds to a probability.

applied in the second sub-section to the same European dataset used in the empirical study in Chapter 6. However, it is shown that the empirical application of the SVM method on real data can be challenging: non-parametric, numerical, methods such as the SVM have some major drawbacks, including computational time, and cannot necessarily be applied efficiently to large datasets, calling for alternative methods. In the second section, a more ad hoc, nonlinear form for π is considered that respects some desired nonlinear properties. Then, an empirical application is proposed that respects the properties described above. It is shown that this ad hoc nonlinear form for π can greatly improve the ARA portfolio performances. The third section concludes the chapter.

7.1 A nonlinear, non-parametric method to estimate π: the Support Vector Machines (SVM)

Assuming the investor has no prior knowledge about the form of the function π, the investor needs, therefore, a non-parametric methodology, to calibrate it on empirical data. The machine learning area has developed numerous methods for nonlinear estimations, for instance learning trees or genetic algorithms. Those neural network algorithms have the advantage of being able to approximate nonlinear functions without a priori assumptions about the data.

7.1.1 The SVM theory

Here, a new, non-parametric machine learning method is presented, which has been proposed recently by Vapnik (1995) to empirically parametrize the function π. Many authors found that SVM can beat other classical machine learning methods to forecast stock returns (see for instance Cao & Tay (2001) who compare the SVM technique to a classical back propagation neural network on S&P 500 data, using a Gaussian kernel SVM, or Kim (2003), who makes the same comparison on Korean data).

7.1.1.1 Basic theory: linear SVM

Here, an overview of the SVM technique is given. For an in depth description, refer to Smola & Scholkopf. (1998).
A set of N data $(x_i, y_i)_{1\leq i\leq N}$ is considered, where the $x_i \in \mathbb{R}^D$, are some inputs (these can be multidimensional, i.e. $D > 1$), and the $y_i \in \mathbb{R}$ are the outputs. The novel aspect of the SVM is that it seeks to minimise an upper bound of the estimation error, rather than minimising the estimation error itself. Indeed, the goal of SVM is to find a function $f(.)$ that has, at most, a deviation ϵ to the outputs $(y_i)_{1\leq i\leq N}$, i.e.:

$$\forall i,\ |f(x_i) - y_i| \leq \epsilon \tag{7.1}$$

The function f is defined as:

$$f(x) =< \omega, x > +\omega_0$$

where $< a, b >\equiv \sum_{d=1}^{D} a_d b_d$ refers to the scalar product.
The SVM separates the input vector, $(x_i)_{1\leq i\leq N}$, through hyperplanes, in such a way that the separated data (or decision classes) are as far as possible from each other. A way to achieve this goal, is to minimise the Euclidian norm of the parameter vector $||\omega||^2 \equiv< \omega, \omega >$. The smaller the norm of the vector ω, the "flatter" the function f is considered to be. Therefore, SVM is aimed at solving the following quadratic programming optimisation problem:

$$\min \frac{1}{2} < \omega, \omega > \tag{7.2}$$

$$\text{subject to, } \forall i \in [1, N], \quad \begin{cases} y_i - < \omega, x_i > -\omega_0 \leq \epsilon \\ < \omega, x_i > +\omega_0 - y_i \leq \epsilon \end{cases}$$

The constraints reflect the fact that it is desirable to respect inequality 7.1.

However, Problem 7.2 may be unfeasible when no function f can approximate all pairs, (x_i, y_i), with precision ϵ. One can introduce slack variables, (ν_i, η_i), and a soft margin loss function parametrized by a parameter c. Problem 7.2 becomes:

$$\min \frac{1}{2} < \omega, \omega > +c \sum_{i=1}^{N} (\nu_i + \eta_i) \tag{7.3}$$

$$\text{subject to } \forall i \in [1, N], \begin{cases} y_i - < \omega, x_i > -\omega_0 \leq \epsilon + \nu_i \\ < \omega, x_i > +\omega_0 - y_i - \leq \epsilon + \eta_i \\ \nu_i, \eta_i \geq 0 \end{cases}$$

The parameter c determines the trade-off between the flatness of f and the amount by which deviations that are larger than ϵ are tolerated. A dual formulation of 7.3 leads to the following maximisation quadratic programme:

$$\max -\frac{1}{2} \sum_{i,j=1}^{N} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) < x_i, x_j > -\epsilon \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) + \sum_{i=1}^{N} y_i(\alpha_i - \alpha_i^*) \tag{7.4}$$

$$\text{subject to } \forall i \in [1, N], \begin{cases} \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) = 0 \\ \alpha_i, \alpha_i^* \in [0, c] \end{cases}$$

Where the (α_i, α_i^*) are Lagrange multiplicators and $\omega = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) x_i$ (the vector parameter ω can be completely described as a linear combination of the inputs x_i). The function f is, therefore, defined as:

$$f(x) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) < x_i, x > +\omega_0$$

and the hyperplanes $(\alpha_i - \alpha_i^*) < x_i, x > +\omega_0$, for such i where $\alpha_i > 0$ and $\alpha_i^* > 0$ are effectively the Support Vectors that separate in classes the inputs x_i.

