UNIVERSITA' DEGLI STUDI DI MILANO BICOCCA

Corso di Dottorato in Matematica Applicata ai Mercati Finanziari (Ciclo XXII)

HIGHER MOMENTS ASSET ALLOCATION

Relatore: Prof.ssa Maria Elena De Giuli

Coordinatore del Dottorato: Prof. Giovanni Maria Zambruno

Tesi di Dottorato di:

Pierpaolo Uberti

Anno Accademico 2009-2010

Aknolegements: I thank Professor A. G. Malliaris for the useful suggestions on the subject of this dissertation.

Contents

List of Figures

List of Tables

Introduction

Modern portfolio theory has been proposed in early 1950's with seminal studies of Markowitz, see Markowitz (1952). In this framework the deterministic calculus of maximization of the agent's utility under budget constraints is not adequate. The attention is focused on choices made under uncertainty. The probabilistic notion of expected return and risk become central. The fundamental instrument for the development of such a theory of choices made under uncertainty is given by the definition of expected utility functions, see Von-Neumann Morgenstern (1947). The second fundamental contribution of Markowitz in portfolio selection analysis is to point out the role of diversification. In other words, the optimal portfolio is the one that minimize the portfolio variance under a certain level of portfolio expected return, where the diversification helps to reduce portfolio variance by investing in securities with low return covariances.

Markowitz portfolio selection model fundamentally relies on the assumption that asset returns distributions are Gaussian or, equivalently, that the investors preferences are described by a quadratic utility function. However, it has long been recognized that financial assets returns are non-normal. While normal distributions are entirely described by the first two moments (mean and variance), asymmetric and fat-tailed distributions are not. Despite this, Markowitz approach has lead to very good results for a wide range of situations. Empirical evidence has shown that the asset allocation obtained with the mean-variance model is very similar to the one obtained by directly maximizing the expected utility, see Levy and Markowitz (1979). An explanation of the good results of the mean-variance criterion could be that the asset returns show elliptical distributions. As a consequence, the expected utility function can be approximated just by taking into account the first two moments of returns distribution.

However, when assets returns distribution strongly depart from the Normal distribution, also elliptical distributions seem to be inadequate in order to model assets returns. Recent research highlights the importance to consider higher moments in asset allocation models. Hwang and Satchell (1999) show that economical agents have preferences for positive skewness and avoid high level of kurtosis. Moreover, empirical studies on assets returns point out that skewness and kurtosis significantly differ from the one of the Normal distribution, see for example Campbell and Siddique (1999), Hwang and Satchell (1999), Fang and Lai (1997).

For asset allocation models, Kraus and Litzemberger (1976), Simaan (1993) and Gamba and Rossi (1998) showed how to consider mean, variance and skewness in the asset allocation scheme. Simaan and Gamba and Rossi approaches differ because of the choice of the objective function to optimize: Simaan minimizes portfolio variance under skewness constraint while Gamba and Rossi directly maximize portfolio skewness. As a consequence, in Simaan approach the sign of the term that describe systematic skewness is not given. Gamba and Rossi approach simply resembles Kraus and Litzemberger model without making assumptions on the agent utility function. In addition, Gamba and Rossi provide a closed form solution for the optimization problem.

In recent years Jondeau and Rockinger (2003), Jondeau and Rockinger (2006), Jondeau and Rockinger (2009) proposed a four moment allocation model. They underline the importance of considering moments in asset allocation problems up to the fourth order. This approach is different from the previous ones because the attention is focused on modelling conditional moments while a closed form solution for the optimization problem is not provided. A individual utility function is defined and the optimal portfolio is derived through a numerical solution procedure.

The aim of the dissertation is to generalize Gamba and Rossi model in order to take into account the third and the fourth moment of returns distribution. The general idea behind the model is to require a structure of financial returns that permits to solve analytically the optimization problem. To do this, it is assumed that asset returns can be separated into the spherical and the non-spherical components. The problem to solve is a non-linear optimization problem where the portfolio kurtosis is minimized under constraints on expected return, variance, skewness (and the usual budget constraint). A structure of preferences for higher moments is proposed such that investors have preferences for odd order moments and dislike even order moments. Due to the assumptions on the preferences structure, the economical agent can choose portfolios with a higher variance compared to mean-variance efficient ones. Moreover, the extra variance of the portfolio is counterbalanced by higher levels of skewness and lower levels of kurtosis. The proposed model is a generalization of Markowitz mean-variance model and of Gamba and Rossi mean-variance-skewness model. Let underline that a closed form solution of the optimization problem is obtained together with a four funds separation property. The optimal portfolio results as the sum of three portfolios: the mean-variance optimal portfolio plus two arbitrage portfolios that reflect skewness and kurtosis preferences.

The dissertation is organized as follows. In chapter 1 the classical mean-variance model is derived after a briefly introduction on expected utility theory. A derivation of the CAPM model is also provided. In chapter 2 some extensions of the mean-variance asset allocation model and of the CAPM model are described. In particular, a three moment portfolio selection model (Gamba and Rossi (1998)), a four moment portfolio selection model (Jondeau and Rockinger (2003)) and the four moment CAPM are described. Chapter 3 represents the core of the dissertation. In chapter 3 a four moment portfolio selection model is proposed. The proposed model results to be a generalization of the classical mean-variance model (Markowitz (1952)) and of the three moment model (Gamba and Rossi (1998)). In chapter 4 an empirical application on real financial data is performed.

Chapter 1

Mean-Variance Asset Allocation

1.1 Utility Function

In the economic theory it is assumed that each economical agent can "value" various possible goods in terms of her own subjective preferences, without assuming the existence of utility functions. In other words, the utility function is just the instrument used in order to translate on a mathematical level the consumption preferences of an individual. Let assume that n goods are available. The nature of "goods" is irrelevant for the development of the utility theory. For example, in the following chapters the attention will be focused on financial assets: stocks, bonds, commodities. Note that, the consumption of the same physical good at different times or in different states of nature can be label to a different level of utility.

Let define the n vector x as the vector of consumption, where x_i represents selected quantity units of good i. Each consumer selects his consumption x from a particular set Ξ. Let also assume the set Ξ to be convex and closed. Preferences are described by the pre-ordering relation $\gtrsim$. The statement

$$x \gtrsim z \tag{1.1.1}$$

is read "x is weakly preferred to z". The pre-ordering relation also induces the related concepts of strict preference $>$ and indifference $\sim$. Let define the strict preference as

$$x \gtrsim z \qquad but \quad not \qquad z \gtrsim x \tag{1.1.2}$$

and the indifference relation as

$$x \gtrsim z \qquad and \qquad z \gtrsim x \tag{1.1.3}$$

In the case of the indifference relation it is also possible to say that x and z are equivalent. For a pre-ordering relation the following properties are assumed.

Axiom 1.1.1. *(**COMPLETENESS**) For every couple of vectors $x, z \in \Xi$ neither $x \gtrsim z$ or $z \gtrsim x$.*

Axiom 1.1.2. *(**TRANSITIVITY**) If $x \gtrsim z$ and $z \gtrsim y$, then $x \gtrsim y$, for $x, z, y \in \Xi$.*

Axiom 1.1.3. *(**REFLEXIVITY**) For every $x \in \Xi$, then $x \gtrsim x$.*

The meaning of the reflexivity axiom is evident and it is deeply linked to the weak preference and the pre-ordering relation. The completeness axiom appears evident too; however, when choices are made under uncertainty, many commonly used preference functions do not provide complete orderings over all possible choices. The transitivity axiom also seems intuitive, although among certain choices, it's possible to imagine comparisons that are not transitive. The three axioms are insufficient to guarantee the existence of an ordinal utility function, which describes the preferences in the pre-ordering relation. Let define an ordinal utility function as a function U from Ξ into the real numbers $\mathbf{R}$, $U : \Xi \to \mathbf{R}$. An ordinal utility function satisfies the following properties

$$U(x) > U(z) \qquad \Leftrightarrow \qquad x > z \tag{1.1.4}$$

$$U(x) = U(z) \qquad \Leftrightarrow \qquad x \sim z \tag{1.1.5}$$

In order to guarantee the existence of a utility function a fourth axiom is needed. For this purpose let introduce the continuity axiom. The advantage of introducing the continuity axiom is that it also ensures the utility function to be continuous.

Axiom 1.1.4. *(**CONTINUITY**) For every $x \in \Xi$, the two subset of all strictly preferred and all strictly worse complexes are both open.*

The openness of the two subsets of Ξ guaranties continuity of the utility function U because it requires that the utility function takes all values close to $U(x^*)$ in a neighborhood of x^*.

The four axioms are sufficient to guarantee the existence of an ordinal utility function on Ξ consistent with the preferences relation defined on the elements of Ξ. Note that once a utility function exists, it is not necessary unique. Furthermore, the derived utility function is an ordinal one and, apart from continuity guaranteed by the closure axiom, contains no more information than the pre-ordering relation defined on Ξ. No meaning can be attached to the utility level other than that inherent in the "‘greater than"’ relation in arithmetic.

In this respect, if a particular utility function $U(x)$ is a valid representation of some pre-ordering, then $\Phi(x) = \Phi[U(x)]$, where $\Phi(.)$ is any strictly increasing function. This property permits to distinguish from ordinal and cardinal utility functions. In fact, this last property is not true for cardinal utility functions. To proceed further to the development of consumer demand theory, let assume the utility function U to be twice differentiable, increasing, and strictly concave.

These assumptions guarantee that all of the first partial derivatives of U are positive everywhere, except possibly at the upper boundaries of the feasible set. Therefore, a marginal increase in income can always be profitably spent on any good to increase utility. The assumption of strict concavity guarantees that the indifference surfaces, i.e. the sets of points that have the same utility, are strictly concave upwards. That is, the set of all complexes preferred to a given complex must be strictly convex. This property can be used in order to show that a consumer's optimal choice is unique.

1.2 Expected Utility Function

Let now extend the concept of utility function to cover situations involving risk. Let assume that the agent knows the true probabilities of the events. In such a context the economical agent chooses among "lotteries" L_i described by the vector $x = (x_1, \dots, n)$ of payoffs and the corresponding vector $p = (p_1, \dots, p_n)$ of the probabilities of every single payoff.

Axioms 1.1.1, 1.1.2, 1.1.3 and 1.1.4 are still to be considered as governing choices among the various payoffs. Let assume that there is a pre-ordering relation on the set of lotteries that satisfies the following axioms:

Axiom 1.2.1. *(**COMPLETENESS**) For every couple of lotteries L_1 and L_2, neither $L_1 \gtrsim L_2$ or $L_2 \gtrsim L_1$.*

Axiom 1.2.2. *(**TRANSITIVITY**) If $L_1 \gtrsim L_2$ and $L_2 \gtrsim L_3$, then $L_1 \gtrsim L_3$, for every lottery L_1, L_2, L_3.*

Axiom 1.2.3. *(**REFLEXIVITY**) For every lottery L_i, then $L_i \gtrsim L_i$, $i = 1, \dots, n$.*

These axioms are equivalent to those used before and have the same intuition. With them it can be proved that each agent's choice is consistent with an ordinal utility function defined over lotteries or an ordinal utility functional defined over probability distributions of payoffs. The next three axioms are used in order to develop the concept of choice through the maximization of the expectation of a cardinal utility function over payoff complexes.

Axiom 1.2.4. ***(INDEPENDENCE)*** *Let* $L_1 = \{(x_1, \ldots, x_v, \ldots, x_n), p\}$ *and* $L_2 = \{(x_1, \ldots, z, \ldots, x_n), p\}$*. If* $x_v \sim z$*, then* $L_1 \sim L_2 z$ *is either a complex or another lottery. If* z *is a lottery* $z = \{(x_1^v, \ldots, x_n^v), p^v\}$ *then*

$$L_1 \sim L_2 \sim \{(x_1, \ldots, x_{v-1}, x_1^v, \ldots, x_n^v, x_{v+1}, \ldots, x_n),$$
$$(p_1, \ldots, p_{v-1}, p_v p_1^v, \ldots, p_v p_n^v, p_{v+1}, \ldots, p_n)'\}$$

In the last axiom the interpretation of the probabilities plays a central role: p_i^v is the probability of getting x_i^v conditional on outcome v after the choice of lottery L_1. $p_v p_i^v$ is the unconditional probability of getting x_i^v. This axiom is equivalent to say that only the utility of the final payoff is considered.

Axiom 1.2.5. ***(CONTINUITY)*** *If* $x_1 \gtrsim x_2 \gtrsim x_3$*, then there exists a probability* p *such that* $x_2 \sim \{(x_1, x_3), (p, 1-p)'\}$*. the probability is unique unless* $x_1 \sim x_3$*.*

Axiom 1.2.6. ***(DOMINANCE)*** *Let* $L_1 = \{(x_1, x_2), (p_1, 1-p_1)'\}$ *and* $L_2 = \{(x_1, x_2), (p_2, 1-p_2)'\}$*. If* $x_1 > x_2$*, then* $L_1 > L_2$ *if and only if* $p_1 > p_2$*.*

The last two axioms permit to define the problem of choice of the economical agent as an optimization problem with the expected utility function as objective function.

This utility function is usually called a Von Neumann- Morgenstern utility function, see Von Neumann and Morgenstern (1947). It has the properties of any ordinal utility function, but, in addition, it is a "cardinal" measure. The cardinal utility function is different from the ordinal one because the values of the function for every consumption choice have a specifical economical meaning beyond the simple ranking of lotteries. Note that in the next chapters the generic lottery will be substitute with financial activities.

Von Neumann-Morgenstern utility functions are often called measurable rather than cardinal. Each cardinal utility function corresponds to a specific ordinal utility function. Since the latter are distinct only up to a monotone transformation, two very different cardinal utility functions may have the same ordinal properties. Thus, two consumers who always make the same choice under certainty may choose among lotteries differently.

Let p_x denote the probability density for a particular lottery, and let $\phi(x)$ be the cardinal utility function of a consumer. Then the ordinal utility functional over all the possible lotteries is

$$U[p(x)] = E[\Phi(x)] = \int \Phi(x)p(x)dx$$

If lottery payoffs are drawn from a common family of probability distributions, it is possible to write the ordinal functional for lotteries as an ordinal utility function defined over parameters of the distributions of the lotteries.

For example, if there is a single good, the consumer's cardinal utility function is $\Phi(x) = x^\gamma/\gamma$, and the payoff from each lottery is log-normally distributed [i.e., $\ln(x) = N(\mu, \sigma^2)$], then

$$U[p_i(x)] = \frac{1}{\gamma} \exp\left[\gamma\mu_i + \frac{\gamma^2\sigma_i^2}{2}\right] = U(\mu_i, \sigma_i)$$

where μ_i and σ_i^2 can be interpreted as "'goods"'. Since this is an ordinal utility function, choices can be expressed equivalently by

$$\Phi(\mu, \sigma) = \frac{\ln[\gamma U(\mu, \sigma)]}{\gamma} = \mu + \frac{\gamma\sigma^2}{2} \tag{1.2.1}$$

Utility function like the one in equation 1.2.1 can be called derived utility function and it depends only on mean and variance of the probability distribution. This family of utility functions has a very important role in finance.

1.3 Independence

In the previous section an axiom about independence has been introduced. Let underline that this kind of independence is strictly related with the structure of preferences defined on the set of available goods. In other words, the independence axiom states that the utility of a lottery doesn't depend on the mechanism that determines the awards of the lottery itself.

In order to describe the preferential independence let consider a simply lottery with two outcomes x_1 and x_2 respectively with probability p_1 and p_2. Let now consider a set of $2n$ goods. Goods y, through y_n, denote quantities of goods $\mathbf{x}_1$ through $\mathbf{x}_n$ under outcome 1, and goods $\mathbf{y}_{n+1}$, through $\mathbf{y}_{2n}$ denote quantities of goods $\mathbf{x}$, through $\mathbf{x}_2$ under outcome 2. The lottery payoffs are $\mathbf{y}'_1 = (\mathbf{x}'_1, \mathbf{0}')$ and $\mathbf{y}'_2 = (\mathbf{0}', \mathbf{x}'_2)$. The expected utility of this lottery is

$$E[U(\tilde{x})] = p_1 U(\mathbf{x}_1) + p_2 U(\mathbf{x}_2) = v_1 \mathbf{y}_1 + v_2 \mathbf{y}_2 = U(\mathbf{y})$$

It is clear in the latter form that the ordinal representation displays additivity, which guarantees preferential independence. In the formulation just described, the ordinal utility functions v, depend on the lottery through its probabilities. It is possible to express choices in a general way with the use of a utility functional as outlined in the previous section. Only the additive (integral) representation (or monotone transformations) displays preferential independence. Moreover, it's possible to construct other functionals in order to satisfy the other axioms. Note that the other functionals that satisfy the axioms can't be written as expectation of the cardinal utility function.

Another kind of independence is the so called utility independence. This independence can be used in order to simplify the maximization of the expected utility function. By definition, a subset of goods is utility independent of its complement subset when the conditional preference ordering over all lotteries with fixed payoffs of the complement goods does not depend on this fixed payoff. If the subset of goods $\mathbf{y}$ is utility independent of its complement $\mathbf{z}$, it's equivalent to write the following equation

$$U(\mathbf{y},\mathbf{z}) = a(\mathbf{z}) + b(\mathbf{z})c(\mathbf{y}). \tag{1.3.1}$$

The utility function in the equation 1.3.1, $U(y, z) = \Theta[a(\mathbf{z}) + b(\mathbf{z})c(\mathbf{y})]$, displays preferential independence but not utility independence for any nonlinear function Θ. Note that ordinal utility independence can be described so any monotone transformation would be permitted while it's not the same in this case. Utility independence is not symmetric. By assuming symmetry, i.e. the goods are mutually utility independent, then it can be shown that the utility function can be represented as

$$U(\mathbf{x}) = k^{-1}(\exp[k \sum k_i U_i(x_i)] - 1)$$

with $k_i > 0$ and U_i a cardinal utility function for marginal decision involving i-th good. If the utility function has the additive form, then the preference ordering among lotteries depends only on the marginal probability distributions of the goods. The converse is also true. For multiplicative utility this simple result does not hold. This can simple be shown with a simple example, see Ingersoll (1987).

1.3.1 Expected Wealth and Risk Aversion

In finance it is more common to express outcomes in terms of wealth. For this reason the problem of the optimal choice for an economical agent can be described trough a maximization problem, where the objective function is the expected wealth. In this case, utility measures the the satisfaction associated with a particular level of wealth expressed in monetary values. With a single consumption good, this can be the numeraire for wealth, while if there are n consumption goods, then utility can be expressed as a function of wealth and the vector of consumption good prices.

When talking about economical choices the assumption of risk averse agents is common. Risk aversion is a technical definition used to describe the way the economical agent faces the risk. In other words, risk aversion can be defined analyzing the expected payoff of a lottery. A decision maker with a Von Neumann-Morgenstern utility function is said to be risk averse (at a particular wealth level) if he is unwilling to accept every actuarially fair and immediately resolved gamble with only wealth consequences, that is, those that leave consumption good prices unchanged. If the decision maker is risk averse at all wealth levels, he is said to be globally risk averse.

For state-independent utility of wealth, the utility function is risk averse at W if $U(W) > EU(W + \epsilon)$ for all gambles with $E(\epsilon) = 0$ and positive dispersion. If this relation holds at all levels of wealth, the utility function is globally risk averse. In this way it's possible to describe risk aversion using the properties of the agent's utility function.

Theorem 1.3.1. *A decision maker is (globally) risk averse if and only if his Von Neumann-Morgenstern utility function of wealth is strictly concave.*

If a utility function is twice differentiable, then it is concave, representing risk-averse choices, if and only if $U"(W) < 0$. To induce a risk-averse individual to undertake a fair gamble, a compensatory risk premium would have to be offered, making the package actuarially favorable. Similarly, to avoid a present gamble a risk-averse individual would be willing to pay an insurance risk premium. These two premia are closely related but not identical.

$$E[U(W + p_i + \epsilon)] = U(W) \tag{1.3.2}$$

$$E[U(W+\epsilon)] = U(W-p_i) \tag{1.3.3}$$

Equation 1.3.2 describes one of the more commonly risk used in economic analysis. It corresponds in an obvious way to a casualty or liability insurance premium. Equation 1.3.3 describes the compensatory risk. This risk is the extra return expected on risky activities. This second kind of risk is the most common in financial problems. In general, economical agents requires risk premia in order to play in a certain lottery. The quantity $U(W-p_i)$ is usually called certainty equivalent, i.e. the certain amount which provides the same expected utility while gambling at the lottery $W+\epsilon$.

1.3.2 Useful Utility Functions

This section provides a list of utility functions. For every utility function risk-tolerance and risk aversion properties are described.

HARA Utility Function. This utility functions class is said Hyperbolic Absolute Risk Aversion, or, equivalently, Linear Risk Tolerance (LRT). Those utility functions are commonly used and can be defined as

$$U(W) = \frac{1-\gamma}{\gamma}\left[\frac{aW}{1-\gamma}b\right]^{\gamma} \qquad b>0 \tag{1.3.4}$$

The domain of this utility function is $b + a\frac{W}{1-\gamma} > 0$. Note that the parameter γ is bounded: $\gamma < 1$ and $\gamma > 1$. For $\gamma > 1$ is defined for levels of W above the upper bound, but the marginal utility is negative. The absolute risk-tolerance is

$$R(W) = \frac{1}{A(W)} = \frac{W}{1-\gamma} + \frac{a}{b}.$$

The absolute risk-tolerance in this case is linear in W. If $\gamma < 1$ the absolute risk-tolerance is a increasing function of W while for $\gamma > 1$ is decreasing. Therefore risk aversion is decreasing for $\gamma < 1$ and increasing for $\gamma > 1$. As special cases let underline the following: linear risk neutrality corresponds to

$\gamma = 1$, quadratic for $\gamma = 2$, negative exponential utility for $\gamma \to -\infty$ and $b = 1$, iso-elastic or power utility for $b = 0$ and $\gamma < 1$, and logarithmic for $b = \gamma = 0$.

Negative Exponential Utility functions. This class of utility functions is useful due to the risk aversion properties. Let define the negative exponential utility functions as

$$U(W) = -\exp^{aW}.$$

The absolute risk aversion in this case:

$$A(W) = a. \tag{1.3.5}$$

Note that the absolute risk aversion is constant among all the values of W.

Power Utility Functions. Power utility functions are defined as:

$$U(W) = \frac{W^\gamma}{\gamma}.$$

This class of utility functions displays constant relative risk aversion

$$R(W) = 1 - \gamma$$

and, therefore, decreasing absolute risk aversion.

Logarithmic Utility Functions. The logarithmic utility function can be defined as

$$U(W) = \log(W).$$

This utility function shows constant relative risk aversion $R(W) = 1$. The logarithmic utility function can be obtained as a special case of HARA utility functions, see equation 1.3.4, for $\gamma = 0$. Note that for $\gamma = 0$ equation 1.3.4 is not defined. Let solve the limit for $\gamma \to 0$ with the use of L'Hospital's rule:

$$\lim_{\gamma \to 0} \frac{W^\gamma - 1}{\gamma} = \lim_{\gamma \to 0} \frac{W^\gamma \log(W)}{1} = \log(W)$$

Let underline that the negative utility function, the power utility function and the logarithmic utility function can be seen as special cases of the HARA family with particular values of the parameters.

1.4 Stochastic Dominance

Stochastic dominance is a pre-ordering relation on a defined set of stochastic variables. In asset allocation problems, this concept is particularly useful because it permits to rank risky activities. In finance stochastic dominance is used to order financial activities on the base of the returns distribution. From the previous sections is evident the important role of ordering a set of possible choices before defining a individual utility function coherent with the order relation. Let underline that it is always possible to define a utility function coherent with the order relation between the possible choices. Despite of this, such a utility function is not unique, i.e. two different agents with different individual utility functions can have the same preferences on the set of available choices.

Let define stochastic dominance in a rigorous way following Fishburn (1980). For any right-continuous distribution function F, with $F(x) = 0$ for all $x < 0$, let define the function F^n recursively as

$$F^1 = F \qquad \text{and} \qquad F^{n+1}(x) = \int_0^x F^n(y)dy \qquad \forall x \geq 0 \quad \text{and} \quad n = 1, 2, \ldots$$

Definition 1.4.1. ***First Order Stochastic Dominance.*** *For any given couple of distribution functions F and G, with F^1 and G^1 both finite, F dominates G in the sense of stochastic dominance of first order ($F \geq_1 G$) if and only if*

$$F^1(x) \leq G^1(x), \quad \forall x \in [0, +\infty)$$

First order stochastic dominance only takes into account the first moment of the stochastic variable distribution. Translating this concept in a financial choice framework it is equivalent to say that investors prefer higher returns to lower returns. In terms of agent's utility function first order stochastic dominance is equivalent to a monotonically increasing utility function (i.e. non-negative first derivative of the utility function). This partial ordering is not really useful in the portfolio selection framework because it doesn't take into account the risk of a certain choice.

Definition 1.4.2. ***Second Order Stochastic Dominance.*** *For any given couple of distribution functions F and G, with F^i and G^i finite for $i = 1, 2$, F dominates G in the sense of stochastic dominance of second order ($F \geq_2 G$) if and only if*

$$F^2(x) \leq G^2(x), \quad \forall x \in [0, +\infty)$$

When risk aversion is taken into account, an additional selection rule can be maintained, namely second-order stochastic dominance. This risk aversion (i.e. decreasing marginal utility) assumption is equivalent to a negative second derivative of the investors utility function, which implies concavity. In other words, investors prefer higher returns and lower variances. In a mean-variance framework, the solution of the standard Markowitz's optimization problem leads to the identification of a set of portfolios that are efficient compared to other portfolios in terms of second-order stochastic dominance.

In classical portfolio theory second order stochastic dominance plays a central role and it's linked to the assumption of normality distribution for financial returns. In fact, in this case the moments of the returns distribution are function only of the first two moments of the distribution itself. In classical mean-variance framework second order stochastic dominance is equivalent to assume the quadratic utility function for the economical agent.

Definition 1.4.3. ***Stochastic Dominance of Order n.*** *For any given couple of distribution functions F and G, with F^i and G^i finite for $i = 1, \dots, n$, F dominates G in the sense of stochastic dominance of order n ($F \geq_n G$) if and only if*

$$F^n(x) \leq G^n(x), \quad \forall x \in [0, +\infty)$$

As shown from the previous definition the concept of stochastic dominance can be easily generalized to order n. This definition will be useful in the next chapters when the portfolio selection problem will be generalized taking into account also higher moments of returns distribution. In particular in chapter 4 a solution for asset allocation problem considering skewness and kurtosis will be proposed. In this case two assumptions play a central role. First, the return distributions are considered not to be normal, i.e. skewness and kurtosis takes values respectively significantly different from 0 and 3. Second, on the other hand, investors show preferences for third and fourth moments of the returns distribution.

1.5 Introduction to Mean-Variance Analysis

Modern portfolio selection theory was born in early 1950's with seminal studies of Markowitz, see Markowitz (1952). In this framework the deterministic calculus of maximization of the agent's utility under budget constraints is not adequate. The attention is focused on choices made under uncertainty. The probabilistic notion of expected return and risk become central. The fundamental instrument for the development of such a theory of choices made under uncertainty is given by the definition of expected utility functions, see Von-Neumann Morgenstern (1947).

Before Markowitz, the widely accepted principle for portfolio selection was to select securities that maximize discounted expected return. Markowitz showed that, following this rule, the optimal portfolio is composed only by the security with the highest discounted expected return. The second fundamental contribution of Markowitz in portfolio selection analysis is to point out the role of diversification. In other words, the optimal portfolio is the one that minimize the portfolio variance under a certain level of portfolio expected return, where the diversification helps to reduce portfolio variance by investing in securities with low return covariances.

In the following sections the classical mean-variance approach is described. It will be assumed that investor's preferences are represented by an utility function that doesn't enter explicitly in the optimization problem. It is assumed that an utility function exists and that is defined over the mean and variance of the portfolio returns. A convenient approach consists in approximating the utility function using a Taylor series expansion around the current value of the portfolio return. In this context, the expected utility function is re-written as

$$E_t[U(W_{t+1})] \approx U(W_t) + U^{(1)}(W_t)E_t[W_{t+1} - W_t] + \frac{1}{2}U^{(2)}(W_t)E_t[W_{t+1} - W_t]^2 \tag{1.5.1}$$

up to some remainder in the Taylor expansion of U, where U is the utility of the investor and W represents the invested amount of wealth. The utility function has the further property of favoring higher mean and smaller variance.

1.6 Mean-Variance Portfolio Selection

In this section the following notation is used. Let

- $x = (x_1, ..., x_n)'$ be the vector of fractions of wealth allocated to the various assets
- $r = (r_1, ..., r_n)'$ be the vector of returns of n risky assets. Let also assume that not all the elements of vector r are equal.
- V be the (n, n) covariance matrix with entries $\sigma_{i,j}, \quad i, j = 1, ..., n$. The covariance matrix V is symmetric and positive definite.
- $x'Vx$ and $x'r$ be respectively the variance, denoted by σ_P^2, and the expected return, r_P, of portfolio P.

The formulation of portfolio selection problem can be stated as

$$\begin{aligned} \min_x \quad & \sigma_P^2 = x'Vx \\ s.t. \quad & x'\mathbf{1} = 1 \\ & x'r = r_P \end{aligned} \tag{1.6.1}$$

where $\mathbf{1}$ is an n-row vector of ones. In problem 1.6.1 the variance σ_P^2 of the portfolio is minimized subject to two constraints; the first constraint is also called budget constraint and it means that all wealth is invested, while the second constraint set the portfolio expected return equal to r_P. Technically, problem 1.6.1 requires the minimization of a convex function under linear constraints. The variance of the portfolio σ_P^2 is a convex function because of the positive definiteness of matrix V. Therefore, the problem has a unique solution and only first order conditions are needed. The solution for the optimization problem in 1.6.1 is

$$x^* = V^{-1}[\mathbf{r} \quad \mathbf{1}]A^{-1}[r_P \quad 1]'.$$

Let now derive the optimal portfolio in a analytical way. From the Lagrangian function

$$L = x'Vx - \lambda_1(x'\mathbf{r} - r_P) - \lambda_2(x'\mathbf{1} - 1). \tag{1.6.2}$$

The first order conditions are

$$\frac{\partial L}{\partial x} = 2Vx - \lambda_1\mathbf{r} - \lambda_2\mathbf{1} = \mathbf{0}, \tag{1.6.3}$$

where $\mathbf{0}$ is an n-vector of 0 and

$$\frac{\partial L}{\partial \lambda_1} = r_P - x'\mathbf{r} = 0, \tag{1.6.4}$$

$$\frac{\partial L}{\partial \lambda_2} = 1 - x'\mathbf{1} = 0. \tag{1.6.5}$$

From equation 1.6.3

$$x = \frac{1}{2}V^{-1}(\lambda_1\mathbf{r} + \lambda_2\mathbf{1}) = \frac{1}{2}V^{-1}[\mathbf{r} \quad \mathbf{1}][\lambda_1 \quad \lambda_2]'. \tag{1.6.6}$$

In the last equation the term $\lambda_1\mathbf{r} + \lambda_2\mathbf{1}$ is written in matrix form because it will be used 1.6.4 and 1.6.5 in order to solve for $[\lambda_1 \quad \lambda_2]'$. Then, it is possible to rewrite 1.6.4 and 1.6.5 as

$$[\mathbf{r} \quad \mathbf{1}]'x = [r_P \quad 1]'. \tag{1.6.7}$$

Pre-multiply both sides of 1.6.6 by $[\mathbf{r} \quad \mathbf{1}]'$ and use 1.6.7 in order to obtain

$$[\mathbf{r} \quad \mathbf{1}]'x = \frac{1}{2}[\mathbf{r} \quad \mathbf{1}]'V^{-1}[\mathbf{r} \quad \mathbf{1}][\lambda_1 \quad \lambda_2]' = [r_P \quad 1]'. \tag{1.6.8}$$

For notational convenience let denote by

$$A = [\mathbf{r} \quad \mathbf{1}]'V^{-1}[\mathbf{r} \quad \mathbf{1}] = \begin{bmatrix} a & c \\ c & d \end{bmatrix} \tag{1.6.9}$$

the above 2×2 symmetric matrix. The entries of the matrix A are equal to: $a = r'V^{-1}r$, $c = r'V^{-1}\mathbf{1}$ and $d = \mathbf{1}'V^{-1}\mathbf{1}$. The matrix A is a definite positive matrix by the properties of the matrix V^{-1}. In fact, for any y_1, y_2 such that at least one of the elements y_1, y_2 is non-zero,

$$[y_1 \quad y_2]A[y_1 \quad y_2]' = [y_1 \quad y_2][\mathbf{r} \quad \mathbf{1}]'V^{-1}[\mathbf{r} \quad \mathbf{1}][y_1 \quad y_2]' > 0. \tag{1.6.10}$$

By substituting A in equation 1.6.8 the following equation is obtained

$$\frac{1}{2}A[\lambda_1 \quad \lambda_2]' = [r_P \quad 1]' \tag{1.6.11}$$

from which it's possible to solve for the multipliers since A is non-singular. Thus

$$\frac{1}{2}[\lambda_1 \quad \lambda_2]' = A^{-1}[r_P \quad 1]'. \tag{1.6.12}$$

From these manipulations, the n-vector of portfolio weights x^* that minimize portfolio variance for a given mean return is

$$x^* = \frac{1}{2}V^{-1}[\mathbf{r} \quad \mathbf{1}][\lambda_1 \quad \lambda_2]' = V^{-1}[\mathbf{r} \quad \mathbf{1}]A^{-1}[r_P \quad 1]'. \tag{1.6.13}$$

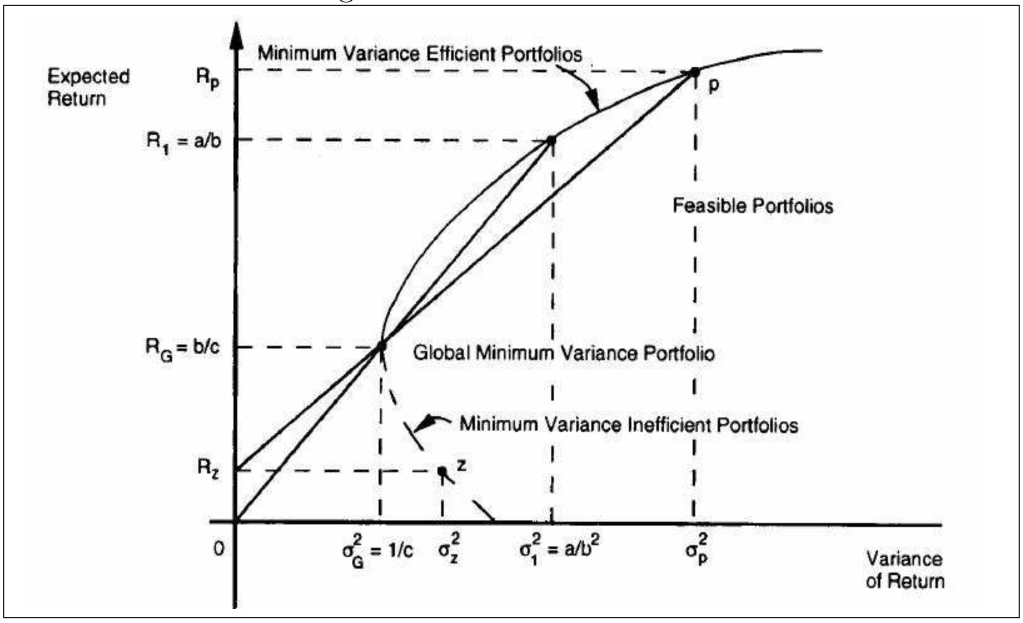
Let compute the variance of any minimum variance portfolio with a given expected return equal to r_P. Using the definitions of the variance of the portfolio σ_P^2, the matrix A and the optimal weights x^*, the set of optimal portfolios can be written as

$$\begin{aligned}\sigma_p^2 = x^{*'}Vx^* &= [r_P \quad 1]A^{-1}[\mathbf{r} \quad \mathbf{1}]'V^{-1}VV^{-1}[\mathbf{r} \quad \mathbf{1}]A^{-1}[r_P \quad 1]' \qquad (1.6.14)\\ = [r_P \quad 1]A^{-1}[r_P \quad 1]' &= [r_P \quad 1]\frac{1}{ad-c^2}\begin{bmatrix} d & -c \\ -c & a \end{bmatrix}[r_P \quad 1]' \\ &= \frac{a - 2cr_P + dr_P^2}{ad-c^2}\end{aligned}$$

In equation 1.6.14 the relation between the variance of the minimum variance portfolio σ_P^2 for any given mean r_P is expressed as a parabola.

Figure 1.1 graphs equation 1.6.14 and distinguishes between the upper half (solid curve) and the bottom half (dashed curve). The upper half of the

Figure 1.1: **Efficient Frontier**



minimum variance portfolio frontier identifies the set of portfolios having the highest return for a given variance; these are called mean-variance efficient portfolios. The portfolios on the bottom half are inefficient. Portfolios to the right of the parabola are called feasible and they are still inefficient in the sense of second order stochastic dominance. For a given variance, the mean return of a feasible portfolio is less than the mean return of an efficient portfolio and higher than the mean return of an inefficient one.

Figure 1.1 also identifies the global minimum variance portfolio. This is the portfolio with the smallest variance for any expected return. Its mean r_G and variance σ_G^2 are respectively given by

$$r_G = \frac{c}{d},$$

$$\sigma_G^2 = \frac{a - 2cr_G + dr_G^2}{ad - c^2} = \frac{a - 2c(\frac{c}{d}) + d(\frac{c}{d})^2}{ad - c^2} = \frac{1}{d}.$$

The global minimum variance portfolio is on the efficient frontier and the correspondent vector of weights x_G can be evaluated as follows

$$x_G = V^{-1}[\mathbf{r} \quad \mathbf{1}]A^{-1}\begin{bmatrix} r_G \\ 1 \end{bmatrix} = \frac{V^{-1}[\mathbf{r} \quad \mathbf{1}]\begin{bmatrix} d & -c \\ -c & a \end{bmatrix}\begin{bmatrix} \frac{c}{d} \\ 1 \end{bmatrix}}{ad - c^2} = \frac{V^{-1}\mathbf{1}}{d}.$$

An additional notion also illustrated in Figure 1.1 is that of orthogonal portfolio. Two minimum variance portfolios x_P and x_Z are said to be orthogonal if their covariance is zero, that is

$$x_Z' V x_P = 0.$$

It's possible to show that for any minimum variance portfolio, except the global minimum variance portfolio, the orthogonal portfolio is unique. Furthermore, if the first portfolio has expected return r_P, its orthogonal one has expected return r_Z equal to

$$r_Z = \frac{a - cr_P}{c - dr_P}$$

Figure 1.1 also shows the geometry of orthogonal portfolios. Given an arbitrary efficient portfolio P on the efficient frontier, the line between P and the global minimum variance portfolio can be shown to intersect the expected return axis at r_Z.

1.7 Mean-Variance Portfolio With a Risk-less Asset

In order to build the efficient mean-variance portfolios in presence of n risky assets and one risk-less asset, let first consider the two fund separation property. The economic implications of this property are significant because the following theorem establishes that the minimum variance portfolio frontier can be generated by any two distinct frontier portfolios.

Theorem 1.7.1 (Two-fund separation). *Let x_a and x_b be two minimum variance portfolios with expected return r_a and r_b respectively, such that $r_a \neq r_b$.*

- *Then every minimum variance portfolio x_c is a linear combination of x_a and x_b.*
- *Conversely, every portfolio which is a linear combination of x_a and x_b, i.e. $\beta x_a + (1-\beta)x_b$, $\beta \in \boldsymbol{R}$, is a minimum variance portfolio.*
- *In particular, if x_a and x_b are minimum variance efficient portfolios, then $\beta x_a + (1-\beta)x_b$ is a minimum variance efficient portfolio for $0 \leq \beta \leq 1$.*

It is of historical interest that this fact was discovered by Tobin (1958). Tobin uses only two assets (risk-less cash and a risky consol), and demonstrates that nothing essential is changed if there are many risky assets. He argues that the risky assets can be viewed as a single composite asset (mutual fund) and investors find it optimal to combine their cash with a specific portfolio of risky assets. In particular, the two fund separation property shows that any mean variance efficient portfolio can be generated by two arbitrary distinct mean-variance efficient portfolios. In other words, an investor who wants a mean-variance efficient portfolio can achieve this goal by investing in an appropriate linear combination of any two mutual funds that are also mean-variance efficient.

Let now return to Tobin's original idea of introducing a risk-less asset. The portfolio selection problem with n risky assets and a risk-less one can easily be formulated and solved. Let there be $n+1$ assets, $i = 0, ..., n$, where 0 denotes the risk-less asset with return r_0. The vector of expected excess returns has elements defined as $r_i = r_i - r_0, i = 1, ..., n$, and is denoted by $\mathbf{r}$. Wealth is now allocated among $(n+1)$ assets with weights $x_0, x_1, ..., x_n$. In the various calculations the vector of weights $x_1, ..., x_n$ is denoted by $\mathbf{x}$ and $x_0 = 1 - x'\mathbf{1}$.

For a given portfolio P, the expected excess return is

$$r_P = x'r + (1 - x'\mathbf{1})r_0 - r_0 = x'\mathbf{r}. \tag{1.7.1}$$

The variance of p is

$$\sigma_P^2 = \mathbf{x}'V\mathbf{x}, \tag{1.7.2}$$

where in equations 1.7.1 and 1.7.2, $\mathbf{r}$ and V are defined as in the previous section. Obviously, the risk-less asset doesn't contribute to the variance of the portfolio.

The mean-variance optimization problem with a risk-less asset can be stated as

$$\begin{aligned} \min_{\mathbf{x}} \quad & \sigma_P^2 = \mathbf{x}'V\mathbf{x} \\ s.t. \quad & \mathbf{x}'\mathbf{r} = r_P. \end{aligned} \tag{1.7.3}$$

In problem 1.7.3, the variance of the $n-$risky assets is minimized subject to a given expected excess return r_P. Note that $\mathbf{x}'\mathbf{1} = 1$ is not a constraint

because the whole wealth need not all to be allocated into the $n-$risky assets but some can be held in the risk-less asset.
Writing the Lagrangian function $L(x, \lambda_1)$ for problem 1.7.3

$$L = x'Vx - \lambda_1(x'\mathbf{r} - r_p)$$

where λ_1 is the Lagrange multiplier, the first order conditions are

$$\frac{\partial L}{\partial x} = \frac{1}{2}Vx - \lambda_1 \mathbf{r} = \mathbf{0} \tag{1.7.4}$$

$$\frac{\partial L}{\partial \lambda_1} = r_p - x'\mathbf{r} = 0. \tag{1.7.5}$$

As in the previous case, V being positive definite, the first order conditions are sufficient for a minimum for problem 1.7.3. By solving equations 1.7.4 and 1.7.5, one obtains the solution

$$\mathbf{x}^* = \left(\frac{r_P}{\mathbf{r}'V^{-1}\mathbf{r}}\right) V^{-1}\mathbf{r} \tag{1.7.6}$$

which gives the variance of the minimum variance portfolio with a given excess expected return

$$\begin{aligned}\sigma_P^2 = \mathbf{x}'V\mathbf{x} &= \left(\frac{r_P}{\mathbf{r}'V^{-1}\mathbf{r}}\right)^2 \mathbf{r}'V^{-1}VV^{-1}\mathbf{r} \qquad (1.7.7)\\ &= \frac{r_P^2}{\mathbf{r}'V^{-1}\mathbf{r}}\end{aligned}$$

The tangency portfolio T, as shown in figure 1.2, is the minimum variance portfolio for which

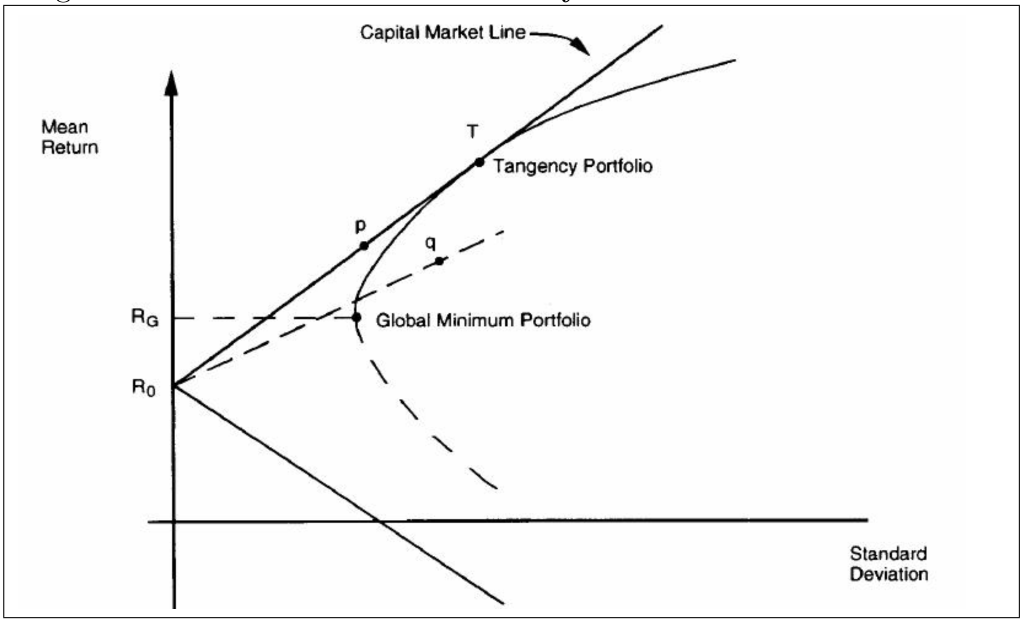
$$\mathbf{1}'\mathbf{x}_T = 1. \tag{1.7.8}$$

Combining equations 1.7.6 and 1.7.8 it can be obtained

$$r_T = \frac{\mathbf{r}'V^{-1}\mathbf{r}}{\mathbf{1}'V^{-1}\mathbf{r}} >< 0. \tag{1.7.9}$$

It is economically plausible that risk-less return is lower than the expected return of the minimum global variance portfolio of the risky assets, that is,

Figure 1.2: **Efficient Frontier: n risky assets and 1 risk-less asset.**



$r_0 < r_G$. It is possible to prove that $\mathbf{1}'V^{-1}\mathbf{r} > 0$. Also $\mathbf{r}'V^{-1}\mathbf{r} > 0$ by the positive definiteness of matrix V. It then follows that r_T and the slope of the tangency line in figure 1.2 is positive. This positive slope line is called capital market line and defines the set of minimum variance efficient portfolios.