7.1.1.2 Nonlinear SVM

The SVM algorithm has the property of being entirely defined by scalar products of the inputs. This allows the algorithm to be made nonlinear by simply preprocessing the inputs. The goal of the nonlinear SVM is to estimate the following function:

$$f(x) =< \omega, \pi(x) > +\omega_0$$

Where π is a nonlinear function of the input x. A kernel, k, that transforms the original inputs, $(x_i)_{1 \leq i \leq N}$, is considered. The fact that k can be nonlinear allows the SVM to operate nonlinear estimations. Because the SVM algorithm only depends on dot products, it, therefore, suffices to know $k(x,y) \equiv < \pi(x), \pi(y) >$, instead of the nonlinear transformation function $\pi(.)$. Common examples of the kernel function are the polynomial kernel: $k(x,y) = (xy+1)^d$ and the Gaussian kernel: $k(x,y) = \exp^{-\frac{(x-y)^2}{\delta^2}}$. In the empirical example displayed here, Gaussian kernels will be considered, as the polynomial kernel requires longer time in the training of the SVM algorithm (as pointed out by Kim (2003)).

7.1.2 Empirical application: nonlinear ARA portfolio calibrated with the SVM algorithm

In this section, the ARA portfolio allocation is constructed when the RARA function is estimated through the SVM methodology, and is applied to the European dataset previously used in Chapter 3 and Chapter 6. In a first sub-section, the framework for the empirical application is defined, then the results are presented and discussed.

7.1.2.1 Framework and calibration of the SVM algorithm

In order to conduct an empirical test on real data, it is necessary to specify the framework of the test. Indeed, some simplifications are required in order to run an SVM algorithm on a very large data sample. The precise calibration of the SVM algorithm parameters is also discussed.

To compute the SVM analysis, the MATLAB 7.5.0 package developed by Canu *et al.* (2008) is used. As an input to the SVM algorithm, the outputs of the different 9 models previously studied are used (The Equally Weighted (EW), Minimum Variance (MN), Mean Variance (MV), CAPM, External Factor Model (EFM), Fundamental Factor Model (FFM), Principal Component Analysis (PCA), Independent Component Analysis (ICA) and Cluster Analysis (CA) portfolios). The idea is to find the nonlinear function π that achieves the best mix of single strategy outputs to explain stock returns. Therefore, the idea is to solve the following nonlinear regression:

$$\phi_t^{i,ARA} \equiv \pi(\phi_{t-1}^{i,EW}, \phi_{t-1}^{i,MN}, \phi_{t-1}^{i,MV}, \phi_{t-1}^{i,CAPM}, \phi_{t-1}^{i,EFM}, \phi_{t-1}^{i,FFM}, \phi_{t-1}^{i,PCA}, \phi_{t-1}^{i,ICA}, \phi_{t-1}^{i,CA}) + \epsilon_t^i$$

In order to find the relevant support vectors at each date t, the SVM algorithm is computed over a sample rolling window of $w \equiv 100$ days, i.e. a sample covering the time interval $[t-w, t-1]$. Due to computational time issues, the more explanative variables are considered the longer the algorithm runs. The nine different single strategies are, therefore, aggregated in three relevant groups in order to consider only three explanative variables to calibrate the SVM (as a reference Cao & Tay (2001) consider only 5 variables to explain the returns of some financial index futures). The three subsequent aggregated models are, therefore, considered:

- Classical portfolio (CP): defined as the average of the EW, MN and MV portfolios.
- Exogenous factor model portfolio (EP): defined as the average of the CAPM, EFM and FFM portfolios.
- Statistical factor model portfolio (SP): defined as the average of the PCA, ICA and CA portfolios.

The allocations obtained for each aggregated model, over the window considered, are concatenated across all stocks. Finally, the SVM algorithm is run in order to calibrate the following nonlinear equation:

$$sign(\boldsymbol{\phi}_t^{1:N,ARA}) \equiv \pi(\boldsymbol{\phi}_t^{1:N,CP}, \boldsymbol{\phi}_t^{1:N,EP}, \boldsymbol{\phi}_t^{1:N,SP}) + \boldsymbol{\epsilon}_t$$

where $\boldsymbol{\phi}_t^{1:N,ARA}$ stands for the vector of all risky assets weights at date t, and the $\boldsymbol{\phi}_t^{1:N,\cdot}$ represents the vectors of the aggregated CP, EP and SP portfolio allocations.

To estimate π_t at each date t, the concatenated vectors of stock returns and CP, EP and SP portfolio allocations over the time interval $[t-w, t-1]$ are considered. Note that the SVM algorithm does not give a prediction for the returns themselves, but for the sign of the returns. Indeed, it is computationally

too challenging to run an SVM algorithm that can predict the scale of returns for such a large dimension problem.
As pointed out by Cao & Tay (2001), the SVM algorithm is very sensitive to the calibration of the Gaussian kernel parameter, δ^2, the loss function parameter, c, and the deviation parameter, ϵ. To calibrate these parameters, a validation test was conducted on a sub-sample of the dataset (i.e. the last 110 days of the data set considered: 100 days being considered for the in sample calibration of the SVM algorithm, and the last 10 days being used as an out of sample validation set), where a range of different values have been tested for the three parameters considered. The set of parameters that gives the best Hit ratio[1] for the out of sample dataset is selected[2]:

$$\begin{cases} \delta^2 = 0.01 \\ c = 10 \\ \epsilon = 0.3 \end{cases}$$

Ideally, one may want to re-calibrate the three parameters, δ^2, c and ϵ, for each day where the SVM is re-computed. However, due to long computation time (each SVM in sample calibration over a 100 day window takes between one and three minutes), it is not possible to re-calibrate those parameters at each step of the empirical study. Therefore, for this study, the decision has been made to keep those parameters constant over the whole period considered.