1.8 The Capital Asset Pricing Model

The CAPM model was first introduced by Sharpe (1964), Lintner (1965) an Mossin (1966). The first step is to consider the universe of investors as a whole, which leads to the deduction of the Capital Asset Pricing Model (CAPM). The idea is to derive a theory of asset valuation in an equilibrium situation, drawing together expected returns and market risk. This model is considered as the first to explain asset valuation by using the notion of risk. This risk can be split into a systematic risk, common to all assets in the same market segment and an unsystematic risk attributed to the specific asset. The CAPM is built on the following main assumptions:

- Investors are risk averse and maximize the expected utility of wealth at the end of the period.
- Investors are considered to be homogeneous, i.e. agents have the same preferences and they are considered to be rational.
- The asset returns are normally distributed or the investor only considers the first two moments of their return distribution, that is equivalent to assume a quadratic utility function for the investors.
- Investors only consider one investment period which is the same for all investors.
- Investors have limitless access to financial markets and can borrow and lend at a risk-free rate r_F.
- Markets are complete (perfect information) and perfect (no taxes, no transaction costs).

In order to consider market equilibrium all the investors have to be taken into account. The introduction of a risk-free asset, whose return is denoted r_F , enables the investor to spread her wealth between an efficient portfolio and this risk free asset, which leads to the following equation, where r_E denotes the return of the chosen efficient portfolio and r_P the return of the resulting portfolio composed of the risky and the risk-less asset:

$$E(r_P) = xr_F + (1 - x)E(r_E) \tag{1.8.1}$$

with x the proportion of wealth invested in the risk-free asset. Equation 1.8.1 is obtained by combining the classical mean variance approach and the two funds separation theorem, see the previous section. The portfolio's risk is than simply given by:

$$\sigma_P = (1 - x)\sigma_E \tag{1.8.2}$$

Combining the equations 1.8.1 and 1.8.2 let write the following expression:

$$E(r_P) = r_F + \frac{(E(r_E) - r_F)}{\sigma_E}\sigma_P$$

For each efficient portfolio a line represents all the combinations between the efficient portfolio and the risk-free asset. Among this set of lines there is one that dominates all others. This line corresponds to the point (portfolio) T, the tangency portfolio. Thus, the efficient frontier in a world with a risk-free asset is a straight line from the point r_F to T. This line is also called capital market line, as already mentioned in the previous sections.

Investors benefit from the risk-free opportunity, since for a given expected return the risk on this new efficient frontier is less than or equal to the risk of the corresponding portfolio consisting of risky assets. In other words, the efficient frontier with a risk-free asset dominates the efficient frontier without the risk-free asset. As in the Markowitz Model the choice of a specific point on this line depends on the individual utility function or, more precisely, on the level of risk-aversion of the investor. If the investor has unlimited access to an efficient financial market, i.e. can borrow money to a rate r_F , the efficient frontier to the right of point T corresponds to the extension of the line between r_F and T. If the investor is constrained to borrow to a rate $(\overline{r}_F) > r_F$ the efficient frontier is flatter to the right of point T.

If the investor has no access at all to financial markets the efficient frontier with a risk-free asset corresponds to the efficient frontier without a risk-free asset for $E(r_P) > E(r_M)$. These results do not only depend on the respective financial market and its accessibility but also on the assumption that all assets are infinitely divisible.

The following derivation of the CAPM assumes an efficient financial market $(\overline{r}_F) = r_F$ and infinite divisibility of assets. This kind of market is called frictionless. The portfolio decision problem can be divided into two parts: first the investor chooses the optimal risky portfolio corresponding to one point on the efficient frontier without the risk-free asset and second, the investor chooses how split his portfolio between the risk-free asset and the risky portfolio. This is known as Tobin's (1958) two-fund separation theorem, as already mentioned.

Taking all investors into account, let move on to the market equilibrium framework. Since every investor is supposed to hold a mean-variance portfolio and assuming homogenous expectations of moments, all agents have to solve the same optimization problem and derive the same efficient frontier, since the one passing through T dominates all the others. Depending on their level of risk aversion they invest a certain proportion in the risk-free asset and the rest in the portfolio T. In equilibrium, all assets are held and since the only traded risky portfolio is portfolio T it must contain all assets. Hence, this portfolio is the market portfolio and it holds all the assets in proportions of their market capitalization.

The Capital Asset Pricing Model aims to value each asset by considering an equilibrium situation. By applying the two-fund theorem only the risky part of the portfolio has to be taken into account in order to price each asset, since the two decisions are independent. Similar to equations 1.8.1 and 1.8.2 let define

$$E(r_P) = xr_i + (1-x)E(r_T)$$

$$\sigma_P = [x^2\sigma_i + (1-x)^2\sigma_T^2 + 2x(1-x)\sigma_{iT}]^{\frac{1}{2}}$$

By varying x, all possible efficient portfolios consisting of the risky asset and the market portfolio in the $E(r_P) - \sigma_P -$ space can be described. In the optimum the leading coefficient of the tangent to this efficient frontier must be equal to the slope coefficient of the capital market line:

$$\frac{\partial E(r_P)}{\partial \sigma_P} = \frac{E(r_T) - r_F}{\sigma_T} \tag{1.8.3}$$

Considering the functional relations above let write:

$$\frac{\partial E(r_P)}{\partial \sigma_P} = \frac{\partial E(r_P)}{\partial x}\frac{\partial x}{\partial \sigma_P}$$

The two derivatives are:

$$\frac{\partial E(r_P)}{\partial x} = E(r_i) - E(r_T)$$

$$\frac{\partial \sigma_P}{\partial x} = \frac{2x\sigma_i - 2(1-x)\sigma_T^2 + 2(1-2x)\sigma_{iT}}{2\sigma_P}$$

which gives:

$$\frac{\partial E(r_P)}{\partial \sigma_P} = \frac{[E(r_i) - E(r_T)]\sigma_P}{x(\sigma_i^2 + \sigma_T^2 - 2\sigma_{iT}) + \sigma_{iT} - \sigma_T^2}$$

In equilibrium the market portfolio contains all assets. The proportion x is therefore an excess in asset i in the portfolio P that must be zero at equilibrium (considering all investors). The Portfolio P is then the market portfolio ($\sigma_P = \sigma_T$) and for the point T:

$$\frac{\partial E(r_P)}{\partial \sigma_P}(T) = \frac{[E(r_i) - E(r_T)]\sigma_T}{\sigma_{iT} - \sigma_T^2}$$

From equation 1.8.3, the following equation holds:

$$\frac{[E(r_i) - E(r_T)]\sigma_T}{\sigma_{iT} - \sigma_T^2} = \frac{E(r_T) - r_f}{\sigma_T}$$

That last expression can also be written as

$$E(r_i) = r_f + \frac{E(r_T) - r_f}{\sigma_T^2}\sigma_{iT} \tag{1.8.4}$$

Defining

$$\beta_i = \frac{\sigma_{iT}}{\sigma_T^2}$$

the characteristic relationship of the CAPM:

$$E(r_i) = r_f + \beta_i(E(r_M) - r_f) \tag{1.8.5}$$

This equation can be understood as follows: the expected return of the risky asset i equals the return of the risk-free asset plus a risk premium. β_i is also called the systematic risk of asset i. By consequence, the risk-free asset has a beta of zero and the market portfolio a beta of one. The CAPM establishes a theory for valuing individual securities and highlights the importance of taking risk into account. It states that there are two kinds of risk. First, the systematic risk, common to all assets, which is rewarded by the market (risk premium). Second, each asset has an individual non-rewarded risk which can also be called diversifiable risk, since it can be avoided by constructing well diversified portfolios. Let rewrite equation equation 1.8.5:

$$r_i = r_f + \beta_i(E(r_M) - r_f) + \epsilon_i \tag{1.8.6}$$

with $E(\epsilon_i) = 0$ and $Var(\epsilon)$ the specific risk.

Chapter 2

The Reasons to Consider Higher Moments in Asset-Allocation

To solve asset-allocation problems, the well-known mean-variance criterion proposed by Markowitz (1952) has provided very sensible results for a very wide range of situations. While some authors have argued that the expected utility function may be more appropriately approximated by a function of higher moments, for example see Arditti (1967), and Samuelson (1970), early empirical evidence suggests that mean-variance criterion results in allocations that are very similar to the ones obtained using a direct optimization of the expected utility function, see Levy and Markowitz (1979), Pulley (1981), and Kroll, Levy and Markowitz (1984).
An explanation of the good performance of the mean-variance criterion may be that, although returns are non-normal, they are driven by an elliptical distribution, for which the mean-variance approximation of the expected utility remains good for all utility functions, see Chamberlain (1983). In contrast, under large departure from normality, in particular when the distribution is severely asymmetric, Chunachinda et al. (1997), Athayde and Flôres (2004)

and Jondeau and Rockinger (2003) show that the mean-variance criterion may fail to correctly approximate the expected utility. In such a case, a three- or four-moment optimization strategy should provide a better approximation of the expected utility.

A limitation of this previous evidence is that it assumes constant investment opportunities while a huge literature has argued that asset returns have time varying moments. However, extension to a conditional asset-allocation is much more difficult to implement.

Modelling asset returns requires rather general distributions that are able to incorporate volatility clustering, asymmetry, and fat-tails features found in empirical data.

2.1 About Mean-Variance Hypothesis

The presence of skewness and kurtosis in asset return distributions apart from normality is well known. Here the research is focused instead on the analysis of co-skewness and co-kurtosis and, if any, their relevance in modelling asset pricing and asset allocation.

On one hand, peculiar return distribution patterns may be originated from the use of specific trading strategies. Hedge fund managers pursue varied hedging and arbitrage strategies that engender pay-off profiles extremely different from traditional assets. The number of new assets traded on global market is increasing continuously and these new financial products may have different returns distribution compared to classical products (stocks and bonds).

On the other hand, skewed and / or kurtotic return distributions may be seen as the statistical expression of market inefficiency and market frictions.

Specifically, non-normal return distributions may be due to illiquidity, lack of divisibility, and low information transparency. All these factors contrast with the assumptions underpinning the standard CAPM model and the standard Mean-Variance model.

Active investing strategies. It is worth emphasizing that trading strategies applied by hedge fund managers engender return distribution typically different from equity market or mutual fund returns. Hedge funds are often able to protect investors against declining markets. Hedge fund managers pursue downside protection by utilizing a variety of hedging strategy and investment styles. As a result, some hedge funds generate non-negative returns even in declining markets. This automatically leads to a positive asymmetry of returns distribution.

The use of leverage and derivatives contributes to the realization of particular risk-performance profiles characterized by low correlation with traditional asset markets. The hedge fund trading strategies widely benefit from options, option-like trading strategies and, in general, financial engineering.

Hedge funds are less regulated than mutual funds. The weaker restrictions allow short selling to boost performance.

Sample Data frequency. A lot of empirical studies show how data frequency can affect the distribution of financial returns. In particular, higher frequency data (like daily or hourly) show non-elliptic distributions of returns. This fact can be relevant in active investing strategies when high frequency data are used to frequently update the asset allocation.

Illiquidity. Illiquid assets do not allow trading of any volume size with an immediate execution and / or without price impact. In particular, hedge funds are generally considered illiquid assets. In fact, the investing strategies of hedge funds are frequently based on highly illiquid and / or volatile assets. Illiquidity is in contrast with one of the main assumptions of the standard CAPM model and standard asset-allocation models. A low level of liquidity in hedge funds would require a risk premium, and hence an asymmetry in returns distribution.

Lack of divisibility. Another assumption behind the CAPM and standard asset-allocation models is that assets are infinitely divisible. This means that investors could take any position in an investment, regardless of the investment size. On the contrary, a minimum investment in a hedge fund is always required. This is a high barrier to entry. High entry barriers may represent a considerable opportunity cost to exit or to undertake short-run trading strategies. Another example of entry and exit barriers in hedge funds is represented by the number of entry and exit dates.

Information transparency. Market inefficiency may be also due to opaque or asymmetric information. It is well known that hedge funds do not easily disclose information about their current positions. The low degree of information transparency is partially justified by the short positions and arbitrage strategies undertaken by hedge fund managers. This kind of trading strategies implies disguising trading positions especially in illiquid markets. In fact, full and transparent information disclosure would jeopardize trading opportunities.

2.2 Three Moment Asset Allocation Problem

Some asset allocation models has been proposed in the literature in order to generalize mean-variance approach and taking into accounts also higher moments. From this point of view, let follow Gamba and Rossi (1998). This approach can be classified as a classical approach to asset allocation due to the hypothesis made on the assets returns. These assumptions permit to give a closed form solution to the optimization problem. Let assume that n risky assets and 1 risk-less asset are available. Let

- x be the column vector of portfolio weights.
- r_i be the return of i-th asset, with μ_i, σ_i and ξ_i respectively the expected value, the standard deviation and the skewness of asset i.
- μ_0 be the return of the risk-less asset.
- y be a random variable with $E(y) = 0$, $E(y^2) = \sigma_y^2$, $E(y^3) = \xi_y^3 \neq 0$.
- ϵ_i be the i-th entry of a random vector of joint Gaussian variables conditional to y.
- b_i be a real number.

Let assume that i-th asset return is:

$$r_i = \mu_i + \epsilon_i + b_i y.$$

According with the assumptions, the covariance matrix of returns is

$$D = [E(\epsilon_i \epsilon_j | y) + b_i b_j \sigma_y^2] = \sigma_{ij}, \qquad i, j = 1, \ldots, n \tag{2.2.1}$$

while the co-skewness is

$$E[(r_i - \mu_i)(r_j - \mu_j)(r_k - \mu_k)] = \xi_{ijk} = b_i b_j b_k \xi_y^3. \tag{2.2.2}$$

A stochastic dominance rule on mean-variance and skewness is proposed. Let r_a and r_b be two random returns with moments respectively (μ_a, σ_a, ξ_a) and (μ_b, σ_b, ξ_b).

Definition 2.2.1. *Return r_a E-V-S (expected return-variance-skewness) dominates r_b if $\mu_a \geq \mu_b$, $\sigma_b \geq \sigma_a$, $\xi_a \geq \xi_b$, with at least a strict inequality.*

The structure of preferences is defined such that an investor prefers the higher values for odd moments (expected return and skewness) and lower levels for even (variance in this case). This structure of preferences, i.e. the expected return-variance-skewness is in general not compatible with the Expected Utility theory. Although no such general compatibility exists, but, by suitably restricting the set of possible distributions and the set of Von Neumann-Morgenstern utility functions the Expected return-variance-skewness criterion can be made compatible with Expected Utility theory.

2.2.1 No Risk-less Asset

In the case of n risky assets, the portfolio optimization problem can be written as follows:

$$\begin{aligned} \max_{x} \quad & x'b\xi_y \\ s.t. \quad & x'Dx = \sigma^2_{P_2} \\ & x'\mu = \mu_P \\ & x'\mathbf{1} = 1 \end{aligned} \tag{2.2.3}$$

where μ_P and $\sigma^2_{P_2}$ are respectively the expected return and the variance of portfolio P. The portfolio optimization problem 2.2.3 can be solved analytically due to the assumptions made on the assets returns. Let define the matrix P_2 and the matrix A as

$$P_2 = M'D^{-1}M = \begin{bmatrix} a & c & f \\ c & d & g \\ f & g & e \end{bmatrix} \qquad A = \begin{bmatrix} a & c \\ c & d \end{bmatrix}$$

where the $(n, 3)$ matrix M is defined as $M = (\mu, \mathbf{1}, b)$. The matrix M is assumed to be of full column rank, i.e. the vectors μ, $\mathbf{1}$ and b are assumed to be linearly independent. The matrix P_2 is symmetric by construction. The entries of matrix P_2 are: $a = \mu'D^{-1}\mu$, $c = \mu'D^{-1}\mathbf{1}$, $f = \mu'D^{-1}b$, $d = \mathbf{1}'D^{-1}\mathbf{1}$, $g = \mathbf{1}'D^{-1}b$, $e = b'D^{-1}b$.

Let also define the quantity σ_A^2 as

$$\sigma_A^2 = (\mu_P \quad \mathbf{1})A^{-1}\begin{pmatrix} \mu_P \\ \mathbf{1} \end{pmatrix} > 0. \tag{2.2.4}$$

Note that σ_A^2 is the set of optimal portfolios represented in mean-variance plane, see Huang and Litzenberger (1988).

Given $\mu_P \in \mathbf{R}$, $\sigma_{P_2} \geq \sigma_A$ and $\xi_y > 0$ the optimal solution x^* for problem 2.2.3 is:

$$x^* = \left[D^{-1}b - D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix} f \\ g \end{pmatrix} \right] \sqrt{\frac{\sigma_{P_2}^2 - \sigma_A^2}{e-h}} + D^{-1}(\mu, \mathbf{1})A^{-1}\begin{pmatrix} \mu_P \\ \mathbf{1} \end{pmatrix} \tag{2.2.5}$$

where $h = (f \quad g)A^{-1}(f \quad g)'$.For more details on the derivation of the optimal portfolio x^* see Gamba and Rossi (1998).

The optimal portfolio in equation 2.2.5 can be written as the sum of two portfolios x_1^* and x_2^*

$$x_1^*(\mu_P) = D^{-1}(\mu, \mathbf{1})A^{-1}\begin{pmatrix} \mu_P \\ \mathbf{1} \end{pmatrix}$$

$$x_2^*(\mu_P, \sigma_P) = \left[D^{-1}b - D^{-1}(\mu, \mathbf{1})A^{-1} \begin{pmatrix} f \\ g \end{pmatrix} \right] \sqrt{\frac{\sigma_{P_2}^2 - \sigma_A^2}{e - h}}.$$

Portfolio x_1^* is the variance-minimizing one for a given μ_P, in the classical mean-variance framework. Portfolio x_2^* represents an arbitrage portfolio, i.e. $\mu' x_2^* = 0$. It can be thought of as a fair bet made by the agent to exploit the distributional asymmetry of returns. The arbitrage portfolio increases volatility of portfolio x_1^* from σ_A^2 up to $\sigma_{P_2}^2$. This higher variance is counter - balanced by a higher skewness.

2.2.2 With a Risk-less Asset

In the case of n risky assets and a risk-less one, the optimization problem can be written as:

$$\begin{aligned} \max_x \quad & x' b \xi_y \\ s.t. \quad & x' D x = \sigma_{Q_2}^2 \\ x'(\mu - \mu_0 \mathbf{1}) &= \mu_P - \mu_0 \end{aligned} \tag{2.2.6}$$

where μ_0 is the return of the risk-less asset. Also in the case of $n+1$ assets, the portfolio optimization problem 2.2.6 can be solved analytically due to the assumptions made on the assets returns structure.

Let define the matrix Q_2 as

$$Q_2 = N' D^{-1} N = \begin{bmatrix} m & l \\ l & e \end{bmatrix}$$

where the $(n, 2)$ matrix N is defined as $N = (\mu - \mu_0 \mathbf{1} \quad b)$. The matrix N is assumed to be of full column rank, i.e. the vectors $\mu - \mu_0 \mathbf{1}$ and b are assumed to be linearly independent. The matrix Q_2 is symmetric by construction. The entries of matrix Q_2 are: $m = (\mu - \mu_0 \mathbf{1})' D^{-1} (\mu - \mu_0 \mathbf{1})$, $l = (\mu - \mu_0 \mathbf{1})' D^{-1} \mathbf{1}$, $e = b' D^{-1} b$.

Let define the quantity σ_m^2 as

$$\sigma_m^2 = \frac{\mu - \mu_0}{\sqrt{m}}$$

Given $\mu_P \in \mathbf{R}$, $\sigma_{Q_2}^2 \geq \sigma_m^2$ and $\xi_y > 0$ the optimal solution for problem 2.2.6 is $\overline{x} = (1 - x^{*'}\mathbf{1}, x^{*'})$ with

$$x^* = \left[D^{-1}b - \frac{l}{m}D^{-1}(\mu - \mu_0\mathbf{1})\right]\sqrt{\frac{\sigma_{Q_2}^2 - \sigma_m^2}{e-k}} + \frac{\mu_P - \mu_0}{m}D^{-1}(\mu - \mu_0\mathbf{1}) \tag{2.2.7}$$

where $k = \frac{l^2}{m}$. For more details on the derivation of the optimal portfolio see Gamba and Rossi (1998).

The portfolio x^* in equation 2.2.7 can be split into two different portfolios x_1^* and x_2^*

$$x_1^* = \frac{\mu_P - \mu_0}{m}D^{-1}(\mu - \mu_0\mathbf{1})$$

$$x_2^* = \left[D^{-1}b - \frac{l}{m}D^{-1}(\mu - \mu_0\mathbf{1})\right]\sqrt{\frac{\sigma_{Q_2}^2 - \sigma_m^2}{e-k}}$$

The optimal portfolio is the sum of two portfolios. Portfolio $\overline{x}_1 = (1 - x_1^{*'}\mathbf{1}, x_1^{*'})$, which is the usual variance minimizing portfolio given μ_P in the mean-variance framework. Portfolio $\overline{x}_2^*$ has the same features and play the same role of portfolio x_2^* for the n risky asset optimization problem. For more details on the calculations of the optimal portfolio, see Gamba Rossi (1998).

2.3 The Four Moment CAPM

Let start with the simple idea that the investor has a specific utility function and is willing to optimize her expected utility of wealth. In the mean-variance approach of Markowitz, upon which the CAPM is built, the risk was represented by the variance (or the standard deviation) of the portfolio returns. Consequently, the investor tried to maximize the expected portfolio return given a certain standard deviation or tried to minimize the standard deviation given a level of expected return of her portfolio. Depending on her level of risk aversion the investor chooses a point on the efficient frontier representing an expected utility-standard-deviation couple. Considering only the first two moments of the portfolio return distribution is only an approximation of the real portfolio allocation game, except in two situations:

- When the portfolio returns are normally distributed.
- When the representative agent (investor) has an utility function only depending on the first two moments, i.e. quadratic utility function.