Now that the framework of the SVM calibration has been specified, the results of a nonlinear ARA portfolio performance, when the function π is estimated through the SVM technique are displayed and discussed in the next section.

[1] number of good signed predictions divided by the total number of predictions

[2] The data ranges chosen for each parameter are as follows:

$$\begin{cases} \delta^2 \in [0.01 : 0.1] \text{ with a step of } 0.01 \\ c \in [10 : 100] \text{ with a step of } 10 \\ \epsilon \in [0.1 : 1] \text{ with a step of } 0.1 \end{cases}$$

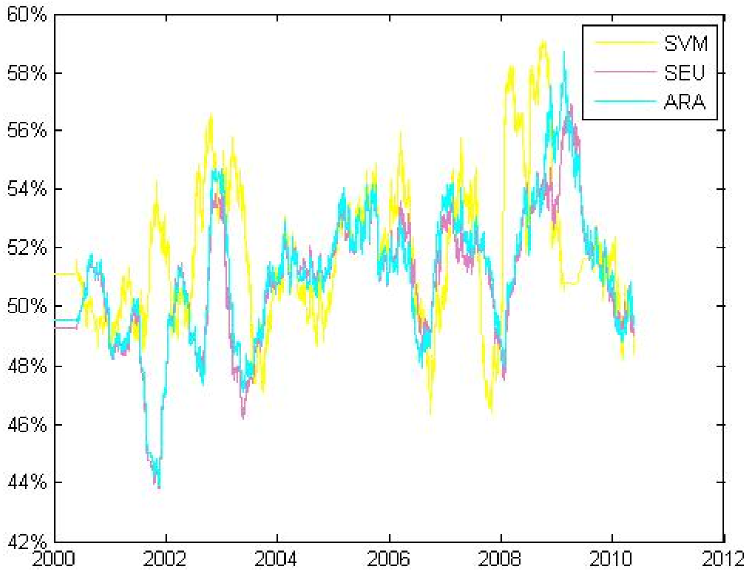
7.1.2.2 Comparison of the SVM generated ARA portfolio performance against the SEU and linear ARA portfolio performances

The performance of predictions is evaluated using the Hit ratio (HIT) metrics, as considered in Chapter 6. The Hit ratio for a given portfolio P (where P is either the Subjected Expected Utility portfolio - denoted SEU, the ARA portfolio, where a simple linear form for π is considered - denoted ARA - as described in Chapter 6 or the ARA portfolio where a nonlinear form for π is evaluated through the SVM technique and denoted SVM) is defined as:

$$HIT_{t+1}^{P} \equiv \frac{\sum_{i=1}^{N}\{sign(r_t^i) = sign(r_t^{P,i})\}}{N}$$

The Hit ratio effectively represents the percentage of times the sign of the predicted return is equal to the sign of the realised return. Consequently, the higher the Hit ratio is, the better. In Figure (7.1), the Hit ratios for the three considered portfolios, SEU, ARA linear and SVM, are plotted.

Figure 7.1: SEU, ARA and SVM strategies Hit ratios



The Hit ratio of the SVM portfolio is higher on average than that of the SEU or linear ARA portfolios (around 52% of the time, the SVM Hit ratio is higher than the ARA or SEU Hit ratios). The SVM technique, allowing for a consideration of a nonlinear form for the Relative Ambiguity Robust Adjustment, π, improves the ARA portfolio performance in terms of Hit ratio, i.e. the SVM technique allows a better prediction of the signs corresponding to future risky asset returns. However, many drawbacks have to be taken into account when implementing the SVM algorithm on real empirical data.

7.1.2.3 Main drawbacks of the SVM algorithm to calibrate π

The non-parametric SVM technique is very flexible in the form the function π can take. However, the main issue with the SVM algorithm is the computation time (the algorithm considered takes several minutes to run to estimate a one day prediction over an in sample data set of 100 days). Because the SVM algorithm is highly dependent on the specification of the parameters, c, ϵ and δ^2, the risk of data mining (where the algorithm is over-fitted for the in sample data set) is high. Also, for the large dataset, as the one considered here, the number of support vectors necessary to fit the Gaussian kernel to the in sample data can be very large as well, making the SVM approach not very tractable (when numerous support vectors are necessary to fit the function π, it makes it difficult to have a good apprehension of the estimated form of the function π); this challenges some of the original motivations for developing the ARA methodology, i.e. simplicity and tractability.
Such drawbacks have led to serious consideration of a more ad hoc approach, where the form of the function π is pre defined according to its desired properties.

7.2 A more ad hoc method to calibrate a nonlinear form for π

In this section, the approach to calibrate a nonlinear form for π, to compute the ARA final allocation, is inverted: some prior nonlinear properties, desired by the investor, are discussed and the form of the Relative Ambiguity Adjustment π is deduced from them. More precisely, the approach is as follows. In the first subsection, three nonlinear properties are described in detail, which are the product of ambiguity aversion. In the second subsection, an ad hoc, nonlinear form for

π, fulfilling those properties, is proposed and tested on real data. The third subsection concludes.