However, empirical results have proved that the corresponding third and fourth moment (skewness and kurtosis) significantly differ from those of the normal distribution, see Kraus and Litzenberger (1976), Campbell and Siddique (1999), Hwang and Satchell (1999), Fang and Lai (1997). Hence, an equilibrium analysis such as the CAPM, that is built on the expected utility-risk duality, should take these higher moments into account. In addition, this approach is justified by the fact that the most used utility functions yield existing derivatives of higher order different from zero. Instead of fixing one particular utility function, let concentrate on a general method that is applicable to a large class of functions with the objective to show the need of considering higher moments.

Let consider any arbitrary utility function. As in Markowitz (1952), the investor only tries to maximize her wealth stemming from her asset investment, assuming a world without labor income. Hence, she will consider the utility of the investment return (r). The fourth-order Taylor expansion is:

$$U(r) = \sum_{i=0}^{4} \frac{U^{(i)}E(r)}{i!}(r - E(r))^i + o[(r - E(r))^4]$$

where $U^{(n)}$ denotes the n-th derivative of the function U. Taking the expectation on both sides yields:

$$E[U(r)] = U[E(r)] + \frac{U^{(2)}[E(r)]}{2}\sigma^2 + \frac{U^{(3)}[E(r)]}{6}s + \frac{U^{(4)}[E(r)]}{24}k \qquad (2.3.1)$$

with s the non-standardized skewness and k the non-standardized kurtosis of the portfolio return distribution. Note that the usual definitions of skewness and kurtosis coefficients are normalized:

$$S = \frac{E[r - E(r)]^3}{\sigma^3}$$

$$K = \frac{E[r - E(r)]^4}{\sigma^4}$$

By neglecting the influence of the third and fourth moments, equation 2.3.1 coincides with the mean-variance approach, since the second derivative is negative. Maximizing expected utility is equivalent to the portfolio trade-off between mean and variance and depends on the level of risk aversion. Similarly, if the utility function only depends on the first two moments (i.e. $U = U(\mu, \sigma^2)$), the third and fourth derivatives are zero and the last two terms are consequently equal to zero. This is the underlying quadratic utility function in the Markowitz approach.

Analyzing equation 2.3.1, let assume that investors have preference for a higher skewness and an aversion towards kurtosis which had already been found by Horvath and Scott (1980). Looking at the distribution this is rather easy to understand. A positive skewness means a higher probability for higher values of wealth relative to lower values.

Concerning the fourth moment, a high kurtosis reflects the so called "fat tails", i.e. higher probability for extreme values than in the case of a normal distribution. Following expected utility theory, the negative value that is attributed to the chance of highly negative returns excesses the positive value that is attributed to the opportunity of higher returns. Moreover, Jondeau and Rockinger (2004) showed that the fourth-order Taylor expansion of a CARA Utility function leads to excellent approximations of the underlying utility function in the framework of optimal portfolio allocation even under large departure from normally distributed portfolio returns. On the contrary, the mean-variance approach yields large deviation from the optimal portfolio constructed by direct expected utility maximization.

As a result, let consider the approximation of the utility function by a fourth-order Taylor expansion as satisfactory. Beside using the variance as a risk and uncertainty measure, this approach also incorporates skewness and kurtosis. Since the traditional CAPM does not consider the latter, the model needs a correction by introducing two more factors.

Approximating a utility function as a function of the expected return, the standard deviation, the skewness and the kurtosis of the portfolio return distribution function, the maximization problem is:

$$\begin{aligned} \max_{x_p, x_0} \quad & \Phi(\mu_P, \sigma_p^2, s_p, k_p) \\ s.t. \quad & \sum_{i=1}^{n} x_{ip} = 1 - x_0 \end{aligned} \tag{2.3.2}$$

with

$$\mu_P = x_0 r_0 + E(x_p' r) = x_0 r_0 + x_p' \mu$$

$$\sigma_p^2 = x_p' E[(r-\mu)(r-\mu)'] x_p = x_p' \Sigma x_p$$

$$s_p = x_p' E[(r-\mu)(r-\mu)' \otimes (r-\mu)'](x_p \otimes x_p) = x_p' \Omega_{xp}$$

$$k_p = x_p' E[(r-\mu)(r-\mu)' \otimes (r-\mu)' \otimes (r-\mu)'](x_p \otimes x_p \otimes x_p) = x_p' \Psi_{xp}$$

where $\otimes$ denotes the Kronecker-product, $r = (r_1, \ldots, r_n)'$ the vector of asset returns, $x_p = (x_{1p}, \ldots, x_{np})'$ the vector of portfolio portions invested in a single asset, x_{op} the part invested in the risk-free asset, with return r_0, and

$\mu = (E(r_1), \dots, E(r_n))'$. Ω_{xp} is the vector of co-skewness for the weighting vector x_p and Ψ_{xp} the vector of co-kurtosis respectively. Writing the Lagrangian function for problem 2.3.2 gives:

$$L(x_p, \lambda) = \Phi(\mu_P, \sigma_p^2, s_p, k_p) - \lambda(x_p'\mathbf{1} + x_{0p} - 1)$$

The first order conditions are:

$$\frac{\partial L(x_p, \lambda)}{\partial x_p'} = \Phi_1\mu + 2\Phi_2\Sigma x_p + 3\Phi_3\Omega_{xp} + 4\Phi_4\Psi_{xp} - \lambda\mathbf{1} = 0 \tag{2.3.3}$$

$$\frac{\partial L(x_p, \lambda)}{\partial x_{0p}} = \Phi_1 r_0 - \lambda = 0 \tag{2.3.4}$$

$$\frac{\partial L(x_p, \lambda)}{\lambda} = x_p'\mathbf{1} + x_{0p} - 1 = 0 \tag{2.3.5}$$

where Φ_i is the partial derivative of the i-th variable. Assuming equilibrium, every investor behaves optimal. From equations 2.3.3, 2.3.4 and 2.3.5, let write

$$(\mu - r_0)\mathbf{1} = \alpha_1\Sigma x_p + \alpha_2\Omega_{xp} + \alpha_3\Psi_{xp} \tag{2.3.6}$$

with

$$\alpha_1 = -\frac{2\Phi_2}{\Phi_1}, \qquad \alpha_2 = -\frac{3\Phi_3}{\Phi_1} \qquad \alpha_3 = -\frac{4\Phi_4}{\Phi_1}$$

In order to move from optimal conditions for individuals to the resulting market equilibrium, let refer to Cass and Stiglitz (1970). Under identical agent's probability beliefs, a necessary and sufficient condition to apply a two-fund separation theorem is that all agents have a hyperbolic absolute risk aversion utility function (HARA), i.e. that each investor's risk tolerance is a linear function of his wealth $(-U_i'/U_i'' = a_i + b_i W)$ with the same parameter b_i. In this case the portfolio weights of each portfolio are the same. Summing up all these portfolios, let conclude that the condition in equation 2.3.6 must hold also for the market portfolio. Moreover, let define

$$\beta = \frac{\Sigma x_M}{\sigma_M^2} \tag{2.3.7}$$

$$\gamma = \frac{\Omega_{x_M}}{s_M} \tag{2.3.8}$$

$$\theta = \frac{\Psi_{x_M}}{k_M} \tag{2.3.9}$$

it's possible to write the four moment CAPM as follows:

$$\mu - r_0 = b_1\beta + b_2\gamma + b_3\theta \tag{2.3.10}$$

The coefficients b_i can be understood as the corresponding premia associated with the respective risk. Moving to a single asset, the four moment CAPM model can be written in a more intuitive and comprehensible version:

$$E(r_i) - r_0 = b_1\beta_i + b_2\gamma_i + b_3\theta_i \tag{2.3.11}$$

Referring to the traditional CAPM β_i denotes the systematic beta, γ_i the systematic skewness and θ_i the systematic kurtosis of asset i. It is easy to verify that:

$$\beta_i = \frac{Cov(r_i, r_m)}{\sigma_M^2} \tag{2.3.12}$$

$$\gamma_i = \frac{CoS(r_i, r_m)}{s_M} \tag{2.3.13}$$

$$\theta_i = \frac{CoK(r_i, r_m)}{k_M} \tag{2.3.14}$$

with

$$CoS(X,Y) = E[(X - E(X))(Y - E(Y))^2] \tag{2.3.15}$$

$$CoK(X,Y) = E[(X - E(X))(Y - E(Y))^3] \tag{2.3.16}$$

the corresponding co-moments, and

$$s_M = E[(r_M - E(r_M))]^3 \tag{2.3.17}$$

$$k_M = E[(r_M - E(r_M))]^4 \tag{2.3.18}$$

respectively the skewness and the kurtosis of the market portfolio. Even if the four moment CAPM can be seen as a multi-factor model, the three factors go back to the same root. Indeed, they measure the relation of the asset with the market portfolio concerning the respective risk. Hence, only one appropriated index is needed in contrast to "real" multi-factor models.

In order to estimate the four moment CAPM it's needed to assume that it's possible to observe the market portfolio in some way. The difficulty in approximating and defining the market portfolio refers to the critique of Roll (1977).

As a consequence of the four-moment CAPM, the agent will be rewarded a risk premium not only for the volatility (variance) of the market portfolio but also premia for the market skewness and the market kurtosis. In other words, the asset is correlated with the market portfolio in the sense of the specific order (i.e. the co-moments of higher order are significant).

2.3.1 Risk Premia

The expected excess return of the market portfolio becomes:

$$E(r_m) - r_0 = b_1 + b_2 + b_3$$

Let now discuss about the premia: investors have a positive preference for expected returns and skewness, on the contrary they have an aversion towards high variance (standard deviation) and high kurtosis[1]. As a consequence

$$\Phi_1 > 0, \qquad \Phi_2 < 0, \qquad \Phi_3 > 0, \qquad \Phi_4 < 0$$

and b_i coefficients can be written as follows:

$$b_1 = -\frac{2\Phi_2}{\Phi_1}\sigma_M^2 > 0$$

$$b_2 = -\frac{3\Phi_3}{\Phi_1}s_M >=< 0$$

[1]The individual preferences for higher moments will be studied also in the next sections.

$$b_3 = -\frac{4\Phi_4}{\Phi_1} k_M > 0$$

For the systematic beta, the sign is the same as already determined in the framework of the traditional CAPM. The risk premium that is rewarded for a beta reduction is assumed to be positive ($b_1 > 0$). Higher market risk results in a higher risk premium.

For the systematic skewness the result concerning the sign is ambiguous. Let underline that s_M can take either positive or negative values. As a consequence, b_2 will take the opposite sign of the market skewness. Since agents have a preference for higher values of skewness, a negative market skewness is considered as a risk and will be rewarded with a positive risk premium. When the market skewness is positive, b_2 can be seen as a risk discount. The risk discount is directly related to the fact that a positive market skewness means that positive returns have higher probability than negative returns. The analysis of the sign of b_2 gives an intuitive interpretation of the concepts of risk premia and risk discount.

For the Kurtosis the argument is the same as for the second moment: high kurtosis ("fat tails") is a negative investment incentive and the corresponding kurtosis risk premium will be positive. Hwang and Satchell (1999) showed that the four-moment CAPM yields better results in terms of explanation of cross-sectional returns than the standard mean-variance approach. This is especially true in the case of emerging markets or hedge funds, since skewness and kurtosis are particulary significant in these contexts.

2.4 Four Moment Asset Allocation Problem

Let begin with the investor's asset allocation problem. In general the four moment asset allocation problem cannot be solved analytically. Then, let describe how the Taylor series expansion can be used to approximate the allocation problem. Conditions for the expansion to be convergent are detailed for the utility function under study. Last, portfolio moments are computed from asset return moments.

Let consider an investor who allocates her portfolio in order to maximize the expected utility $U(W)$ over her end-of-period wealth W. The initial wealth is arbitrarily set equal to one. There are n risky assets with return vector $r = (r_1, ..., r_n)'$ and joint cumulative distribution function $F(r_1, ..., r_n)$. End-of-period wealth is given by $W = (1 + r_p)$, with $r_p = x'r$, where the vector $x = (x_1, ..., x_n)'$ represents the fractions of wealth allocated to the various risky assets. Let assume that the investor does not have access to a risk-less asset, implying that the portfolio weights sum to one ($\sum_{i=1}^{n} x_i = 1$). In addition, portfolio weights are constrained to be positive, so that short-selling is not allowed. Formally, the optimal asset allocation is obtained by solving the following problem:

$$\begin{aligned} \max_{x_t} \quad & E_t[U(W_{t+1})] \\ s.t. \quad & \sum_{i=1}^{n} x_{i,t} = 1 \\ x_{i,t} \geq 0 \quad & i = 1, \ldots, n \end{aligned} \tag{2.4.1}$$

The n first-order conditions (FOCs) of the optimization problem are

$$\frac{\partial E[U(W)]}{\partial x} = E[U^{(1)}(W)] = 0 \tag{2.4.2}$$

where $U^{(j)}(W)$ denotes the j-th derivative of U. Let assume that the utility function satisfies the usual properties so that a solution exists and is unique.

On one hand, when the empirical distribution of returns is used, the solution to this problem can be easily obtained, see Levy and Markowitz (1979), Pulley (1981), Kroll et al (1984). On the other hand, when a parametric joint distribution for returns is used, the first order conditions in equation 2.4.2 do not have a closed-form solution.

This approach has been applied to normal iid returns by Campbell and Viceira (1999) or to a regime-switching multivariate normal distribution by Ang and Bekaert (2002). The difficulty for non-normal distributions and in particular for distributions that involve asymmetry and fat tails, is that the required number of quadrature points is likely to increase exponentially with the number of assets. Therefore, solving the optimization problem using numerical integration becomes tricky for more than two or three assets. For more general distributions of returns, Monte-Carlo simulations may be necessary to approximate the expected utility function.

Being interested in measuring the effect of higher moments on the asset allocation, let's approximate the expected utility by a Taylor series expansion around the expected wealth. For this purpose, the utility function is expressed in terms of the wealth distribution, so that

$$E[U(W)] = \int_{-\infty}^{+\infty} U(w) f(w) dw \tag{2.4.3}$$

where $f(w)$ is the probability distribution function of end-of-period wealth, that depends on the multivariate distribution of returns and on the vector of weights x. Hence, the infinite-order Taylor series expansion of the utility function is

$$U(W) = \sum_{k=0}^{\infty} \frac{U^{(k)}(\overline{W})(W - \overline{W})^k}{k!} \tag{2.4.4}$$

where $\overline{W} = E(W) = 1 + x'\mu$ denotes the expected end-of-period wealth, with $\mu = E(r)$ the expected return vector. Under rather mild conditions, see

Lhabitant (1997), the expected utility is given by

$$E[U(W)] = E\left[\sum_{k=0}^{\infty} \frac{U^{(k)}(\overline{W})(W-\overline{W})^k}{k!}\right] = \sum_{k=0}^{\infty} \frac{U^{(k)}(\overline{W})}{k!} E[W-\overline{W}]^k. \tag{2.4.5}$$

Therefore, the expected utility depends on all central moments of the distribution of the end-of-period wealth.

Necessary conditions for the infinite Taylor series expansion to converge to the expected utility have been explored by Loistl (1976) and Lhabitant (1998). The region of convergence of the series depends on the considered utility function. In particular, the exponential or polynomial utility functions do not put any restriction on the wealth range, while the power utility function converges for wealth levels in the range $[0, 2(\overline{W})]$. It is worth emphasizing that such a range is likely to be large enough for bonds and stocks when short-selling is not authorized. In contrast, it may be too small for options, due to their leverage effect. These results hold for arbitrary return distributions. Now, since the infinite Taylor series expansion is not suitable for numerical implementation, a solution is to approximate the expected utility by truncating the infinite expansion at a given order $\overline{k}$. For instance, the standard mean-variance criterion proposed by Markowitz (1952) corresponds to the case $\overline{k} = 2$. More generally, an expansion truncated at $\overline{k}$ provides an exact solution to the expected utility when utility is described by a polynomial function of order $\overline{k}$. This result holds because such a utility function depends only on the first $\overline{k}$ moments of the return distribution and the Taylor expansion is an exact approximation for polynomials of order less than or equal to $\overline{k}$. This avenue has been followed for instance by Levy (1969), Hanoch and Levy (1969), or Jurczenko and Maillet (2001) for $\overline{k} = 3$ (cubic utility function) and by Benishay (1992) and Jurczenko and Maillet (2006) for $\overline{k} = 4$.

However, it is not clear a priori the order the Taylor series expansion should be truncated. For instance, Hlawitschka (1994) provides examples in which, even if the infinite expansion converges, adding more terms may worsen the approximation at a given truncation level. In contrast, Lhabitant (1998) describes an example in which omitted terms are of importance. Some arguments put forward by Ederington (1986) as well as Berényi (2001) suggest that introducing the fourth moment will generally improve the approximation of the expected utility.

It should be noticed, at this point, that the approximation of the expected utility by a Taylor series expansion is related to the investor's preference (or aversion) towards the moments of the distribution, that are directly given by derivatives of the utility function. Scott and Horvath (1980) show that, under the assumptions of positive marginal utility, decreasing absolute risk aversion at all wealth levels together with strict consistency for moment preference, the following inequalities hold:

$$U^{(k)}(W) > 0 \quad \forall W, \quad if \quad k \quad is \quad odd \tag{2.4.6}$$

$$U^{(k)}(W) < 0 \quad \forall W, \quad if \quad k \quad is \quad even \tag{2.4.7}$$

Further discussion on the conditions that yield such moments preferences or aversions may be found in Pratt and Zeckhauser (1987), Kimball (1993), and Dittmar (2002). Brockett and Garven (1998) provide examples indicating that expected utility preferences do not necessarily translate into moment preferences. Under rather mild assumptions, a general condition for the smoothness of the convergence of the Taylor series expansion, so that the inclusion of an additional moment will improve the quality of the approximation, is that preference-weighted odd central moments are not dominated by their consecutive preference-weighted even central moments, so that

$$\frac{U^{(2n+1)}[E(W)]}{2n+1}E[W-E(W)]^{(2n+1)} < -\frac{U^{(2n+2)}[E(W)]}{2n+2}E[W-E(W)]^{(2n+2)}$$

with n integer. In this case, including skewness and kurtosis always leads to a better approximation of the expected utility. Focusing on terms up to the fourth one, the expected utility function can be written as

$$\begin{aligned} E_t[U(W_{t+1})] \approx U(W_t) + U^{(1)}(W_t)E_t[W_{t+1}-W_t] + \\ \frac{1}{2}U^{(2)}(W_t)E_t[W_{t+1}-W_t]^2 + \frac{1}{6}U^{(3)}(W_t)E_t[W_{t+1}-W_t]^3 + \\ \frac{1}{24}U^{(4)}(W_t)E_t[W_{t+1}-W_t]^4 + O(W^4) \end{aligned}$$

where $O(W^4)$ is the Taylor remainder. Let define the expected return, variance, skewness, and kurtosis of the end-of-period return as

$$\mu_P = E[r_p] = x'\mu$$

$$\sigma_p^2 = E[(r_p-\mu_P)^2] = E[(W-\overline{W})^2]$$

$$s_p^3 = E[(r_p-\mu_P)^3] = E[(W-\overline{W})^3]$$

$$k_p^4 = E[(r_p-\mu_P)^4] = E[(W-\overline{W})^4].$$

Hence, the expected utility is simply approximated by the following preference function

$$\begin{aligned} E[U(W)] \approx U(\overline{W}) + \frac{1}{2}U^{(2)}(\overline{W})\sigma_p^2 + \\ \frac{1}{3!}U^{(3)}(\overline{W})s_p^3 + \frac{1}{4!}U^{(4)}(\overline{W})k_p^4. \end{aligned}$$

Under conditions established by Scott and Horvath (1980), the expected utility depends positively on expected return and skewness and negatively on variance and kurtosis. Let consider now the CARA (for Constant Absolute Risk Aversion) utility function. The CARA utility function is defined by:

$$U(W) = -\exp(-\lambda W)$$

where λ measures the investor's constant absolute risk aversion. This specification has been widely used in the literature because of the appealing interpretation of the associated parameter. The approximation for the expected utility is given by

$$E[U(W)] \approx -\exp(-\lambda\overline{W})\left[1+\frac{\lambda^2}{2}\sigma_p^2+\frac{\lambda^3}{3!}s_p^3+\frac{\lambda^4}{s!}k_p^4\right] \tag{2.4.8}$$

After simplifications, the FOCs (first order conditions) can be written as:

$$\mu\left[1+\frac{\lambda^2}{2}\sigma_p^2+\frac{\lambda^3}{3!}s_p^3+\frac{\lambda^4}{s!}k_p^4\right]=\frac{\lambda}{2}\frac{\partial\sigma_p^2}{\partial x}+\frac{\lambda^2}{3!}\frac{\partial s_p^3}{\partial x}+\frac{\lambda^3}{4!}\frac{\partial k_p^4}{\partial x} \tag{2.4.9}$$

Optimal portfolio weights can be obtained alternatively by maximizing expression 2.4.8 or by solving equalities 2.4.9. Equation 2.4.9 reveals that computing this expression would be rather simple if the variance, skewness, and kurtosis of the portfolio return and their derivatives are known.

Let briefly describe how the moments of a portfolio return can be expressed in a very convenient way and how their derivatives may be obtained. First the definition of the (n, n^2) co-skewness matrix is needed

$$M_3 = E[(r-\mu)(r-\mu)' \otimes (r-\mu)'] = \{s_{ijk}\}$$

The (n, n^3) co-kurtosis matrix can be defined as

$$M_4 = E[(r-\mu)(r-\mu)' \otimes (r-\mu)' \otimes (r-\mu)'] = \{k_{ijkl}\}$$

The entries of M_3 and M_4 are respectively

$$\{s_{ijk}\} = E[(r_i-\mu_i)(r_j-\mu_j)(r_k-\mu_k)] \quad i,j,k = 1,\ldots,n$$

$$\{k_{ijkl}\} = E[(r_i-\mu_i)(r_j-\mu_j)(r_k-\mu_k)(r_l-\mu_l)] \quad i,j,k,l = 1,\ldots,n$$

Such notation has been used by Harvey et al. (2000), Prakash et al. (2003). It should be noticed that, because of certain symmetries, not all the elements of these matrices need to be computed. The dimension of the covariance matrix is (n, n), but only $n(n+1)/2$ elements have to be computed. Similarly, the co-skewness and co-kurtosis matrices have dimensions (n, n^2) and (n, n^3), but involve only $n(n+1)(n+2)/6$ elements and $n(n+1)(n+2)(n+3)/24$ different elements respectively.

Now, using these notations, moments of the portfolio return can be computed in a very tractable way. For a given portfolio weight vector, expected return, variance, skewness, and kurtosis of the portfolio return are, respectively:

$$\mu_P = x'\mu$$

$$\sigma_p^2 = x'M_2x$$

$$s_p^3 = x'M_3(x \otimes x)$$

$$k_p^4 = x'M_4(x \otimes x \otimes x)$$

where M_2 denotes the usual (n, n) covariance matrix. The moments of the portfolio returns may also be expressed as follows:

$$\sigma_p^2 = E\left[\sum_{i=1}^{n} x_i(r_i - \mu_i)(r_p - \mu_P)\right] = x'\Sigma_p$$

$$s_p^3 = E\left[\sum_{i=1}^{n} x_i(r_i - \mu_i)(r_p - \mu_P)^2\right] = x'S_p$$

$$k_p^4 = E\left[\sum_{i=1}^{n} x_i(r_i - \mu_i)(r_p - \mu_P)^3\right] = x'K_p$$

where

$$\Sigma_p = E\left[(r_i - \mu_i)(r_p - \mu_P)\right] = M_2x$$

$$S_p = E\left[(r_i - \mu_i)(r_p - \mu_P)^2\right] = M_3(x \otimes x)$$

$$K_p = E\left[(r_i - \mu_i)(r_p - \mu_P)^3\right] = M_3(x \otimes x \otimes x)$$

are the $(n, 1)$ vectors of covariances, co-skewness, and co-kurtosis between the asset returns and the portfolio return respectively. These notations are obviously equivalent to the previous ones and they offer the advantage of requiring only small-dimensional vectors. Notations above allow a straightforward computation of the derivatives with respect to the weight vector, that is:

$$\frac{\partial \mu_P}{\partial x} = \mu$$
$$\frac{\partial \sigma_p^2}{\partial x} = 2M_2 x$$
$$\frac{\partial s_p^3}{\partial x} = 3M_3(x \otimes x)$$
$$\frac{\partial k_p^4}{\partial x} = 4M_4(x \otimes x \otimes x)$$

Equations 2.4.9 can thus be rewritten as

$$\mu - \delta_1(x)[M_2 x] + \delta_2(x)[M_3(x \otimes x)] - \delta_3(x)[M_4(x \otimes x \otimes x)] = 0$$

where δ_1, δ_2, and δ_3 are non-linear functions of x, such that $\delta_i(x) = \delta^i/(i!A)$, $i = 1, 2, 3$, with $A = 1 + \frac{\lambda^2}{2}\sigma_p^2 - \frac{\lambda^3}{3!}s_p^3 + \frac{\lambda^4}{4!}k_p^4$. The n equations in condition 2.4.9 can be easily solved numerically, using a standard optimization package. The difficulty to solve this problem is not of the same order as compared to problems involving numerical integration of the utility function. This approach provides an alternative way of solving the asset allocation problem to the PGP approach developed by Lai (1991) and Chunhachinda et al. (1997). The main advantage of the approach proposed here is that weights attributed to the various portfolio moments in equation 2.4.9 are selected on the basis of the utility function, while they are arbitrarily chosen in the PGP approach. Solving equation 2.4.9 also provides an alternative to the rather time-consuming approach based on maximizing the expected utility numerically. Here, a very accurate solution is obtained in just a few seconds, even in the case of a large number of assets. The price to pay is that the focus is put on a finite number of moments only.

2.4.1 Model For Returns

Following Joundeau and Rockinger (2003) a conditional set-up that incorporates most statistical features of stock market returns is proposed. First, the model takes into account properties of volatility clustering, see Engle (1982), and time-varying correlations, see Engle and Sheppard (2002). Second, in order to describe asymmetry and fat-tails of asset returns distributions a multivariate skewed Student t distribution is proposed for assets returns, see Hansen (1994) for the univariate case, Bauwens and Laurent (2002) for the multivariate case.

Let define the following notations: $r_{i,t}$ is the rate of return of asset i from time $t-1$ to time t, in excess of the risk-free rate. Let $\mu_{i,t}$ be the expected excess return of asset i conditional on information available at time $t-1$. Then, $\epsilon_{i,t} = r_{i,t} - \mu_{i,t}$ is the unexpected return of asset i. Let define $r_t = (r_{1,t}, \ldots, r_{n,t})'$ the vector of asset returns, and $\epsilon_t = (\epsilon_{1,t}, \ldots, \epsilon_{n,t})'$ the vector of unexpected returns. $\sigma_{i,t}$ is the conditional variance of $r_{i,t}$. $\sigma_{ij,t}$ denotes the conditional covariance between $r_{i,t}$ and $r_{j,t}$. The conditional covariance matrix is denoted $\Sigma_t = \{\sigma_{ij,t}\}$.

2.4.2 The DCC Model

Let start with a description of the dynamics of the first two moments of the return distribution. The DCC model is able to capture both volatility clustering and persistence in correlation. The conditional mean of returns is described as an AR(1) dynamic to capture the possible first-order serial correlation of returns:

$$r_t = \mu + \varphi r_{t-1} + \Sigma_t^{\frac{1}{2}} z_t \tag{2.4.10}$$

where μ is a vector of size $(n,1)$, φ is a diagonal matrix and z_t is a standardized residual drawn from multivariate Sk-t distribution. The conditional covariance matrix is

$$\Sigma_t = D_t \Gamma_t D_t \tag{2.4.11}$$

$$D_t = \begin{bmatrix} \sqrt{\sigma_{11,t}} & 0 & \dots & 0 \\ 0 & \sqrt{\sigma_{22,t}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \sqrt{\sigma_{nn,t}} \end{bmatrix}$$

$$\sigma_{ii,t} = \omega_i + \beta_i \sigma_{ii,t-1} + \gamma_i \epsilon_{i,t-1}^2 \qquad i = 1, \dots, n$$

$$\Gamma_t = [diag(Q_t)]^{-1} Q_t [diag(Q_t)]^{-1}$$

$$Q_t = (1 - \delta_1 - \delta_2)(\overline{Q}) + \delta_1 (u_{t-1} u'_{t-1}) + \delta_2 Q_{t-1} \tag{2.4.12}$$

$$\overline{Q} = \begin{bmatrix} 1 & \rho_{12} & \dots & \rho_{1n} \\ \rho_{12} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho_{n-1,n} \\ \rho_{1n} & \dots & \rho_{n-1,n} & 1 \end{bmatrix}$$

where $\epsilon_t = \Sigma^{\frac{1}{2}} z_t$ is the vector of error terms, $u_t = D_t^{-1} \epsilon_t$ the vector of standardized errors, $diag(Q_t)$ denotes the diagonal of the matrix Q_t and $\overline{Q}$ is the unconditional covariance matrix of u_t. In equation 2.4.11 the covariance matrix Σ_t is a function of the univariate standard deviations in D_t and of the conditional correlation matrix Γ_t, where each conditional variance $\sigma_{ii,t}$ is represented by a GARCH model. Note that the correlation matrix Γ_t is time-varying and that the restriction $\gamma_i + \beta_i < 1$ guarantees stationarity of the variance process. Other restrictions are imposed on the values of the coefficients δ in order to assure the positive definiteness of the conditional correlation matrix: in particular, $\delta_1 \geq 0$, $\delta_2 \leq 1$ and $\delta_1 + \delta_2 \leq 1$.

2.4.3 The Multivariate Sk-t Distribution

To extend the univariate t distribution to the multivariate case it is possible to assume that the χ^2 which appears in the definition of the t variable is the same for each component. This feature could be a strong restriction when analyzing financial data because it imposes the same fat-tails to every component.

A second way to generalize the t distribution to the n-dimensional case is by using a covariance matrix. In this case, the equality of the degree-of-freedom parameters has to be tested in a second step. Let assume that z_t is drawn from the multivariate Sk-t distribution as

$$t(z_t|\nu_1,\ldots,\nu_n,\lambda_1,\ldots,\lambda_n) = \prod_{i=1}^{n} b_i c_i \left(1+\frac{\zeta_{i,t}^2}{\nu_i-2}\right)^{-\frac{\nu_i+1}{2}} \tag{2.4.13}$$

where

$$\zeta = \begin{cases} \frac{(b_i z_{i,t}+a_i)}{(1-\lambda_i)} & if \quad z_{i,t} < -\frac{a_i}{b_i} \\ \frac{(b_i z_{i,t}+a_i)}{(1+\lambda_i)} & if \quad z_{i,t} \geq -\frac{a_i}{b_i} \end{cases}$$

$$a_i = 4\lambda_i c_i \frac{\nu_i-2}{\nu_i-1}$$

$$b_i^2 = 1+3\lambda_i^2-a_i^2.$$

Fixing the i-th component, λ_i introduces asymmetry on the standard t distribution. The parameters a and b are scale parameters and can be set in order to obtain mean equal to zero and unit variance. Each component has a well-defined distribution if $\nu_i > 2$ and $-1 < \lambda_i < 1$. Note that this distribution is a generalization of the t distribution; when $\lambda_i = 0$ the i-th component shows standard t distribution. If $\nu_i < \infty$ the i-th component shows positive excess kurtosis. In fact, it is well known that the normal distribution can be obtained as a limit of Student t distribution when the degree of freedom parameter tend to infinity. Let define $M_r = E[z_{i,t}^{*r}]$ as the r-th moment of the non-standardized random variable $z^*_{i,t}$ drawn from the univariate Sk-t distribution. The first four moments of the returns distribution can be written as

$$M_{i,1} = 4c_i\lambda_i\frac{\nu_i-2}{\nu_i-1} = a_i$$

$$M_{i,2} = 1+3\lambda_i^2 = b_i^2+a_i^2$$

$$M_{i,3} = 16c_i\lambda_i(1+\lambda_i^2)\frac{(\nu_i-2)^2}{(\nu_i-1)(\nu_i-3)} \qquad if \qquad \nu_i > 3$$

$$M_{i,4} = 3\frac{\nu_i-2}{\nu_i-4}(1+10\lambda_i^2+5\lambda_i^4) \qquad if \qquad \nu_i > 4$$

As a consequence the moments of the standardized variable $z_{i,t} = \frac{(z^*_{i,t}-a_i)}{b_i}$ are respectively

$$\mu_i^{(1)} = 0$$

$$\mu_i^{(2)} = 1$$

$$\mu_i^{(3)} = \frac{M_{i,3}-3a_iM_{i,2}+2a_i^3}{b_i^3}$$

$$\mu_i^{(4)} = \frac{M_{i,4}-4a_iM_{i,3}+6a_i^2M_{i,2}-3a_i^4}{b_i^4}$$

where $\mu_i^{(r)} = E[z^r_{i,t}]$. The skewness and kurtosis are non-linear functions of the parameters λ_i and ν_i that describe respectively the asymmetry and the degrees-of-freedom.

2.4.4 Higher Moments Dynamic

Time variability in higher moments has been analyzed more directly in Hansen (1994), Harvey and Siddique (1999), and Jondeau and Rockinger (2003). Parameters λ_i and ν_i following recent contributions. For the asymmetry parameter, Ang and Bekaert (2002) show that bearish and bullish markets tend to be persistent, suggesting that there may be some clustering in this parameter. Second, the lower the degree-of-freedom parameter, the higher the probability of extreme events in market returns. As argued by Das and Uppal (2004), such extreme events are not likely to be persistent. As a consequence, after a large shock, it's more likely to expect a decrease in kurtosis. In other words the degree-of-freedom parameter is expected to be negatively correlated with the size of shocks. Let underline that in case of a distribution very close to the normal one, the parameter $\nu_{i,t} \to +\infty$, and so, it could be easier to estimate the inverse $\frac{1}{\nu_{i,t}}$.

For higher moments dynamic the following models are proposed:

$$\frac{1}{\tilde{\nu}_{i,t}} = \frac{1}{\tilde{\nu}_i} + b_1 \sum_{i=1}^{p} \omega_i |\epsilon_{t-i}| \tag{2.4.14}$$

$$\tilde{\lambda}_{i,t} = \tilde{\lambda}_i + b_2 \sum_{i=1}^{p} \omega_i \epsilon_{t-i} \tag{2.4.15}$$

where $\tilde{\nu}_{i,t}$ and $\tilde{\lambda}_{i,t}$ are respectively mapped into $[2, +\infty) \times [-1, 1]$ and $\omega_i = 1 - \frac{i}{p}$ is the weight on lag i.

The dynamic of the degree-of-freedom parameter $\nu_{i,t}$ depends on the absolute value of residuals, because it translates the heaviness of the distribution's tails regardless of the sign of shocks over the recent period. Since the degree-of-freedom parameter is very large in the case of very small recent shocks, $\tilde{\nu}_i$ can be set to a large value in order to describe normality in asset returns.

In contrast, the dynamic of the asymmetry parameter naturally depends on signed residuals, $\lambda_{i,t}$ being likely to reflect the sign and size of shocks over the recent period. Equations 2.4.14 and 2.4.15 look like ARCH(p) models because of the introduction of some lags in the function of unexpected returns. It is assumed that parameters b_1 and b_2 are the same for all markets. This assumption is made in order to avoid the curse of dimensionality, for example see the DCC model of Engle and Sheppard (2002). Jondeau and

Rockinger (2003) showed that different estimation of parameters b_1 and b_2 doesn't change significantly the results. Moreover, it can increase the uncertainty related to the estimation of the parameters.

2.4.5 Parameters Estimation

In order to solve the asset allocation problem a set of parameters need to be estimated. Let call this set of parameters θ, with $\theta = (\mu_i, \varphi_i, \omega_i, \beta_i, \gamma_i, (i = 1, \ldots, n), \delta_1, \delta_2, \rho_{jk}, (1 \leq j \leq k \leq n))$. This first set of parameters refers to the DCC model. A second set of parameters needs to be estimated. Let call this set κ. The parameters in κ refer to the shape of the distribution. Assuming that innovations are drawn from a multivariate Sk-t distribution with constant shape parameters, κ becomes $\kappa = (\lambda_1, \ldots, \lambda_n, \nu_1, \ldots, \nu_n)$. In contrast, if the shape parameters of the multivariate Sk-t distribution are assumed to be time varying, κ becomes $\kappa = (\lambda_1, \ldots, \lambda_n, b_1, b_2)$.

Let now define the sample log-likelihood function of the DCC model, see section 2.4.11, when returns are drawn from a multivariate Sk-t distribution as

$$\begin{aligned}\ln L(r_1, \ldots, r_t|\theta, \kappa) &= \sum_{t=1}^{T} \ln \left[t \left(\sigma_t(\theta)^{-\frac{1}{2}} (r_t - \mu_t(\theta))|\kappa \right) \right] \\ &= \sum_{t=1}^{T} \ln \left[t(z_t|\kappa) \right] - \frac{1}{2} \sum_{t=1}^{T} \ln |\sigma_t(\theta)| \qquad (2.4.16)\end{aligned}$$

where T is the sample size and $(z_t|\kappa)$ is defined as in equation 2.4.13. The maximization of equation 2.4.16 leads to the log-likelihood estimation of the parameters.

For large dimensional systems, the estimation can be significantly speed up by performing the estimation in two steps. In the first step, the quasi-ML (maximum likelihood) estimation of the univariate conditional mean and variance equations is obtained assuming normality. The unconditional correlation matrix of standardized residuals is then used to estimate the matrix $\overline{Q}$. In the second step, the parameters referring to the dynamics of correlation (δ_1 and δ_2) and to the shape of the distribution κ are estimated simultaneously. As shown by Joundeau and Rockinger (2003), the two step estimation procedure gives similar results to the estimation obtained by the direct maximization of the log-likelihood function.

Chapter 3

Four Moment Asset Allocation Model: a Proposal

3.1 Introduction

In this chapter a four moment asset allocation model is proposed. Some assumptions are made in order to simplify the optimization model and to obtain a closed form solution for the optimal portfolio. In particular, the key assumption concerns the representation of skewness and kurtosis. The obtained optimal portfolio is a generalization of the classical two moments optimal portfolio, see Markowitz (1952). This generalization permits to write the optimal portfolio as the sum of three portfolios: the first one is the mean-variance optimal portfolio, the second one depends on the skewness only and the third one on the kurtosis only. When the kurtosis is equal to 3, i.e. no kurtosis in excess from the case of normal distribution returns is present, then the optimal portfolio is composed only by the mean variance one and the component due to the skewness. This portfolio is the mean-variance-skewness optimal portfolio as obtained in Gamba and Rossi (1998). Therefore, the four moments model moves a step forward in the direction proposed by Gamba and Rossi (1998).

The general idea behind this model is that in presence of skewness and kurtosis the optimal portfolio can present an higher variance compared to the mean-variance optimal one due to the investor preferences for skewness and kurtosis. The problem of taking into account higher moments in the definition of the individual utility function is not directly treated in the next sections. The preference structure is easily defined starting from the idea that an investor has preferences for mean and positive skewness while avoids high volatility and kurtosis. This is intuitive: an higher value of skewness means that there's an higher probability of positive returns, while, kurtosis is a dispersion measure as variance, and so, a higher value of it could mean an increasing risk for the portfolio, see Horvath and Scott (1980).

Let underline that the evidence that returns of financial activities are not normally distributed is not enough to justify the introduction of a portfolio selection model that considers also the the third and fourth moment. It has to be assumed also that the individual utility function depends on the higher order moments.

Some empirical evidence suggested that mean-variance criterion results in allocations could be good also when returns are non normal, Levy and Markowitz (1979), Pulley (1981), and Kroll, Levy and Markowitz (1984). An explanation of the good performance of the mean-variance criterion may be that the returns are driven by an elliptical distribution, for which the mean-variance approximation of the expected utility remains good for all utility functions (Chamberlain (1983)). In contrast, under large departure from normality, in particular when the distribution is severely asymmetric, Chunachinda et al. (1997), Athayde and Flôres (2004) and Jondeau and

Rockinger (2004) show that the mean-variance criterion can lead to unsatisfactory results. In such a case, a three- or four-moment optimization strategy can improve the results.

The model proposed in this chapter is very general and it permits to express the optimal portfolio as a function of the first four moments of the returns distribution. Moreover, the result of the optimization problem is the generalization of the efficient frontier in the E-V-S-K framework, where E-V-S-K stands for expected return-variance-skewness-kurtosis.

3.2 Assumptions

Let now define the assumptions that will be used in the next section. Let assume that $n+1$ assets are available, n with random return distributions and one with a deterministic return, i.e. the risk-less asset. Let define

- x : the column vector of portfolio weights.
- r : the column vector of returns with mean, standard deviation, skewness and kurtosis respectively represented by the vectors μ, σ, ξ and k.
- μ_0 : the risk-less return.
- y : a random variable such that $E(y) = 0$, $E(y^2) = \sigma_y^2$, $E(y^3) = \xi_y^3 \neq 0$ and $k_y^4 = E(y^4) - 3 = 0$.
- z : a random variable such that $E(z) = 0$, $E(z^2) = \sigma_z^2$, $E(z^3) = \xi_z^3 = 0$ and $k_z^4 = E(z^4) - 3 \neq 0$[1].

[1]The variable k_z^4 is the excess kurtosis of the random variable z. Considering excess kurtosis implies to obtain negative values for the variable k_z^4.

- y and z are assumed to be independent random variables.

- ϵ : a column vector where the entries are Gaussian variables conditional to y and z with $E(\epsilon|y,z) = 0$ and conditional covariance matrix $C = [E(\epsilon\epsilon'|y,z)]$.

- b, t : two column vector of real numbers.

From the previous assumptions the role of the two random variables y and z is clear. Note that, y is assumed to have skewness different from zero and kurtosis equal to 3. The value of the kurtosis is chosen not to give any a contribution to the kurtosis of the returns due to y. Moreover, note that for $E(y^3)$ is just asked to be different from 0 and the sign of the skewness itself is not important. In other words, as it will be more clear in the following, the skewness of the variable y only play the role of a numeraire, useful to measure the assets skewness in a proportional way.

Similarly, the same considerations are valid for the random variable z. In this case, the kurtosis of variable z play the role of a numeraire[2] to measure assets returns kurtosis. As a consequence, in this framework it is indifferent to directly calculate the kurtosis or express it as excess kurtosis with respect to 3, where 3 is the value of kurtosis for a random variable with normal distribution. Obviously, when considering excess kurtosis, it's possible to obtain also negative values.

Following Ingersoll (1987), let assume that the vector of the returns r can be written as follows:

$$r = \mu + \epsilon + by + tz \tag{3.2.1}$$

[2]In this case k_z plays the role of a numeraire only when excess kurtosis to the value three is taken into account.

Note that the previous equation it's a generalization of the representation of returns made by Ingersoll (1987) that permits to take into account also the kurtosis. Moreover, note that ϵ describes the Gaussian component of the returns while the variables y and z add separately the asymmetry (skewness) and the fat-tails (kurtosis) to the returns distribution. Moreover, let recall that the 2 random variables y and z are assumed to be independent. This assumption plays a central role in the following model because it allows to handle separately the skewness and the kurtosis, y and z being the only sources of asymmetry and fat taildness of the returns. The implications of the last assumption will be underlined more specifically in the empirical section. Let now define the covariance matrix of returns as:

$$D = [E(\epsilon\epsilon'|y,z) + bb'\sigma_y^2 + tt'\sigma_z^2] \tag{3.2.2}$$

The covariance matrix D is positive definite being the sum of a positive definite matrix, the covariance matrix of ϵ conditional to y and z, and two semi-positive definite matrices by construction, the results being of outer products. Furthermore, the covariance matrix D is a non-singular matrix.

The co-skewness between i, j, l assets and the co-kurtosis between i, j, l, m assets can be defined as:

$$E[(r_i - \mu_i)(r_j - \mu_j)(r_l - \mu_l)] = \xi_{i,j,l} = b_i b_j b_l \xi_y^3 \tag{3.2.3}$$

$$E[(r_i - \mu_i)(r_j - \mu_j)(r_l - \mu_l)(r_m - \mu_m)] = k_{i,j,l,m} = t_i t_j t_l t_m k_z^4 \tag{3.2.4}$$

where the indexes i, j, l, m represent the entries of the vectors. Note that at least three assets are needed in order to define the co-skewness and at least four for the definition of co-kurtosis. Moreover, the definitions in equations 3.2.3 and 3.2.4 are linear functions of ξ_y^3 and k_z^4.

In particular, due to the assumptions made on the returns, the moments of the portfolio can be defined as follows

$$\mu_P = x'\mu + (1 - \mathbf{1}'x)\mu_0$$

$$\sigma_P^2 = x'Dx$$

$$s_P^3 = x'b\xi_y$$

$$k_P^4 = x'tk_z.$$

Let underline that the first and second moments are computed like in the mean-variance framework while skewness and kurtosis are linear functions respectively of the skewness of y and the kurtosis of z.

The introduction of the third and fourth moments has strong implications on the definition of the structure of preferences. In general it is assumed that the agent has a preference for high values of skewness and for low values of kurtosis. Formally, let define a stochastic dominance rule (E-V-S-K: expected return, variance, skewness, kurtosis).

Definition 3.2.1. *Return r_A E-V-S-K dominates r_B ($r_A >_{EVSK} r_B$) if $\mu_A \geq \mu_B$ and $\sigma_B \geq \sigma_A$ and $\xi_A \geq \xi_B$ and $k_B \geq k_A$.*

The E-V-S-K dominance is a simple generalization of the classical second order stochastic dominance. As shown in section 1.4, it's also possible to define the general concept of stochastic dominance of order n.