7.2.1 Some nonlinear desired properties of the Relative Ambiguity Robust Adjustment π

The function π is a universal function applied to all the different models taken into account by the investor. The nonlinear properties of the function π should allow the investor to express their relative aversion to ambiguity. A number of desirable properties are required for the function π, so that the nonlinearity of π reflects some economical rationalities expressed by the investor. In particular:

- The weight dispersion
- The precautionary principle
- The global ambiguity aversion

7.2.1.1 Weight dispersion across the different models

A desirable characteristic of the function π is that it should shrink the weights of the assets on which the different classes of models tend to disagree. The concavity of the function π should increase for weights for which the models disagree.
To give a more precise idea of this theoretical property, two assets i and j are considered. A measure of dispersion v is considered (for instance, the standard deviation, but it can also be the variance, the absolute mean deviation...). The dispersion and mean of all the Absolute Ambiguity Adjusted weights $(\psi(\phi^{i,\mathbb{Q}}, \gamma))_{\mathbb{Q}\in\mathcal{Q}}$ for any asset i are denoted v^i and μ^i. The weight dispersion property for the function π is defined as:

Property 7.1 (Weight dispersion). *If $v^i < v^j$ (meaning that the models disagree more on asset j than on asset i), then the final ARA weight $\phi^{i,ARA}$ should be closer relatively to the mean μ^i than the weight $\phi^{j,ARA}$ to the mean μ^j i.e:*

$$v^i < v^j \Rightarrow \frac{|\phi^{i,ARA} - \mu^i|}{|\mu^i|} < \frac{|\phi^{j,ARA} - \mu^j|}{|\mu^j|}$$

7.2.1.2 Precautionary principle

If the models disagree widely on a given asset allocation, the investor should require the function π, to set the questionable asset allocation to zero. The investor is able to set a threshold, v^{max}, such that if the dispersion of the different models considered is above this threshold for a given asset, the final robust ARA weight should be set to zero. More precisely:

Property 7.2 (Precautionary principle)**.** *If for a given asset i, $v^i > v^{max}$, then* $\phi^{i,ARA} = 0$

Note that the dispersion measure, v, can be quantitative (for instance the standard deviation), but it also can be qualitative. Indeed, an investor may consider more qualitative measures for dispersion; for instance, the number of agreeing or disagreeing experts.

7.2.1.3 The global ambiguity aversion

Finally, the ARA methodology allows the investor to compute a cash buffer that represents the amount of investment money not allocated to the risky assets due to the investor's ambiguity aversion. The first main conceptual difference between the ARA and the SEU blending is that the Relative Ambiguity Adjustment, π, is not necessarily a probability distribution, in the sense that the different models weights, $\pi(\mathbb{Q})$, do not necessarily sum to one. Indeed, investors may consider that the combination of all their models remains insufficient to explain stock returns, and, therefore, allows a proportion of their wealth to not be invested in risky assets. In the case when π is a linear function, and when $\sum_{\mathbb{Q}\in\mathcal{Q}} \pi(\mathbb{Q}) < 1$, the amount $1 - \sum_{\mathbb{Q}\in\mathcal{Q}} \pi(\mathbb{Q})$ is disinvested from the risky assets to the risk-free asset, due to Relative Ambiguity Aversion of the investor. Depending on the context, this buffer can increase and decrease. During uncertain periods (as for instance during financial crisis), the buffer gets more important, whilst during more stable times it can shrink back. This RAA is, therefore, precisely measurable, and can for instance set the rules to define the amount of reserves required by a risk management policy: if financial markets become more risky, the amount of investment wealth set aside should increase; and vice versa. As an analogy, the

RAA is similar, conceptually, to the capital required for a given level of VaR. Although in this case, the RAA is not model dependant, and is, therefore, more flexible than the VaR methodology.
As already mentioned in Chapter 5 (where a linear form of π is considered) the cash buffer is more formally defined in this case as:

Property 7.3 (cash buffer)**.** *The ARA cash buffer ρ is defined as:*

$$\rho \equiv \phi^{0,ARA} - \pi(\phi^{0,\mathbb{Q}^1}, ..., \phi^{0,\mathbb{Q}^q})$$

where the set of priors $\mathcal{Q} \equiv \{\mathbb{Q}^1, ..., \mathbb{Q}^q\}$.
The ARA cash buffer can be decomposed into two parts: the AARA cash buffer for each prior $\mathbb{Q} \in \mathcal{Q}$ denoted $\rho^{\mathbb{Q}}$ and the RARA cash buffer for the set of priors $\mathcal{Q}$ denoted $\rho^{\mathcal{Q}}$:

$$\begin{cases} \forall \mathbb{Q} \in \mathcal{Q},\ \rho^{\mathbb{Q}} \equiv \psi(\phi^{0,\mathbb{Q},\gamma^{\mathbb{Q}}}) - \phi^{0.\mathbb{Q}} \\ \rho^{\mathcal{Q}} \equiv \rho - \pi(\rho^{\mathbb{Q}_1}, ..., \rho^{\mathbb{Q}_q}) \end{cases}$$

The cash buffer represents the global ambiguity aversion of an investor. It corresponds to the difference between the specific allocation given to the risk-free asset considered as a "refuge value" in the ARA portfolio allocation and the Relative Ambiguity Adjusted combined allocations of the risk-free asset under the different models, when only ambiguity towards the set of models is considered. Note that the cash buffer represents, effectively, an overall cap on the sum of total asset allocations for the ARA portfolio.