Under suitable regularity conditions on the individual utility function U, the E-V-S-K can be characterized as $U' > 0$, $U'' < 0$, $U''' > 0$ and $U'''' < 0$. Despite of this, it's always possible to find a couple r_A and r_B and a utility function U^* such that $r_A >_4 r_B$ but $U^*(r_B) \geq U^*(r_A)$. In other words, in general there is no general compatibility between expected utility theory and stochastic dominance, see for example Brockett and Kahane (1992). This incompatibility doesn't affect the proposed model for two order of reasons. First, it is always possible to restrict the set of Von Nuemann-Morgenstern expected utility functions in order to obtain a correspondence with E-V-S-K dominance rule. Second, as it will be clearer in the next pages, the model doesn't need to assume a particular utility function. The optimization problem can be solved explicitly: the optimal solution then is a function of the first four moments of returns distribution.

3.3 Four Moment Optimal Portfolio

In this section the optimal portfolio in a four-moment framework is derived. The optimal portfolio is the one that minimize the kurtosis subject to expected return, variance, skewness and budget constraints. In order to simplify the calculations the optimization problem will be written as a maximization problem. Short sellings are allowed in the following model and negative portfolio weights have the usual interpretation of short selling.

3.3.1 No Risk-less Asset

Let consider the problem with no risk-less asset. The rational agent, according to the preference to skewness and kurtosis, chooses a portfolio that minimizes the kurtosis, given the mean, skewness and kurtosis . The objective function can be written as:

$$F(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{l=1}^{n} \sum_{m=1}^{n} x_i x_j x_l x_m k_{i,j,l,m} \tag{3.3.1}$$

where x is the column vector of portfolio weights. Let underline that in order to perform the four moment allocation model at least four assets are needed. Using the assumptions on returns distribution, the objective function in equation 3.3.1 can be expressed in an equivalent form as

$$F(x) = (x't)^4 k_z^4$$

The optimization problem given the expected return μ_P, the variance σ_P^2 and the skewness ξ_P^3 is

$$\max_x \quad -x'tk_z \tag{3.3.2}$$

$$s.t. \quad x'Dx = \sigma_P^2 \tag{3.3.3}$$

$$x'\mu = \mu_P \tag{3.3.4}$$

$$x'\mathbf{1} = 1 \tag{3.3.5}$$

$$x'b\xi_y = \xi_P \quad \Rightarrow \quad x'b = \frac{\xi_P}{\xi_y}. \tag{3.3.6}$$

The optimization problem presents a linear objective function, three linear constraints and a quadratic constraint. Following the assumptions on the investors's preferences (E-V-S-K dominance), the optimization problem in 3.3.2 can be written equivalently as a minimization problem where the objective function is the portfolio kurtosis. The first constraint in equation 3.3.3 sets the level of portfolio variance. Note that, in the classical mean-variance framework, the agent's problem is written as a quadratic optimization problem with linear constraints. Equation 3.3.4 is the usual constraint on the portfolio expected return. Equation 3.3.5 is the usual budget constraint: no assumptions are made on the sign of the portfolio weights, i.e. short selling is allowed. Equation 3.3.6 is the constraint on skewness. Note that it's possible to write the constraint as in equation 3.3.6 and divide by ξ_y because the variable y is assumed to have skewness different from 0.

Let now define the matrix P as:

$$P = M'D^{-1}M = \begin{bmatrix} a & c & f & p \\ c & d & g & q \\ f & g & e & r \\ p & q & r & s \end{bmatrix}$$

where the $(n, 4)$ matrix M is defined as $M = (\mu, \mathbf{1}, b, t)$. The matrix M is assumed to be of full column rank, i.e. the vectors $\mu, \mathbf{1}, b$ and t are assumed to be linearly independent. The matrix P is symmetric by construction. The entries of matrix P are: $a = \mu' D^{-1}\mu$, $c = \mu' D^{-1}\mathbf{1}$, $f = \mu' D^{-1} b$, $p = \mu' D^{-1} t$, $d = \mathbf{1}' D^{-1}\mathbf{1}$, $g = \mathbf{1}' D^{-1} b$, $q = \mathbf{1}' D^{-1} t$, $e = b' D^{-1} b$, $r = b' D^{-1} t$, $s = t' D^{-1} t$. Let define also the matrices A and P_2 as sub-matrices of P:

$$P_2 = \begin{bmatrix} a & c & f \\ c & d & g \\ f & g & e \end{bmatrix} \qquad A = \begin{bmatrix} a & c \\ c & d \end{bmatrix}.$$

The matrices A and P_2 will be useful in the next pages for further calculations. The matrix A coincides to the matrix A defined in equation 1.6.9 in the mean-variance framework. Let $\psi' = (p \quad q \quad r)$ and $H = \psi' P_2^{-1} \psi$.

Lemma 3.3.1. *If $rank(M) = 4$ then*

$$(s - H) = \frac{detP}{detP_2} > 0.$$

Proof: Let $y = Mx$, by substitution it's possible to write:

$$x'Px = x'M'D^{-1}Mx = y'D^{-1}y.$$

D being a positive definite matrix, then also P is a positive definite matrix, i.e. $det(P) > 0$ and $det(P_2) > 0$ (P and P_2 are leading principal minors of matrix P). Moreover, it's easy to show that $(s - H)det(P_2) = det(P)$, and then $(s-H)>0$. ⋆

Let define the quantity $\sigma^2_{P_2} = \beta' P_2^{-1}\beta$, where $\beta = (\mu_P \quad 1 \quad \frac{\xi_P}{\xi_y})'$. Note that the number $\sigma^2_{P_2}$ is a positive number by construction, P_2^{-1} being a positive definite matrix. As it will be more clear in the next pages, the quantity $\sigma^2_{P_2}$ can be interpreted as a variance. Let now calculate the optimal portfolio for problem 3.3.2.

Proposition 3.3.1. *Given μ_P, $\sigma^2_P \geq \sigma^2_{P_2}$, $\xi_P > 0$ and $k_z < 0$ the optimal portfolio for problem 3.3.2 is*

$$x^* = D^{-1}(\mu \quad \mathbf{1} \quad b)P_2^{-1}\beta + \tag{3.3.7}$$
$$+\sqrt{\frac{\sigma^2_P - \sigma^2_{P_2}}{(s-H)}}\left[D^{-1}t - D^{-1}(\mu \quad \mathbf{1} \quad b)P_2^{-1}\begin{pmatrix} p \\ q \\ r \end{pmatrix}\right]$$

and the optimal kurtosis k^* is

$$k^* = k_z\left[(p \quad q \quad r)P_2^{-1}\begin{pmatrix} \mu_P \\ 1 \\ \frac{\xi_P}{\xi_y} \end{pmatrix} + \sqrt{s-H}\sqrt{\sigma^2_P - \sigma^2_{P_2}}\right]. \tag{3.3.8}$$

Proof: The Lagrangian function for problem 3.3.2 is:

$$L(x,\lambda) = -x'tk_z - \lambda_1(x'Dx - \sigma^2_P) - \lambda_2(x'\mu - \mu_P) - \lambda_3(x'\mathbf{1} - 1) - \lambda_4(x'b\xi_y - \xi_P)$$

The first order conditions for the optimization problem 3.3.2 are:

$$\frac{\partial L}{\partial x} = -tk_z - 2\lambda_1 Dx - \lambda_2\mu - \lambda_3\mathbf{1} - \lambda_4 b\xi_y = \mathbf{0} \tag{3.3.9}$$

$$\frac{\partial L}{\partial \lambda_1} = x'Dx - \sigma^2_P = 0 \tag{3.3.10}$$

$$\frac{\partial L}{\partial \lambda_2} = x'\mu - \mu_P = 0 \tag{3.3.11}$$

$$\frac{\partial L}{\partial \lambda_3} = x'\mathbf{1} - 1 = 0 \tag{3.3.12}$$

$$\frac{\partial L}{\partial \lambda_4} = x'b\xi_y - \xi_P = 0 \tag{3.3.13}$$

The Hessian matrix for problem 3.3.2 is $H_x(L) = \frac{\partial^2 L}{\partial x^2} = -2\lambda_1 D$. D being positive definite, then $-2\lambda_1 D$ is negative definite when $\lambda_1 > 0$. In the following, the sign of λ_1 will be investigated in order to ensure the first order conditions to be sufficient for problem 3.3.2.

From equation 3.3.9, assuming $\lambda_1 \neq 0$, the optimal portfolio can be written as

$$x^* = -\frac{\lambda_4}{2\lambda_1}\xi_y D^{-1}b - \frac{\lambda_2}{2\lambda_1}D^{-1}\mu - \frac{\lambda_3}{2\lambda_1}D^{-1}\mathbf{1} - \frac{k_z}{2\lambda_1}D^{-1}t \tag{3.3.14}$$

The assumption $\lambda_1 \neq 0$ is equivalent to ask the quadratic restriction to hold. Let re-name some parts of the previous equation in order to simplify calculations. Let define the vector $\gamma' = (\gamma_1 \quad \gamma_2 \quad \gamma_3 \quad \gamma_4)$ where

$$\gamma_1 = -\frac{\lambda_2}{2\lambda_1} \qquad \gamma_2 = -\frac{\lambda_3}{2\lambda_1} \qquad \gamma_3 = -\frac{\lambda_4}{2\lambda_1}\xi_y \qquad \gamma_4 = -\frac{k_z}{2\lambda_1} \tag{3.3.15}$$

Let re-write equation 3.3.14 as a function of γ:

$$x^* = D^{-1}M\gamma.$$

Plugging the optimal x^* into the constraints, equations 3.3.3, 3.3.4, 3.3.5 and 3.3.6, the following equations hold:

$$x'Dx = \gamma' M' D^{-1} D D^{-1} M\gamma = \gamma' P\gamma = \sigma_P^2 \tag{3.3.16}$$

$$\mu' x = \mu D^{-1} M\gamma = \mu_P \tag{3.3.17}$$

$$\mathbf{1}x = \mathbf{1}D^{-1}M\gamma = 1 \tag{3.3.18}$$

$$\xi_y b' x = \xi_y b' D^{-1} M\gamma = \xi_P \quad \Rightarrow \quad b' D^{-1} M\gamma = \frac{\xi_P}{\xi_y} \tag{3.3.19}$$

From the 3 linear constraints, 3.3.17, 3.3.18 and 3.3.19, it's possible to obtain γ_1, γ_2 and γ_3 as functions of γ_4:

$$\begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix} = P_2^{-1} \begin{pmatrix} \mu_P \\ 1 \\ \frac{\xi_P}{\xi_y} \end{pmatrix} - P_2^{-1} \begin{pmatrix} p \\ q \\ r \end{pmatrix} \gamma_4.$$

By substitution in the quadratic restriction, it's possible to solve for γ_4:

$$\gamma_4 = \pm\sqrt{\frac{\sigma_P^2 - \sigma_{P_2}^2}{(s-H)}}.$$

From equation 3.3.15, λ_1 can be expressed as:

$$\lambda_1 = -\frac{k_z}{2\gamma_4}. \tag{3.3.20}$$

The first order conditions are sufficient for problem 3.3.2 if $\lambda_1 > 0$. Therefore, for $k_z < 0$ ($k_z > 0$) the value of γ_4 is

$$\gamma_4 = \sqrt{\frac{\sigma_P^2 - \sigma_{P_2}^2}{(s-H)}} \qquad \left(\gamma_4 = -\sqrt{\frac{\sigma_P^2 - \sigma_{P_2}^2}{(s-H)}}\right)$$

It's easy now to write the optimal kurtosis k^* associated with the optimal portfolio x^* just recalling that $k = x'tk_z$ and substituting the optimal portfolio x^*

$$k^* = k_z \left[\begin{pmatrix} p & q & r \end{pmatrix} P_2^{-1} \begin{pmatrix} \mu_P \\ 1 \\ \frac{\xi_P}{\xi_y} \end{pmatrix} + \sqrt{s-H}\sqrt{\sigma_P^2 - \sigma_{P_2}^2} \right] \quad \star. \tag{3.3.21}$$

The optimal portfolio x^* in equation 3.3.9 is the optimal one in a E-V-S-K framework. This representation is pretty obscure and doesn't allow to get any easy insight about it. It's useful then to re-write the optimal portfolio in a form that allows to make some considerations more directly. Moreover, re-writing the optimal portfolio in a more suitable form permits to decompose it in different components that underline the role of skewness and kurtosis.

Let calculate the values of γ as functions of the matrix A instead of P_2. To do this, it is necessary to write $(\gamma_1 \quad \gamma_2)$ as function of γ_3 and γ_4 using the first 2 linear restrictions, equations 3.3.17 and 3.3.18:

$$\begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} = A^{-1} \begin{pmatrix} \mu_P \\ 1 \end{pmatrix} - A^{-1} \begin{pmatrix} f \\ g \end{pmatrix} \gamma_3 - A^{-1} \begin{pmatrix} p \\ q \end{pmatrix} \gamma_4 \tag{3.3.22}$$

Remark 3.3.1. *From Lemma 3.3.1, the following equation holds:*

$$\gamma' P \gamma = \sigma_P^2 > \sigma_{P_2}^2 = (\gamma_1 \quad \gamma_2 \quad \gamma_3)\, P_2 \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix} \tag{3.3.23}$$

A simple interpretation of the Remark 3.3.1 is that an investor with the given structure of preferences for higher moments are willing to choose portfolios with higher variance compared to the mean-variance efficient ones. In other words, such an investor balances the higher values of portfolio variance with a increasing value of skewness and decreasing values of kurtosis.

Let evaluate the value of γ_3 plugging equation 3.3.22 into the quadratic constraint, equation 3.3.3. The value of γ_3 can be written as follows:

$$\begin{aligned} \gamma_3^2 (e-h) = \sigma_{P_2}^2 + 2\gamma_4 (p \quad q) A^{-1} \begin{pmatrix} \mu_P \\ 1 \end{pmatrix} - \\ -(\mu_P \quad 1) A^{-1} \begin{pmatrix} \mu_P \\ 1 \end{pmatrix} - \gamma_4^2 (p \quad q) A^{-1} \begin{pmatrix} p \\ q \end{pmatrix} \end{aligned} \tag{3.3.24}$$

Let re-write equation 3.3.24 in a more useful way:

$$\gamma_3^2 (e-h) = \sigma_{P_2}^2 - \sigma_B^2 \tag{3.3.25}$$

where

$$\sigma_B^2 = [\mu_P \quad 1 \quad \gamma_4 p \quad \gamma_4 q] T \begin{pmatrix} \mu_P \\ 1 \\ \gamma_4 p \\ \gamma_4 q \end{pmatrix} \qquad \mathit{with} \qquad T = \begin{bmatrix} A^{-1} & A^{-1} \\ A^{-1} & A^{-1} \end{bmatrix} \tag{3.3.26}$$

Let note that the quadratic form that defines σ_B^2 is semi-positive definite. In fact, A being positive definite, then A^{-1} is positive definite. Therefore, the matrix T has two positive eigenvalues and two null eigenvalues by construction. It's useful to define also the quantity σ_A^2 as

$$\sigma_A^2 = (\mu_P \quad 1) A^{-1} \begin{pmatrix} \mu_P \\ 1 \end{pmatrix}$$

The quantity σ_A^2 plays a central role in this model because it permits to directly compare the optimal portfolio in E-V-S-K framework with the one obtained by Gamba and Rossi (1998) in a three moments framework. Moreover, note that σ_A^2 is equivalent to the quantity defined in equation 2.2.4.

Let now evaluate the value of γ_3 for $\xi_y > 0$ ($\xi_y < 0$):

$$\gamma_3 = \sqrt{\frac{\sigma_{P_2}^2 - \sigma_A^2}{(e-h)}} \qquad \left(\gamma_3 = -\sqrt{\frac{\sigma_{P_2}^2 - \sigma_A^2}{(e-h)}} \right) \tag{3.3.27}$$

It's possible now to express the optimal portfolio for problem 3.3.2 as a function of matrix A. This allows to directly compare E-V-S-K optimal portfolio with three moment optimal portfolios , see Gamba and Rossi (1998).

Corollary 3.3.1. *Given μ_P, $\sigma_P^2 \geq \sigma_{P_2}^2 \geq \sigma_B^2$, $\xi_P > 0$ and $k_z < 0$ then the* *optimal portfolio for problem 3.3.2 can equivalently be written as:*

$$x^* = D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix}\mu_P \\ 1\end{pmatrix} +$$
$$+\left[D^{-1}b + D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix}f \\ g\end{pmatrix}\right]\sqrt{\frac{\sigma^2_{P_2}-\sigma^2_B}{e-h}} +$$
$$+\left[D^{-1}t + D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix}p \\ q\end{pmatrix}\right]\sqrt{\frac{\sigma^2_{P}-\sigma^2_{P_2}}{s-H}} \tag{3.3.28}$$

3.3.2 Optimal Portfolio Returns Analysis

Let now analyze the properties of the optimal portfolio in equation 3.3.28. The optimal portfolio in equation 3.3.28 is equivalent to the one in equation 3.4.5 and it can be separated as follows:

$$x_1^* = D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix}\mu_P \\ 1\end{pmatrix} \tag{3.3.29}$$

$$x_2^* = \left[D^{-1}b + D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix}f \\ g\end{pmatrix}\right]\sqrt{\frac{\sigma^2_{P_2}-\sigma^2_B}{e-h}} \tag{3.3.30}$$

$$x_3^* = \left[D^{-1}t + D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix}p \\ q\end{pmatrix}\right]\sqrt{\frac{\sigma^2_{P}-\sigma^2_{P_2}}{s-H}} \tag{3.3.31}$$

Portfolio x_1^* is the classical optimal portfolio given μ_p in the M-V framework, see Huang and Litzemberger (1988). The portfolio x_1^* displays the properties $\mathbf{1}'x_1^* = 1$ and $\mu' x_1^* = \mu_P$.

Note that, by construction, portfolios x_2^* has the property $x_2^{*'}\mathbf{1} = 0$. According to Ingersoll (1987), portfolios x_2^* is said to be an arbitrage portfolio. Portfolio x_2^* is useful to add variance to the optimal mean-variance portfolio following the skewness individual preferences. The extra variance due to the skewness preference depends on the quantity $\sigma^2_{P_2} - \sigma^2_B$. Moreover, portfolio x^*_2 is a null vector if and only if $\xi_y = 0$, i.e. $\sigma^2_{P_2} = \sigma^2_B$.

Portfolio x^*_3 displays the property $x^{*'}_3 \mathbf{1} = 0$. According to Ingersoll (1987), portfolio x^*_3 is said to be an arbitrage portfolio. Portfolio x^*_3 is useful to add variance to the optimal mean-variance portfolio following the kurtosis individual preferences. The extra variance due to the kurtosis preference depends on the quantity $\sigma^2_P - \sigma^2_{P_2}$. Moreover, portfolio x^*_3 is a null vector if and only if $k_z = 0$, i.e. $\sigma^2_P = \sigma^2_{P_2}$.

Let underline that the cases when $\sigma^2_{P_2} = \sigma^2_B$ and $\sigma^2_P = \sigma^2_{P_2}$ are trivial. Those two situations occur respectively when $\xi_y = 0$ and $k_z = 0$. In those cases the rank of matrix P is less than 4 and it simply means that skewness or kurtosis or both of them are not taken into account[3].

Let now analyze what happens to the optimal portfolio x^* for particular values of $\sigma^2_{P_2}$. In this way it will be clear that the proposed model is a generalization of classical mean-variance model, see Markowitz (1952), and of the three moments asset allocation model proposed by Gamba Rossi (1998).

- $\sigma^2_P = \sigma^2_{P_2} = \sigma^2_B$.

  In this case, being $\sigma^2_P = \sigma^2_{P_2} = \sigma^2_B$, no extra variance is taken into account. The optimal portfolio x^* in equation 3.3.28 becomes

  $$x^* = D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix} \mu_P \\ 1 \end{pmatrix} \tag{3.3.32}$$

  This portfolio is exactly the mean-variance optimal portfolio, see Huang and Litzenberger (1988).

[3]The assumptions made in section 3.2 require $\xi_y \neq 0$ and $k_z \neq 0$. The only reason of those assumptions is to guarantee that the model takes into account both the third and the fourth moments.

- $\sigma_P^2 > \sigma_{P_2}^2 > \sigma_B^2$.

  In this case, positive extra variance is taken into account. Portfolio x_2^* and x_3^* are both different from the null vector. The optimal portfolio x^* becomes

$$x^* = D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix}\mu_P \\ 1\end{pmatrix} +$$
$$+\left[D^{-1}b + D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix}f \\ g\end{pmatrix}\right]\sqrt{\frac{\sigma_{P_2}^2 - \sigma_B^2}{e-h}} +$$
$$+\left[D^{-1}t + D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix}p \\ q\end{pmatrix}\right]\sqrt{\frac{\sigma_P^2 - \sigma_{P_2}^2}{s-H}}. \qquad (3.3.33)$$

  Note that, when $\sigma_{P_2}^2 \downarrow \sigma_B^2$ almost all the extra variance taken into account is used in order to increase the skewness of the optimal portfolio. On the contrary, when $\sigma_{P_2}^2 \uparrow \sigma_p^2$ the extra variance is almost all used in order to decrease the kurtosis of the optimal portfolio. In other words, by varying $\sigma_{P_2}^2$, it's possible to give different importance to skewness and kurtosis in the optimal portfolio and build the efficient frontier in E-V-S-K framework.

  Let also underline that portfolio $x_1^* + x_2^*$ is different from the optimal portfolio in obtained by Gamba and Rossi (1998) considering the first three moments of returns distribution. This is reasonable because the rate of substitution between skewness and kurtosis preferences is not explicitly defined in this model. To do that, the definition of the individual utility function is needed.

- $\sigma_P^2 > \sigma_{P_2}^2 = \sigma_B^2$.

  This is the case where the extra-variance is used just to decrease the value of the optimal portfolio kurtosis. The optimal portfolio x^* becomes

$$x^* = D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix} \mu_P \\ 1 \end{pmatrix} +$$

$$+\left[D^{-1}t + D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix} p \\ q \end{pmatrix}\right]\sqrt{\frac{\sigma_P^2 - \sigma_{P_2}^2}{s-H}} \tag{3.3.34}$$

Let underline that this model permits to consider mean, variance and kurtosis by themselves without considering the the moments in a certain order.

- $\sigma_P^2 = \sigma_{P_2}^2 > \sigma_B^2$.

This case is very interesting because it directly shows that the proposed model is a generalization not only of the mean-variance model but also of the three moments model proposed by Gamba and Rossi (1998). When $\sigma_P^2 = \sigma_{P_2}^2$ the kurtosis is not taken into account and, at the same time, the quantity σ_B^2 becomes equal to σ_A^2, see equation 2.2.4. In fact, let recall that

$$\gamma_3^2(e-h) = \sigma_{P_2}^2 + 2\gamma_4(p \quad q)A^{-1}\begin{pmatrix} \mu_P \\ 1 \end{pmatrix} -$$

$$-(\mu_P \quad 1)A^{-1}\begin{pmatrix} \mu_P \\ 1 \end{pmatrix} - \gamma_4^2(p \quad q)A^{-1}\begin{pmatrix} p \\ q \end{pmatrix} \tag{3.3.35}$$

but

$$\gamma_4 = \sqrt{\frac{\sigma_P^2 - \sigma_{P_2}^2}{(s-H)}} = 0 \tag{3.3.36}$$

and so

$$\gamma_3^2(e-h) = \sigma_{P_2}^2 - (\mu_P \quad 1)A^{-1}\begin{pmatrix} \mu_P \\ 1 \end{pmatrix} \qquad \Rightarrow \qquad \gamma_3 = \sqrt{\frac{\sigma_{P_2}^2 - \sigma_A^2}{e-h}}$$

The optimal portfolio x^* becomes

$$x^* = D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix} \mu_P \\ 1 \end{pmatrix} +$$
$$+\left[D^{-1}b + D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix} f \\ g \end{pmatrix}\right]\sqrt{\frac{\sigma^2_{P_2} - \sigma^2_A}{e-h}} \qquad (3.3.37)$$

The optimal portfolio x^* in equation coincides with the optimal portfolio in the three moment framework as obtained in Gamba e Rossi (1998), see equation 2.2.3.

Another interesting case is when the optimal portfolio in equation 3.3.28 can be split exactly into the sum of the optimal portfolio in E-V-S framework plus the component due to the kurtosis. Let note that $\sigma^2_B = \sigma^2_A$ not only for $\gamma_4 = 0$. The equality $\sigma^2_B = \sigma^2_A$ also holds when

$$2\gamma_4(p \quad q)A^{-1}\begin{pmatrix} \mu_P \\ 1 \end{pmatrix} - \gamma_4^2(p \quad q)A^{-1}\begin{pmatrix} p \\ q \end{pmatrix} = 0. \qquad (3.3.38)$$

Solving equation 3.3.38, the condition becomes

$$\gamma_4\left[2(p \quad q)A^{-1}\begin{pmatrix} \mu_P \\ 1 \end{pmatrix} - \gamma_4(p \quad q)A^{-1}\begin{pmatrix} p \\ q \end{pmatrix}\right] = 0$$

and so

$$\gamma_4 = \frac{2(p \quad q)A^{-1}\begin{pmatrix} \mu_P \\ 1 \end{pmatrix}}{(p \quad q)A^{-1}\begin{pmatrix} p \\ q \end{pmatrix}}. \qquad (3.3.39)$$

When the condition on γ_4 in equation 3.3.39 holds, the optimal portfolio

x^* becomes

$$x^* = D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix}\mu_P \\ 1\end{pmatrix} + \\ + \left[D^{-1}b + D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix}f \\ g\end{pmatrix}\right]\sqrt{\frac{\sigma^2_{P_2} - \sigma^2_A}{e-h}} + \\ + \left[D^{-1}t + D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix}p \\ q\end{pmatrix}\right]\sqrt{\frac{\sigma^2_{P} - \sigma^2_{P_2}}{s-H}}. \tag{3.3.40}$$

Portfolio x^* is now the sum of the optimal portfolio in E-V-S framework, see equation 3.3.2, plus portfolio x_3^* that translates the kurtosis correction, i.e. the variation of the optimal weights due to the introduction of the kurtosis.

3.3.3 Four Funds Decomposition

Due to the assumptions made in the previous sections[4] it is possible to decompose the optimal portfolio in presence of higher moments into four funds. The result is obtained re-writing the portfolio selection optimization problem in a more suitable way. The four funds have the shape of four vectors that can be used in order to span the space of optimal portfolios.

Lemma 3.3.2. *When the returns follow equation 3.2.1, the efficient set is*

[4]The structure of returns is chosen in order to guarantee the four funds decomposition property, see equation 3.2.1.

spanned by the following 4 vectors:

$$
\begin{aligned}
v_1 &= \frac{D^{-1}\mu}{\boldsymbol{1}'D^{-1}\mu}\\
v_2 &= \frac{D^{-1}\boldsymbol{1}}{\boldsymbol{1}'D^{-1}\boldsymbol{1}}\\
v_3 &= \frac{D^{-1}b}{\boldsymbol{1}'D^{-1}b}\\
v_4 &= \frac{D^{-1}t}{\boldsymbol{1}'D^{-1}t}
\end{aligned}
$$

Proof: Let rewrite the optimization problem as a quadratic optimization problem with linear restrictions.

$$
\begin{aligned}
&\min_x \quad \frac{1}{2}x'Vx\\
&s.t. \quad x'\mu = \mu_P\\
&\qquad x'\mathbf{1} = 1\\
x'b\xi_y = \xi_P \quad &\Rightarrow \quad x'b = \frac{\xi_P}{\xi_y}\\
x'tk_z = k_P \quad &\Rightarrow \quad x't = \frac{k_P}{k_z}
\end{aligned}
$$

where $V = E(\epsilon\epsilon'|y,z)$. From the definition of the covariance matrix, see equation 3.2.2, the matrix V can be written as a function of D as $V = [D - bb'\sigma_y^2 - tt'\sigma_z^2]$.

The new objective function of the optimization problem is a quadratic form in V, with linear constraints that represent respectively portfolio expected return, the budget constraint, portfolio skewness and portfolio kurtosis. Note that the covariance matrix in equation 3.2.2 is positive definite by construction and, therefore, the first order conditions are sufficient for a global minimum point.

The langrangian function for the optimization problem is:

$$L(x,\delta) = \frac{1}{2}x'Dx - \frac{1}{2}\left(\frac{\xi_P}{\xi_y}\right)^2\sigma_y^2 - \frac{1}{2}\left(\frac{k_P}{k_z}\right)^2\sigma_z^2 +$$
$$+\delta_1(\mu_P - x'\mu) + \delta_2(1 - x'\mathbf{1}) + \delta_3\left[\frac{\xi_P}{\xi_y} - x'b\right] + \delta_4\left[\frac{k_P}{k_z} - x't\right]$$

The first order condition for x is:

$$\frac{\partial L}{\partial x} = Dx - \frac{\xi_P}{\xi_y}\sigma_y^2 b - \frac{k_P}{k_z}\sigma_z^2 t - \delta_1\mu - \delta_2\mathbf{1} - \delta_3 b - \delta_4 t = \mathbf{0}$$
$$Dx - \delta_1\mu - \delta_2\mathbf{1} - \left[\frac{\xi_P}{\xi_y}\sigma_y^2 + \delta_3\right]b - \left[\frac{k_P}{k_z}\sigma_z^2 + \delta_4\right]t = 0$$

Solving for x, the optimal solution is:

$$x^* = \delta_1 D^{-1}\mu + \delta_2 D^{-1}\mathbf{1} + \left[\frac{\xi_P}{\xi_y}\sigma_y^2 + \delta_3\right]D^{-1}b + \left[\frac{k_P}{k_z}\sigma_z^2 + \delta_4\right]D^{-1}t \quad (3.3.41)$$

By equation 3.3.41 and using the linear constraints, the optimal portfolio x^* can be written as:

$$x^* = \lambda_1\frac{D^{-1}\mu}{\mathbf{1}'D^{-1}\mu} + \lambda_2\frac{D^{-1}\mathbf{1}}{\mathbf{1}'D^{-1}\mathbf{1}} + \lambda_3\frac{D^{-1}b}{\mathbf{1}'D^{-1}b} + \lambda_4\frac{D^{-1}t}{\mathbf{1}'D^{-1}t}$$

where

$$\lambda_1 = \delta_1(\mathbf{1}'D^{-1}\mu)$$
$$\lambda_2 = \delta_2(\mathbf{1}'D^{-1}\mathbf{1})$$
$$\lambda_3 = \left[\delta_3 + \frac{\xi_P}{\xi_y}\sigma_y^2\right](\mathbf{1}'D^{-1}b)$$
$$\lambda_4 = \left[\delta_4 + \frac{k_P}{k_z}\sigma_z^2\right](\mathbf{1}'D^{-1}t)$$

and $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1$. ⋆

The advantage of the previous procedure is the possibility to directly compare the spanning funds to the ones obtained in the classical mean-variance framework. More precisely, in order to obtain the first two spanning vectors $v_1 = \frac{D^{-1}\mu}{\mathbf{1}'D^{-1}\mu}$ and $v_2 = \frac{D^{-1}\mathbf{1}}{\mathbf{1}'D^{-1}\mathbf{1}}$ equal to the ones of the classical mean-variance framework it's needed to write the optimization problem as the minimization of portfolio variance and add skewness and kurtosis through the constraints.

3.4 With a Risk-less Asset

In the case of presence of a risk-less asset with return μ_0, the optimization problem in E-V-S-K framework changes and it can be written as follows:

$$\max_x \quad -x'tk_z \tag{3.4.1}$$

$$s.t. \quad x'Dx = \sigma_P^2 \tag{3.4.2}$$

$$x'(\mu - \mu_0\mathbf{1}) = \mu_P - \mu_0 \tag{3.4.3}$$

$$x'b = \frac{\xi_P}{\xi_y}. \tag{3.4.4}$$

In the optimization problem the objective function is linear as two of the restrictions while the first restriction is a quadratic form in x. The quadratic constraint in equation 3.4.2 sets the desired value of portfolio variance. Equation 3.4.3 is the usual constraint on portfolio expected return in presence of a risk-less asset. Equation 3.4.4 is the restriction on portfolio skewness and it's equivalent to the one in equation 3.3.6.

Note that in this case the restrictions for problem 3.4.1 are just 3. This depends on the fact that the weight on the risk-less asset is calculated as the residual after allocating in the risky assets. By intuition, the risk-less return μ_0 is deterministic and so it doesn't give any contribution to the standardized central moments of the returns distribution.

Let now define the matrix Q as:

$$Q = N'D^{-1}N = \begin{bmatrix} m_1 & l_1 & f_1 \\ l_1 & e & g_1 \\ f_1 & g_1 & s \end{bmatrix}$$

where the $(n, 3)$ matrix N is defined as $N = [(\mu - \mu_0\mathbf{1}), b, t]$. The matrix Q is assumed to be of full column rank, i.e. the vectors $(\mu - \mu_0\mathbf{1}), b$ and t are assumed to be linearly independent. The matrix Q is symmetric and positive definite by construction. The entries of matrix Q are: $m_1 = (\mu - \mu_0\mathbf{1})' D^{-1}(\mu - \mu_0\mathbf{1})$, $l_1 = (\mu - \mu_0\mathbf{1})' D^{-1} b$, $f_1 = (\mu - \mu_0\mathbf{1})' D^{-1} t$, $e = b' D^{-1} b$, $g_1 = b' D^{-1} t$, $s = t' D^{-1} t$. Let define also the matrix Q_2 as a sub-matrix of Q:

$$Q_2 = \begin{bmatrix} m_1 & l_1 \\ l_1 & e \end{bmatrix}$$

The matrix Q_2 will be useful in the next pages for further calculations.

Let $\iota' = (f_1 \quad g_1)$ and $H_1 = \iota' Q_2^{-1} \iota$.

Lemma 3.4.1. *If $rank(N) = 3$ then*

$$(s - H_1) = \frac{detQ}{detQ_2} > 0.$$

Proof: Let $y = Nx$, by substitution it's possible to write:

$$x'Qx = x'N'D^{-1}Nx = y'D^{-1}y.$$

D being a positive definite matrix, then also Q is a positive definite matrix, i.e. $det(Q) > 0$ and $det(Q_2) > 0$ (Q and Q_2 are leading principal minors of matrix Q). Moreover, it's easy to show that $(s - H_1)det(Q_2) = det(Q)$, and then $(s - H_1) > 0$. $\star$

Let define the quantity $\sigma^2_{Q_2} = \beta' Q_2^{-1} \beta$, where $\beta = (\mu_p \quad 1 \quad \frac{\xi_p}{\xi_y})'$. Note that the number $\sigma^2_{Q_2}$ is a positive number by construction Q_2^{-1} being a positive definite matrix. Let now calculate the optimal portfolio for problem 3.4.1.

Lemma 3.4.2. *Given μ_P, $\sigma_P > \sigma^2_{Q_2}$, $\xi_P > 0$ and $k_z < 0$ the optimal portfolio for problem 3.4.1 is $\tilde{x} = (1 - x^{*'}\boldsymbol{1}, x^{*'})$ with*

$$x^* = D^{-1}[(\mu - \mu_0\boldsymbol{1}) \quad b] Q_2^{-1} \begin{bmatrix} \mu_P - \mu_0 \\ \frac{\xi_P}{\xi_y} \end{bmatrix} +$$

$$+\sqrt{\frac{\sigma_P^2 - \sigma_{Q_2}^2}{(s - H_1)}} \left[D^{-1}t - D^{-1}[(\mu - \mu_0\boldsymbol{1}) \quad b] Q_2^{-1} \begin{pmatrix} f_1 \\ g_1 \end{pmatrix} \right] \tag{3.4.5}$$

and the optimal kurtosis k^* is

$$k^* = k_z \left[(f_1 \quad g_1) Q_2^{-1} \begin{pmatrix} \mu_P - \mu_0 \\ \frac{\xi_P}{\xi_y} \end{pmatrix} + \sqrt{\sigma_P^2 - \sigma_{Q_2}^2} \sqrt{s - H_1} \right] \tag{3.4.6}$$

Proof: The Lagrangian function for problem 3.4.1 is:

$$L(x,\lambda) = -x'tk_z - \lambda_1(x'Dx - \sigma_P^2) - \lambda_2[x'(\mu - \mu_0\mathbf{1}) - (\mu_P - \mu_0)] - \lambda_3(x'b\xi_y - \xi_P)$$

The first order conditions for problem 3.4.1 are:

$$\frac{\partial L}{\partial x} = -tk_z - 2\lambda_1 Dx - \lambda_2(\mu - \mu_0\mathbf{1}) - \lambda_3 b\xi_y = \mathbf{0} \tag{3.4.7}$$

$$\frac{\partial L}{\partial \lambda_1} = x'Dx - \sigma_P^2 = 0 \tag{3.4.8}$$

$$\frac{\partial L}{\partial \lambda_2} = x'(\mu - \mu_0\mathbf{1}) - (\mu_p - \mu_0) = 0 \tag{3.4.9}$$

$$\frac{\partial L}{\partial \lambda_3} = x'b\xi_y = \xi_P \tag{3.4.10}$$

The Hessian matrix for problem 3.4.1 is $H_x(L) = \frac{\partial^2 L}{\partial x^2} = -2\lambda_1 D$. D being positive definite, then $-2\lambda_1 D$ is negative definite when $\lambda_1 > 0$. In the following, the sign of λ_1 will be investigated in order to ensure the first order conditions to be sufficient for problem 3.4.1.

From equation 3.4.7, assuming $\lambda_1 \neq 0$, the optimal portfolio can be written as

$$x^* = -\frac{k_z}{2\lambda_1} D^{-1} t - \frac{\lambda_2}{2\lambda_1} D^{-1}(\mu - \mu_0\mathbf{1}) + \frac{\lambda_3}{2\lambda_1} D^{-1}\xi_y b \tag{3.4.11}$$

The assumption $\lambda_1 \neq 0$ is equivalent to ask the quadratic restriction to hold. Let rename some parts of the equation 3.4.7 in order to simplify calculations. Let define the vector $\delta' = (\delta_1 \quad \delta_2 \quad \delta_3)$ where

$$\delta_1 = -\frac{\lambda_2}{2\lambda_1} \qquad \delta_2 = -\frac{\lambda_3}{2\lambda_1}\xi_y \qquad \delta_3 = -\frac{k_z}{2\lambda_1} \tag{3.4.12}$$

Let re-write equation 3.4.11 as a function of δ:

$$x^* = D^{-1}N\delta.$$

The 3 restrictions can now be re-written as functions of δ:

$$x'Dx = \delta' N' D^{-1} D D^{-1} N\delta = \delta' Q\delta = \sigma_P^2 \tag{3.4.13}$$

$$(\mu - \mu_0 \mathbf{1})' x = (\mu - \mu_0 \mathbf{1})' D^{-1} N\delta = \mu_P - \mu_0 \tag{3.4.14}$$

$$\xi_y b' x = \xi_y b' D^{-1} N\delta = \xi_P \quad \Rightarrow \quad b' D^{-1} N\delta = \frac{\xi_P}{\xi_y} \tag{3.4.15}$$

From the 2 linear constraints it's possible to obtain δ_1 and δ_2 as functions of δ_3:

$$\begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} = Q_2^{-1} \begin{pmatrix} \mu_P - \mu_0 \\ \frac{\xi_P}{\xi_y} \end{pmatrix} - Q_2^{-1} \begin{pmatrix} f_1 \\ g_1 \end{pmatrix} \delta_3.$$

By substitution in the quadratic restriction, it's possible to solve for δ_3:

$$\delta_3 = \pm \sqrt{\frac{\sigma_P^2 - \sigma_{Q_2}^2}{(s - H_1)}}.$$

From equation 3.4.12, λ_1 can be expressed as:

$$\lambda_1 = -\frac{k_z}{2\delta_3}. \tag{3.4.16}$$

The first order conditions for problem 3.4.1 are sufficient if $\lambda_1 > 0$. Therefore, for $k_z < 0$ ($k_z > 0$) the value of γ_4 is

$$\delta_3 = \sqrt{\frac{\sigma_P^2 - \sigma_{Q_2}^2}{(s - H_1)}} \qquad \left(\delta_3 = -\sqrt{\frac{\sigma_P^2 - \sigma_{Q_2}^2}{(s - H_1)}} \right)$$

Recalling that the kurtosis of the portfolio is $k = x' t k_z$, the optimal kurtosis associated to x^* can be easily obtained by substitution:

$$k^* = k_z \left[(f_1 \quad g_1) Q_2^{-1} \begin{pmatrix} \mu_P - \mu_0 \\ \frac{\xi_P}{\xi_y} \end{pmatrix} + \sqrt{\sigma_P^2 - \sigma_{Q_2}^2} \sqrt{s - H_1} \right] \quad \star . \tag{3.4.17}$$

The portfolio x^* in equation 3.4.11 is the optimal one in a E-V-S-K framework in presence of n risky assets and a risk-less one. This representation doesn't permit to split the effects of the higher moments on the optimal portfolio. It's useful then to re-write the optimal portfolio in a form that allows to make some considerations. Moreover, re-writing the optimal portfolio in a more suitable form, it permits to decompose it in different components that underline the role of skewness and kurtosis.

Let calculate the values of δ as a function m_1 instead of the matrix Q_2. To do this it is necessary to write δ_1 as functions of δ_2 and δ_3 using the first linear restriction in equation 3.4.14:

$$\delta_1 = \frac{\mu_P - \mu_0}{m_1} - \frac{l_1}{m_1}\delta_2 - \frac{f_1}{m_1}\delta_3 \tag{3.4.18}$$

where m_1 is the first entry of matrix Q.

Remark 3.4.1. *From Lemma 3.4.2, the following equation holds:*

$$\delta' Q \delta = \sigma_Q^2 > \sigma_{Q_2}^2 = (\delta_1, \delta_2)\, Q_2 \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} \tag{3.4.19}$$

A simple interpretation of Remark 3.4.1 is that an investor with the given structure of preferences for skewness and kurtosis is willing to choose portfolios with a higher variance compared to the mean-variance optimal ones. As in the case of n risky assets, the higher variance is counterbalanced by an increased value of portfolio skewness and a lower value of portfolio kurtosis.

Let evaluate the value of δ_2 plugging equation 3.4.18 into the quadratic constraint. The value of δ_2 can be written as follows:

$$\delta_2^2 \left(e - \frac{l_1^2}{m_1}\right) = \sigma_{Q_2}^2 + 2\delta_3(\mu_P - \mu_0)\frac{f_1}{m_1} \tag{3.4.20}$$

$$-\frac{(\mu_P - \mu_0)^2}{m_1} - \frac{f_1^2}{m_1}\delta_3^2 \tag{3.4.21}$$

Let rewrite equation 3.4.21 in a more useful way:

$$\delta_2^2(e-k) = \sigma_Q^2 - \sigma_B^2 \tag{3.4.22}$$

where

$$k = \frac{l_1^2}{m_1} \tag{3.4.23}$$

$$\begin{aligned} \sigma_B^2 = 2\delta_3(\mu_P - \mu_0)\frac{f_1}{m_1} - \\ -\frac{(\mu_P-\mu_0)^2}{m_1} - \frac{f_1^2}{m_1}\delta_3^2 \end{aligned} \tag{3.4.24}$$

Note that the quantity σ_B^2 is non negative by construction; in fact

$$\frac{[(\mu_P-\mu_0)-f_1]^2}{m_1} \quad and \quad m_1 > 0.$$

It's possible now to express the optimal portfolio for problem 3.4.1 as a function of matrix m_1. This allows to directly compare E-V-S-K optimal portfolio in presence of a risk-less asset with the three moment optimal portfolio, see Gamba and Rossi (1998).

Corollary 3.4.1. *Given μ_P, $\sigma_Q^2 \geq \sigma_{Q_2}^2 \geq \sigma_B^2$, $\xi_y > 0$ and $k_z < 0$ then the optimal portfolio for problem 3.4.1 can equivalently be written as, $\tilde{x} = (1 - x^{*'}\mathit{1}, x^{*'})$:*

$$\begin{aligned} x^* = D^{-1}(\mu - \mu_0\boldsymbol{1})\frac{(\mu_P-\mu_0)}{m_1} + \\ + \left[D^{-1}b + D^{-1}(\mu-\mu_0\boldsymbol{1})\frac{l_1}{m_1}\right]\sqrt{\frac{\sigma_{Q_2}^2-\sigma_B^2}{e-k}} + \\ + \left[D^{-1}t + D^{-1}(\mu-\mu_0\boldsymbol{1})\frac{f_1}{m_1}\right]\sqrt{\frac{\sigma_Q^2-\sigma_{Q_2}^2}{s-H_1}} \end{aligned} \tag{3.4.25}$$

3.4.1 Optimal Portfolio Returns Analysis

This section is very similar to section 3.3.2 for the case of n risky assets. Let now analyze the properties of the optimal portfolio in equation 3.4.25. The optimal portfolio in equation 3.4.25 is equivalent to the one in equation 3.4.5 and it can be divided as follows:

$$x_1^* = D^{-1}(\mu - \mu_0\mathbf{1})\frac{(\mu_P - \mu_0)}{m_1} \tag{3.4.26}$$

$$x_2^* = \left[D^{-1}b + D^{-1}(\mu - \mu_0\mathbf{1})A^{-1}\frac{l_1}{m_1}\right]\sqrt{\frac{\sigma_{Q_2}^2 - \sigma_Q^2}{e - k}} \tag{3.4.27}$$

$$x_3^* = \left[D^{-1}t + D^{-1}(\mu - \mu_0\mathbf{1})\frac{f_1}{m_1}\right]\sqrt{\frac{\sigma_P^2 - \sigma_{Q_2}^2}{s - H_1}} \tag{3.4.28}$$

Portfolio x_1^* is the classical optimal portfolio given μ_P in the mean-variance framework when considering $n+1$ assets, see Huang and Litzemberger (1988). The portfolio x_1^* displays the properties $\mathbf{1}'x_1^* = 1$, $\mu' x_1^* = \mu_P$ and $x_1^{*'} V x_1^* = \sigma_Q^2$.