Those three properties are not exhaustive, and an investor could come up with additional properties specific to a given allocation problem. In the next section, a very tractable, ad hoc, form of π is considered, that fulfils all the desired properties presented above.

7.2.2 Ad hoc nonlinear form for π for the ARA allocation

As specified in the preceding section, where the ARA allocation is computed through the SVM algorithm, non-parametric methods to calibrate nonlinear functions are often very computationally heavy (the processing time required to re-compute portfolio allocations is often important when considering real-life trading constraints, when portfolios are re-balanced every day or even several times a day). That is the reason why a more ad hoc version of π is proposed, here, that proves better performing with empirical data than the linear form of π considered in Chapter 6. In the first sub section, the details of the computation, of π, is presented. Then, the performance of the resulting "nonlinear" ARA portfolio (calibrated with respect to the Sharpe of the different single strategies) is displayed in comparison to the equivalent "linear" ARA portfolio performance, computed in Chapter 6.

7.2.2.1 Prior specification of the function π

A simple nonlinear form of π can be considered to fulfil the three properties of: weight dispersion, precautionary principle and global ambiguity aversion. The generic form can be expressed as:

$$\forall i \in [1, N], \phi^{ARA,i} \equiv \max\{\phi^{max}, \frac{\mu_{\mathbb{Q}\in\mathcal{Q}}[\psi(\phi^{\mathbb{Q},i}, \gamma^{\mathbb{Q}})]}{v_{\mathbb{Q}\in\mathcal{Q}}[\psi(\phi^{\mathbb{Q},i}, \gamma^{\mathbb{Q}})]}\}$$

where ϕ^{max} is a maximum cap value for any ARA allocation $\phi^{ARA,i}$, reflecting the precautionary principle and also the cash buffer, as the sum of all ARA risky allocations will also be capped. $\mu_{\mathbb{Q}\in\mathcal{Q}}[\psi(\phi^{\mathbb{Q},i}, \gamma^{\mathbb{Q}})]$ describes the average AARA transformed allocation of the single strategies $\mathbb{Q} \in \mathcal{Q}$ for the asset i, and $v_{\mathbb{Q}\in\mathcal{Q}}[\psi(\phi^{\mathbb{Q},i}, \gamma^{\mathbb{Q}})]$ represents a dispersion value of the AARA transformed allocation of the single strategies $\mathbb{Q} \in \mathcal{Q}$ for the same risky asset i, translating the weight diversification property. Indeed, the higher $v_{\mathbb{Q}\in\mathcal{Q}}[\psi(\phi^{\mathbb{Q},i}, \gamma^{\mathbb{Q}})]$, the smaller the absolute allocation $\phi^{ARA,i}$.

More specifically, considering a cap value of 1 and the variance as dispersion measure, the final nonlinear ARA allocation can be defined as:

$$\forall i \in [1, N], \phi^{ARA,i} \equiv \max\{1, \frac{\mu_{\mathbb{Q}\in\mathcal{Q}}[\psi(\phi^{\mathbb{Q},i}, \gamma^{\mathbb{Q}})]}{\sigma^2_{\mathbb{Q}\in\mathcal{Q}}[\psi(\phi^{\mathbb{Q},i}, \gamma^{\mathbb{Q}})]}\} \tag{7.5}$$

Note that in this case, the nonlinear allocation, $\boldsymbol{\phi}^{ARA}$, is capped by one (no single asset allocation can be greater than one in absolute terms).

7.2.2.2 Results and Comments

The nonlinear version of the ARA portfolio will now be empirically tested, and its performance compared to the linear version of the ARA portfolio, as described in Chapter 6. Note that the nonlinear ARA portfolio is constructed with a similar methodology, and the same dataset as the one employed in Chapter 6 to define the linear ARA allocation.
Two portfolios are considered here:

- The ARA portfolio as described in Chapter 6, where a simple linear form for π is considered, where priors AARA transformed allocations are linearly weighted according to past performance measures, and where the aversion parameters are calibrated with respect to the Sharpe ratio performance measure (denoted "sharpe" thereafter).
- The ARA portfolio, where a nonlinear form for π, as described in Equation 7.5, is considered (denoted "nonlinear" thereafter).

Note that for the "nonlinear" ARA portfolio, the AARA transformation and the calibration of the aversion parameters are specified as for the "sharpe" portfolio. Note also, that the decision has been made to only display the ARA portfolios that have been constructed with a calibration of the absolute ambiguity parameter γ according to the Sharpe measure. This choice is for the sake of result clarity, only, since a calibration with the other performance measures (Sortino, Win Lose or Gain Loss ratios) gives similar results.

As can be seen in Figure 7.2, the nonlinear, ad hoc version of π outperforms the linear form considered in Chapter 6: the overall total return of the nonlinear version of the ARA portfolio is higher (85% overall return against only 65%). Consequently, of course, the nonlinear ARA portfolio outperforms all single strategies and the SEU strategy, as well. Furthermore, the linear form of the ARA portfolio displays a negative cumulative return up until 2004, which is not the case for the nonlinear ARA portfolio: the nonlinear ARA portfolio prevents the investor from being subjected to a drawdown, offering more robust and stable performance than the linear ARA portfolio.

The detailed statistics, presented in Table 7.1, confirm the graphical observation: the nonlinear ARA portfolio performs better than the linear ARA portfolio in terms of Sharpe, Sortino, Gain Loss and especially CER ratios (3.23 against 2.39), for a comparable level of turnover (104% against 99%). The nonlinear ARA portfolio offers, therefore, a much better remuneration for risk than the linear ARA portfolio.

Figure 7.2: Ad hoc nonlinear Sharpe ARA strategy versus linear Sharpe ARA strategy cumulative returns

7.3 Conclusion

Ambiguity aversion implies some nonlinear properties that can be taken into account, not only, through the Absolute Robust Ambiguity Adjustment ψ, but, also through the Relative Ambiguity Robust Adjustment π. Indeed, the more models actually agree on the allocation for a given risky asset, the less the investor should allegedly be averse to it, and vice versa. Also, the investor averse to ambiguity may apply a precautionary principle, and cap final risky asset allocations due to global ambiguity aversion towards the set of models.

	non linear	sharpe
$\mu(\%)$	85.88	63.34
$\overline{\mu}$(Bps)	3.29	2.43
$\sigma(\%)$	12.18	9.23
max(μ)(Bps)	820.39	656.03
min(μ)(Bps)	-504.85	-408.09
Sharpe	0.67	0.66
Sortino	0.86	0.81
GainLoss(%)	53.43	53.20
WinLose(%)	52.84	52.91
CER(Bps)	3.23	2.39
T/O(%)	104.32	99.84

Table 7.1: Non Linear Strategy Performance

When Investigating nonlinear forms for π, empirical evidence shows that it can further enhance portfolio allocation performances: the overall return of the non-linear ARA portfolios, presented in this Chapter, is much higher than the one for a linear ARA portfolio, as the one considered in Chapter 6. Also, the Certainty Equivalent Ratio obtained with a nonlinear ARA portfolio is 35% higher than the CER obtained with the linear ARA portfolio.

Non-parametric numerical methods, which have the advantage of not assuming any a priori form for π, prove, however, challenging to use in practice due to computational time limitations and lack of tractability for the solution selected. Another more practical approach, proposed in this chapter, is to come up with a prior form for π that respects some desired, selected, nonlinear properties. It has been shown, empirically, that this version of π allows a better performing ARA asset allocation than the ones considered previously.

This chapter is a first step; the ad hoc form for π, proposed in this chapter, is by no means a unique, or best solution. It is solely shown, that, via an empirical test, this approach can make a contribution to a better way of allocating assets. Further research is clearly needed to investigate the potentialities of nonlinearity for the RARA function; such as deciding, which nonlinear properties to consider and how to take them into account.

Chapter 8

Conclusion

"*This is not the end. It is not even the beginning of the end. But it is, perhaps, the end of the beginning.*"

Sir Winston Churchill.

In this PhD thesis, a new methodology to account for model ambiguity has been proposed: the Ambiguity Robust Adjustment (ARA). This novel approach differs from classical approaches found in the literature, as it focuses on finding a robust, tractable and flexible solution, rather than an optimal solution, sensu stricto. The methods developed in this research have, thus far, proved to be an easy to implement, robust ambiguity methodology that allows investors to adapt asset allocation decisions to the level of ambiguity aversion held against the different priors considered when modelling asset return dynamics. Through this approach, the investor is able to distinguish between two types of ambiguity: a model specific absolute ambiguity aversion, and relative ambiguity aversion across the set of different priors. Contrary to the classical approaches offered in the literature, the ARA can be applied to complex, high-dimension problems. In particular, empirical studies performed on financial data have shown that the ARA methodology greatly improves the performance of an asset allocation problem solution: the ARA portfolio allocation outperforms the single strategy allocations consistently; as well as the SEU allocation, which operates a plain linear blending of the single strategies. It has also been shown, that, given a more complex nonlinear form for the RARA function, the ARA portfolio performance can be further enhanced, giving scope for significant improvement in the portfolio allocation problem facing financial practitioners.

The novel ARA approach yields better portfolio returns when applied to the asset allocation problem. Mixing models by scaling down extreme asset weights through the AARA function and adjusting the different models allocations through the RARA function allows the investor to smooth the portfolio performance better than when mixing models with a simple linear approach as the SEU that does not take into account ambiguity. It is not to say though that this new methodology is the ultimate solution to deal with ambiguity. There may be other, more effective ways to account for ambiguity in the decision making process. However, the benchmark approach used in this PhD thesis, comparing the SEU method (used as a benchmark, not taking into account ambiguity) to the ARA method (taking into account ambiguity) shows that the treatment of ambiguity made by the later novel approach improves the decision making process. Note that this benchmark approach may remain insufficiant to irrevocably prove that the ARA treatment of ambiguity is the sole cause of improved portfolio performance, other hidden effects could have been overlooked, and this could constitute material for further research.

The research conducted in this PhD thesis has led to numerous questions, and by no means have they been all addressed in this PhD thesis. Further research will almost certainly involve reflections on the following topics:

- Calibration of the absolute and relative ambiguity aversion parameters.
- Other forms for the RARA function.
- Application of the ARA methodology on other fields.

More specifically:

- Testing different forms for the Absolute Robust Ambiguity Adjustment function ψ and also the Relative Ambiguity Robust Adjustment π; indeed, throughout this PhD thesis, only a few forms have been investigated for those two functions. Research should be undertaken, especially, into other nonlinear calibrations for π . For instance, deterministic models such as those stemming from Chaos Theory could be tested; other stochastic approaches could also be used. For instance, many non-parametric neural network algorithms could be considered (the neural network field, with genetic or learning algorithm, has evolved immensely over the past few years), even if numerically challenging, they could offer a good alternative to calibrate the function π. In addition, parametric models, such as the Gaussian Mixture of models, could also be investigated, even though they require a strong prior (i.e. the decomposition of π in different normal distributions). Some qualitative approaches could also be used when blending different models under ambiguity, as has been proposed in political sciences, environmental policies, biology ...

- Another area that requires further research is the calibration of the absolute and relative ambiguity parameters. In this research, a method based on the knowledge of past model performances is used. However, it could be assumed that decision-makers are able to define their aversion to ambiguity in ways that are more qualitative.

- Finally, the ARA methodology should be tested on decision-making problems in areas other than the financial field, which has been the main consideration in this thesis. Indeed, this method could be applied to many areas involving decision-making under uncertainty. For instance, the ARA could be used, not only to characterise robust investment strategies, but also, to design and establish specific risk-management regulation for pension funds, that - although currently underdeveloped - are crucial to the avoidance of dramatic impacts of financial market crashes at a larger scales of the economy. Collective decision-making is another important application of this research proposal. Climate change policy and environmental regulation, more generally, constitute a direct application area. Macroeconomic and monetary policies, where robustness is now the object of an expanding research community, should also benefit from the decision method proposed here.

Because of the increasing complexity of modelling requirements (due to the evolution of modern technology as well as the progression of many different scientific fields) and the diversity of prior models that can be considered to represent various decision variables, it became clear that it was important to consider an adaptable approach to account for ambiguity. This constituted the main motivation to undertake this PhD thesis; the aim has been to introduce an original methodology to account for ambiguity in the decision-making process, generally speaking, in ways that are robust, flexible and tractable.

References

Al-Najjar, N. & Weinstein, J. (2009). The ambiguity aversion literature: A critical assessment. *Economics and Philosophy*, **25**, 249–284. 120

Allais, M. (1953). Le comportement de lhomme rationnel devant le risque, critique des postulats et axiomes de lecole amricaine. *Econometrica*, **21**, 503546. 103

Artzner, P., Delbaen, F., Eber, F. & Heath, J. (1999). Coherent measures of risk. *Mathematical Finance*, **9**, 203–228. 108, 117

Back, A.D. & Weigend, A.S. (1997). A first application of independent component analysis to extracting structure from stock returns. *International Journal of Neural Systems*, **8**, 473–484. 86, 89

Basak, S. (2005). Asset pricing with heterogeneous beliefs. *Journal of Banking and Finance*, **29**, 2849–2881. 110, 114

Bernardo, A.E. & Ledoit, O. (2001). Gain, loss and asset pricing. *Journal of Political Economy*, **108**, 144–172. 40

Black, F. & Litterman, R. (1990). Asset allocation: combining investor views with market equilibrium. *Goldman Sachs Fixed Income Research*. 115

Brennan, M.J. & Xia, Y. (2001). Assessing asset pricing anomalies. *Review of Financial Studies*, **14**, 905–942. 111, 112, 114, 159

Canu, S., Grandvalet, Y., Guigue, V. & Rakotomamonjy, A. (2008). *SVM and Kernel Methods Matlab Toolbox*. 184

Cao, L. & Tay, F. (2001). Financial forecasting using support vector machines. *Neural Computing & Applications*, **10**, 184–192. 181, 185, 186

Cvitanić, J., Lazrak, A., Martellini, L. & Zapatero, F. (2006). Dynamic portfolio choice with parameter uncertainty and the economic value of analysts recommendations. *Review of Financial Studies*, **14**, 1113–1156. 112, 114

DeMiguel, V., Garlappi, L. & Uppal, R. (2007). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The Review of Financial Studies*, **108**, 144–172. 40, 51, 96, 118, 119, 158, 162, 164, 169, 170

Dickey, D. & Fuller, W. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, **74**, 427–431. 62

Ellsberg, D. (1961). Risk, ambiguity and the savage axiom. *Quarterly Journal of Economics*, **75**, 643–669. 102, 105, 124

Epstein, L.G. & Schneider, M. (2007). Learning under ambiguity. *Review of Economic Studies*, **74(4)**, 1275–1303. 116, 133

Epstein, L.G. & Schneider, M. (2008). Ambiguity, information quality and asset pricing. *The Journal of Finance*, **63(1)**, 197–228. 113, 114, 117

Fabozzi, F., Kolm, P., Pachamanova, D. & Focardi, S. (2007). Robust portfolio optimization. *The Journal of Portfolio Management*, **33**, 40–48. 108, 120

Fama, E. & French, K.R. (1993). Common risk factors in the returns of stock and bonds. *Journal of Financial Economics*, **33**, 3–56. 75

Fama, E.F. & French, K.R. (1992). The cross-section of expected stock returns. *Journal of Finance*, **47**, 427–465. 38, 47, 50, 61

Fama, E.F. & French, K.R. (1998). Value versus growth: The international evidence. *Journal of Finance*, **53**, 1975–2000. 47

FÖLLMER, H. & SCHIED, A. (2002). Convex measures of risk and trading constraints. *Finance and Stochastics*, **6**, 429–447. 107, 108, 117

GARLAPPI, L., UPPAL, R. & WANG, T. (2009). Portfolio selection with parameter and model uncertainty: A multi-prior approach. *Review of Financial Studies*, **20 (1)**, 41–81. 115

GILBOA, I. & SCHMEIDLER, D. (1989). Maxmin expected utility with a non-unique prior. *Journal of Mathematical Economics*, **18**, 141–153. 106, 115

HANSEN, L., SARGENT, T.J., TURMUHAMBETOVA, G. & WILLIAMS, N. (2006). Robust control and model misspecification. *Journal of Economic Theory*, **128**, 45–90. 117

HANSEN, L.P. & SARGENT, T.J. (2001). Robust control and model uncertainty. *American Economic Review*, **91(3)**, 60–66. 108, 115

HYVRINEN, A. & OJA, E. (2000). Independent component analysis: Algorithms and applications. *Neural Networks*, **13**, 411–430. 88, 89

JENSEN, M.C. (1969). Risk, the pricing of capital assets and the evaluation of investment portfolios. *Journal of Business*, **42**, 167–242. 26, 31, 47, 49, 58, 61, 63, 65, 96

JORION, P. (1986). Bayes-stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis*, **21**, 279–292. 115

KAHNEMAN, D. & TVERSKY, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, **47**, 263–292. 105

KIM, K. (2003). Financial time series forecasting using support vector machines. *Neurocomputing*, **55**, 307–319. 181, 184

KLIBANOFF, P., MARINACCI, M. & MUKERJI, S. (2005). A smooth model of decision making under ambiguity. *Econometrica*, **73**, 1849–1892. 117, 121, 122, 124, 125, 129, 136

KNIGHT, F.H. (1921). *Risk, Uncertainty, and Profit*. Hart, Schaffner & Marx, Houghton Mifflin Company, Boston. 8, 100

KOGAN, L., T. & WANG (2002). A simple theory of asset pricing under model uncertainty. *working paper*. 116

LEDOIT, O. & WOLF, M. (2004). Honey, i shrunk the sample covariance matrix. *Journal of Portfolio Management*, **31**. 178

LEVIN, A. & WILLIAMS, J. (2003). Robust monetary policy with competing reference models. *Journal of Monetary Economics*, **50**, 945–975. 2

LONGIN, F. & SOLNIK, B. (2001). Extreme correlation of international equity markets. *Journal of Finance*, **56**, 649–676. 39, 55, 57, 68

MACCHERONI, F., MARINACCI, M., A. & RUSTICHINI (2006). Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, **74**, 1447–1498. 117

MARKOWITZ, H.M. (1952). Portfolio selection. *Journal of Finance*, **7**, 77–91. vi, 7, 9, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24

MERTON, R.C. (1973). An intertemporal capital asset pricing model. *Econometrica*, 867–887. 21, 61, 109, 114

NEHRING, K. (2009). Coping rationally with ambiguity: Robustness versus ambiguity aversion. *Economics and Philosophy*, **25**, 303–334. 120

PASTOR, L. (2000). Portfolio selection and asset pricing models. *Journal of Finance*, **55**, 179–223. 115

PHILLIPS, P.C.B. & PERRON, P. (1988). Testing for unit roots in time series regression. *Biometrika*, **75**, 335–346. 63

ROSS, S.A. (1976). Portfolio selection. *Journal of Economic Theory*, **13**, 341–360. vii, 7, 9, 33, 34, 35, 36, 37, 71

SAVAGE, L.J. (1954). *The Foundations of Statistics*. Wiley, New-York. 101, 102, 123, 146

SHARPE, W.F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, **19**, 425–442. vi, 7, 9, 25, 26, 27, 28, 29, 30, 31, 32, 50

SHARPE, W.F. (1994). The sharpe ratio. *The Journal of Portfolio Management*, 49–58. 19, 38

SMOLA, A. & SCHOLKOPF., B. (1998). A tutorial on support vector regression. *Technical Report , Neuro Colt Royal Holloway College*, **TR-1998-030**. 182

SORTINO, F. & PRICE, L. (1994). Performance measurement in a downside risk framework. *The Journal of Investing*, **3**, 59–65. 39

STEFEK, D. (2002). The barra integrated model. *http://www.barra.com/products/pdfs/IntModel.pdf*. 74

THEIL, H. (1971). *Principles of Econometrics*. John Wiley, New York. 80

TOBELEM, S. & BARRIEU, P. (2010a). *Alternative Investments and Strategies. Chapter 13: Robust Asset Allocation Under Model Risk.*. World Scientific. 134

TOBELEM, S. & BARRIEU, P. (2010b). *The Risk Modeling Evaluation Handbook: Rethinking Financial Risk Management Methodologies in the Global Capital Markets. Chapter 30: Robust Asset Allocation Under Model Risk.*. Mc Graw Hill. 140

VAPNIK, V. (1995). *The Nature of Statistical Learning*. Springer-Verlag, New York. 181

WANG, Z. (2005). A shrinkage approach to model uncertainty and asset allocation. *Review of Financial Studies*, **18(2)**, 673–705. 115

WU, G. & GONZALEZ, R. (1999). Nonlinear decision weights in choice under uncertainty. *Management Science*, **45**, 74–85. 105