Note that, by construction, portfolios x_2^* has the property $x_2^*\mathbf{1} = 0$. According to Ingersoll (1987), portfolio x_2^* is said to be an arbitrage portfolio. Portfolio x_2^* is useful to add variance to the optimal mean-variance portfolio following the skewness individual preferences. The extra variance due to the skewness preference depends on the quantity $\sigma_{Q_2}^2 - \sigma_B^2$. Moreover, portfolio x_2^* is a null vector if and only if $\xi_y = 0$, i.e. $\sigma_{Q_2}^2 = \sigma_B^2$.

Portfolio x_3^* displays the property $x_3^{*'}\mathbf{1} = 0$. According to Ingersoll (1987), portfolio x_3^* is said to be an arbitrage portfolio. Portfolio x_3^* is useful to add variance to the optimal mean-variance portfolio following the kurtosis individual preferences. The extra variance due to the kurtosis preference depends on the quantity $\sigma_Q^2 - \sigma_{Q_2}^2$. Moreover, portfolio x_3^* is a null vector if and only if $k_z = 0$, i.e. $\sigma_P^2 = \sigma_{Q_2}^2$.

Let now analyze what happens to the optimal portfolio x^* for particular values of $\sigma^2_{Q_2}$. In this way it will be clear that the proposed model is a generalization of classical mean-variance model, see Markowitz (1952), and of the three moments asset allocation model proposed by Gamba and Rossi (1998).

- $\sigma^2_Q = \sigma^2_{Q_2} = \sigma^2_B$.

  In this case, being $\sigma^2_Q = \sigma^2_{Q_2} = \sigma^2_B$, no extra variance is taken into account. The optimal portfolio x^* in equation 3.4.25 becomes

  $$x^* = D^{-1}(\mu - \mu_0 \mathbf{1})\frac{(\mu_P - \mu_0)}{m_1}$$

  This portfolio is exactly the mean-variance optimal portfolio when the risk-less asset is taken into account, see Huang and Litzenberger (1988).

- $\sigma^2_Q > \sigma^2_{Q_2} > \sigma^2_B$.

  In this case, positive extra variance is taken into account. Portfolio x^*_2 and x^*_3 are both different from the null vector. The optimal portfolio x^* becomes

  $$\begin{aligned} x^* = D^{-1}(\mu - \mu_0 \mathbf{1})\frac{(\mu_P - \mu_0)}{m_1} &+ \\ + \left[D^{-1}b + D^{-1}(\mu - \mu_0 \mathbf{1})\frac{l_1}{m_1}\right]\sqrt{\frac{\sigma^2_{Q_2} - \sigma^2_B}{e - k}} &+ \\ + \left[D^{-1}t + D^{-1}(\mu - \mu_0 \mathbf{1})\frac{f_1}{m_1}\right]\sqrt{\frac{\sigma^2_Q - \sigma^2_{Q_2}}{s - H_1}} & \end{aligned}$$

  Note that, when $\sigma^2_{Q_2} \downarrow \sigma^2_B$ almost all the extra variance taken into account is used in order to increase the skewness of the optimal portfolio. On the contrary, when $\sigma^2_{Q_2} \uparrow \sigma^2_Q$ the extra variance is almost all used in order to decrease the kurtosis of the optimal portfolio. In other words, by varying $\sigma^2_{Q_2}$, it's possible to give different importance to skewness and kurtosis in the optimal portfolio and build the efficient frontier in E-V-S-K framework. Let also underline that portfolio $x_1^*+x_2^*$ is different from the optimal portfolio obtained by Gamba and Rossi (1998) considering the first three moments of returns distribution. This is reasonable because the rate of substitution between skewness and kurtosis preferences is not explicitly defined in this model. To do that, the definition of the individual utility function is needed.

- $\sigma^2_Q > \sigma^2_{Q_2} = \sigma^2_B$.

  This is the case where the extra-variance is used just to decrease the value of the optimal portfolio kurtosis. The optimal portfolio x^* becomes

$$\begin{aligned} x^* = {} & D^{-1}(\mu - \mu_0\mathbf{1})\frac{(\mu_P - \mu_0)}{m_1} + \\ & + \left[D^{-1}t + D^{-1}(\mu - \mu_0\mathbf{1})\frac{f_1}{m_1}\right]\sqrt{\frac{\sigma^2_Q - \sigma^2_{Q_2}}{s - H_1}} \end{aligned} \qquad (3.4.29)$$

  Let underline that this model permits to consider mean, variance and kurtosis by themselves without considering the moments in the increasing order.

- $\sigma^2_Q = \sigma^2_{Q_2} > \sigma^2_B$.

  This case is very interesting because it directly shows that the proposed model is a generalization not only of the mean-variance model but also of the three moments model proposed by Gamba and Rossi (1998). When $\sigma^2_Q = \sigma^2_{Q_2}$ the kurtosis is not taken into account and, at the same time, the quantity σ_Q^2 becomes equal to σ_m^2. In fact, let recall that

$$\delta_2^2\left(e - \frac{l_1^2}{m_1}\right) = \sigma_{Q_2}^2 + 2\delta_3(\mu_P - \mu_0)\frac{f_1}{m_1} \\ -\frac{(\mu_P - \mu_0)^2}{m_1} - \frac{f_1^2}{m_1}\delta_3^2$$

but

$$\delta_3 = \sqrt{\frac{\sigma_Q^2 - \sigma_{Q_2}^2}{(s - H_1)}} = 0.$$

and so

$$\delta_2^2(e - k) = \sigma_{Q_2}^2 - \sigma_m^2 \qquad \Rightarrow \qquad \delta_2 = \sqrt{\frac{\sigma_{Q_2}^2 - \sigma_m^2}{e - k}}$$

where $\sigma_m^2 = \frac{(\mu_P - \mu_0)^2}{m_1}$. The optimal portfolio x^* becomes

$$x^* = D^{-1}(\mu - \mu_0\mathbf{1})\frac{(\mu_P - \mu_0)}{m_1} + \\ + \left[D^{-1}b + D^{-1}(\mu - \mu_0\mathbf{1})\frac{l_1}{m_1}\right]\sqrt{\frac{\sigma_{Q_2}^2 - \sigma_m^2}{e - k}} \tag{3.4.30}$$

The optimal portfolio x^* in equation 3.4.30 coincides with the optimal portfolio in the three moment framework as obtained in Gamba and Rossi (1998), see equation 2.2.7.

Another interesting case is when the optimal portfolio in equation 3.4.1 can be be split exactly in the sum of the optimal portfolio in E-V-S framework plus the component due to the kurtosis. Let note that $\sigma_B^2 = \sigma_m^2$ not only for $\delta_3 = 0$. The equality $\sigma_B^2 = \sigma_m^2$ also holds when

$$2\delta_3(\mu_P - \mu_0)\frac{f_1}{m_1} - \frac{f_1^2}{m_1}\delta_3^2 = 0$$

Solving equation 3.4.31, the condition becomes

$$\delta_3\left[2(\mu_P - \mu_0)\frac{f_1}{m_1} - \delta_3\frac{f_1^2}{m_1}\right] = 0$$

and so

$$\delta_3 = \frac{2(\mu_P - \mu_0)\frac{f_1}{m_1}}{\frac{f_1^2}{m_1}}. \tag{3.4.31}$$

When the condition on δ_3 in equation 3.4.31 holds, the optimal portfolio x^* becomes

$$\begin{aligned} x^* = D^{-1}(\mu - \mu_0\mathbf{1})\frac{(\mu_P - \mu_0)}{m_1} + \\ + \left[D^{-1}b + D^{-1}(\mu - \mu_0\mathbf{1})\frac{l_1}{m_1}\right]\sqrt{\frac{\sigma_{Q_2}^2 - \sigma_m^2}{e-k}} + \\ + \left[D^{-1}t + D^{-1}(\mu - \mu_0\mathbf{1})\frac{f_1}{m_1}\right]\sqrt{\frac{\sigma_Q^2 - \sigma_{Q_2}^2}{s - H_1}} \end{aligned} \tag{3.4.32}$$

Portfolio x^* is now the sum of the optimal portfolio in E-V-S framework, see equation 2.2.7, plus portfolio x_3^* that translates the kurtosis correction, i.e. the variation of the optimal weights due to the introduction of the kurtosis.

3.4.2 Four Funds Decomposition

Also in the case of n risky assets and a risk-less one it's possible to decompose the optimal portfolio in presence of higher moments into three funds. The result is obtained rewriting the portfolio selection optimization problem in a more tractable way as in subsection 3.3.3. The three funds have the shape of three vectors that can be used to span the space of optimal portfolios, i.e. the E-V-S-K space.

Lemma 3.4.3. *When the returns follow equation 3.2.1 and $n+1$ assets are* *considered, the efficient set is spanned by the following 3 vectors:*

$$v_1 = \frac{V^{-1}(\mu - \mu_0 \boldsymbol{1})}{\boldsymbol{1}'V^{-1}(\mu - \mu_0 \boldsymbol{1})}$$
$$v_2 = \frac{V^{-1}b}{\boldsymbol{1}'V^{-1}b}$$
$$v_3 = \frac{V^{-1}t}{\boldsymbol{1}'V^{-1}t}$$

The proof of this result is analogous to the one for the case of n risky assets, for details see subsection 3.3.3 and Simaan (1993).

Chapter 4

Application

4.1 Data Set

Let now apply the model presented in the previous chapter to a real financial data. Let first introduce the data base. The basket of asset classes taken into account is composed of the following 9 financial indexes: ESX Eurostoxx 50 index, DAX German stock market index composed of 30 largest companies in term of book volume and market capitalization, FIB Italian stock market index composed of 30 largest Italian companies, $IBEX$ Spanish stock market index composed of 35 largest Spanish companies, FTI Dutch stock market index, Vix implied volatility index on $S\&P500$ index, $Vnasdaq$ implied volatility index on $Nasdaq$ index, $Vdaxx$ implied volatility index on DAX index, $Vstoxx$ implied volatility index on ESX index. The time series is composed of 240 weekly observations from September 2004 to April 2009.

In table 4.1 the first four moments of the return distribution for equity indexes are reported. The Jarque-Bera test is performed in order to investigate normality of assets returns. As shown in table 4.1, none of the equity indexes show a distribution compatible with the normal distribution. Note

Table 4.1: Equity Indexes

	ESX	DAX	FIB	$IBEX$	FTI
Mean	0.000220468	0.000466	-0.00111	0.000393	-0.00057
Standard Deviation	0.01291557	0.010152	0.013547	0.012523	0.013822
Skewness	-2.029736721	-0.49269	-1.79556	-1.58298	-2.45641
p-value (T statistics)	1	0.99	1	1	1
Kurtosis	13,97719396	4.027769	9.695811	8.705369	17.66857
p-value (T statistics)	0	0	0	0	0
J. Bera test	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-3}$

that Jarque-Bera test isn't able to show if the non-normality is due to asymmetry, fat tails or both of them. In order to verify the reason of non-normality a T test on skewness and kurtosis is performed. Table 4.1 shows that for equity indexes non-normality depends only on kurtosis. Even if skewness for equity indexes is not statistically significant, let underline that generally equity markets show negative skewness values.

In table 4.2 the first four moments of the return distribution for volatility indexes are reported. The Jarque-Bera test is performed in order to investigate normality of assets returns. Has shown in table 4.2 none of the volatility indexes show a distribution compatible with the normal distribution. In this case all the indexes show non-normal distribution both for asymmetry and kurtosis[1]. Let underline that all the volatility indexes show positive asymme-

[1]The values of kurtosis for some volatility indexes are less than three because excess kurtosis compared to the value 3 (the kurtosis for a normal random variable) is taken into account.

Table 4.2: Volatility Indexes

	$Vnasdaq$	Vix	$Vdaxx$	$Vstoxx$
Mean	-0.000655622	9.7E-05	0.000349	-0.00026
Standard Deviation	0.042622394	0.048601	0.043145	0.046794
Skewness	0.572282007	0.471078	0.640776	0.704584
p-value (T statistics)	$1.93(10^{-8})$	$3.544(10^{-10})$	$1.55(10^{-15})$	$6.88(10^{-12})$
Kurtosis	1.702582958	1.517476	2.278429	3.42027
p-value (T statistics)	$1.40(10^{-13})$	0	0	0
J. Bera test	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-3}$

try.

The interesting thing about volatility indexes is that they show a strong negative correlation with equity markets. This feature could be of big interest in asset allocation problems giving the possibility of diversifying portfolio risk. Let underline that usually volatility indexes are considered to represent implied volatility, where implied volatility directly recall the definition of volatility in Black and Scholes formula for option pricing. Despite this, the value of volatility indexes is calculated as an average of out of the money call and put options on the underlying, i.e. the equity index itself.

The opportunity of using indexes instead of usual financial assets is possible thanks to the existence of efficient future markets. Futures markets are usually more liquid than standard equity markets and therefore, the use of futures can be considered as an advantage. The only difference with the use of futures contracts is the interpretation of the portfolio weights. In fact, a short position on a certain asset class can be directly taken without recalling the usual interpretation of short selling.

4.2 Estimation of the Parameters

In order to calculate the optimal portfolio in E-V-S-K framework it's needed to estimate the model parameters. The expected returns μ and the covariance matrix D can be easily evaluated by using the historical returns.

The difference with M-V framework is in the evaluation of the skewness and kurtosis coefficients that are not present in the classical framework. Let recall the key assumption of the proposed approach:

$$r = \mu + \epsilon + by + tz. \tag{4.2.1}$$

Assuming that the asset returns follow the previous equation the estimation of coefficients b and t is needed. The skewness of returns depends only on the random variable y. Note that in this framework the skewness of an asset is assumed to be proportional to the skewness of the non-spherical variable y. The b_i parameter can be estimated as suggested by Simaan (1986) as

$$\widehat{b}_i = \frac{\left[\frac{1}{T}\sum_{t=1}^{T}(r_{it} - E(r_i))^3\right]^{\frac{1}{3}}}{E^{\frac{1}{3}}(y)^3}$$

where b_i is the i-th entry of vector b and T is the sample size. The order between the assets doesn't depend on the variable y. Therefore the distribution of y is totally arbitrary. Moreover, b being proportional to y, the value of the skewness of y can be chosen as a numeraire, for example $E^{\frac{1}{3}}(y)^3 = \xi_y = 1$.

For kurtosis parameters t no estimation has been proposed in the literature. As in the previous case, the kurtosis of the returns depends only from the random variable z because of the assumption made on the returns structure. Simply generalizing the proposal for the estimation of b_i, the following

Table 4.3: Estimated Skewness Parameters

	ESX	DAX	FIB	$IBEX$	FTI	$Vnasdaq$	Vix	$Vdaxx$	$Vstoxx$
$\widehat{b}$	0	0	0	0	0	0.382	0.314	0.427	0.470

Table 4.4: Estimated Kurtosis Parameters

	ESX	DAX	FIB	$IBEX$	FTI	$Vnasdaq$	Vix	$Vdaxx$	$Vstoxx$
$\widehat{t}$	2.795	0.805	1.939	1.741	3.534	0.341	0.303	0.456	0.684

estimation for parameter t_i is proposed:

$$\widehat{t_i} = \frac{\left[\frac{1}{T}\sum_{t=1}^{T}(r_{it} - E(r_i))^4\right]^{\frac{1}{4}}}{E^{\frac{1}{4}}(z)^4}$$

where t_i is the i-th entry of vector t and T is the sample size. Again, the order between the assets doesn't depend on the variable z and therefore the distribution of z is totally arbitrary. Moreover, t being proportional to z, the value of the kurtosis of z can be chosen as a numeraire, for example $E^{\frac{1}{4}}(z)^4 = k_z = 1$.

The complete knowledge of the distribution of the latent variables y and z is not needed in order to estimate the model's parameters. Despite of this, for simplicity it's possible to assume a distribution for the variables y and z. For example, it's possible to assume that y and z are drawn respectively from a $Sk-t$ distribution with the desired skewness ξ_y and a t distribution with the desired kurtosis k_z.

In tables 4.3 and 4.4 are reported the estimated values for b and t respectively. Let note that the skewness parameters for equity indexes are set equal to zero. This choice depends on the fact that equity indexes show a skewness non significantly different from 0, see the previous section.

4.3 The Four Moment Efficient Frontier

Let now draw some frontiers in the E-V-S-K framework. As shown in Lemma 3.3.1, the optimal kurtosis can be written as:

$$k^* = k_z \left[\begin{pmatrix} p & q & r \end{pmatrix} P_2^{-1} \begin{pmatrix} \mu_P \\ 1 \\ \frac{\xi_P}{\xi_y} \end{pmatrix} + \sqrt{s-H}\sqrt{\sigma_P^2 - \sigma_{P_2}^2} \right].$$

Note that k^* is a function of the portfolio mean μ_P, the portfolio variance σ_P^2 and the portfolio skewness ξ_P. In general, k^* is defined on $\mathbf{R}^3$. As consequence, in order to draw the efficient frontier is necessary to set one of the variables equal to a given value. In figures 4.1, 4.2 and 4.3 the E-V-S-K frontier is represented with given values of ξ_P respectively equal to 1.5, 1.8 and 2.

Note that in figures 4.1, 4.2 and 4.3 the kurtosis represented on Z ax takes negative values. This is a consequence of the assumptions made on the random variable z. The kurtosis is defined as excess kurtosis compared to the value 3. Therefore, the portfolios on the E-V-S-K frontier show kurtosis less than 3, i.e. a returns distribution with tails that are less fat than the normal distribution.

As in the case of the mean-variance frontier, the efficient frontier is composed only of a subset of the optimal portfolio set. In other words, some of the portfolios on the frontier are trivially dominated in the sense of the E-V-S-K dominance by other portfolios. Figure 4.4 shows the efficient part

Figure 4.1: **E-V-S-K Frontier:** $\xi_P = 1.5$

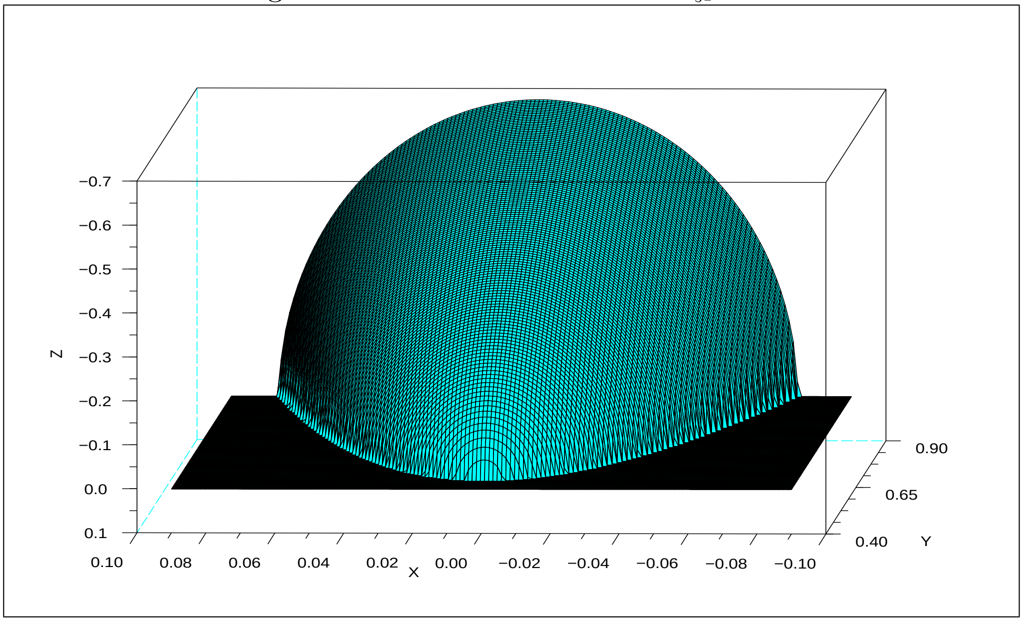
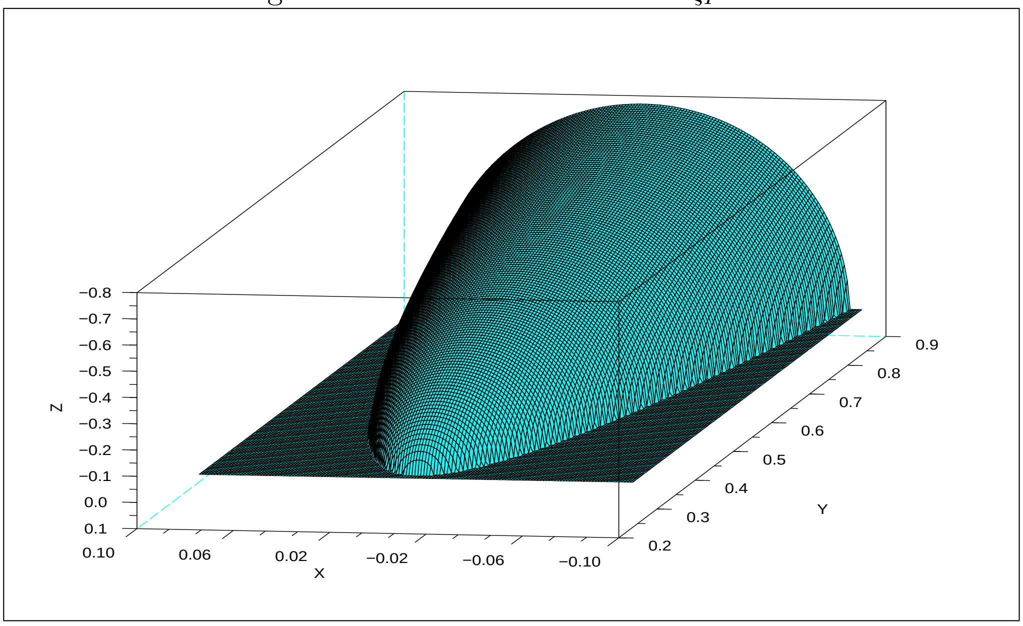


Figure 4.2: **E-V-S-K Frontier:** $\xi_P = 1.8$



of the E-V-S-K frontier for $\xi_P = 2$.

In figure 4.5 the mean-variance efficient frontier and the E-V-S-K efficient

Figure 4.3: **E-V-S-K Frontier:** $\xi_P = 2$

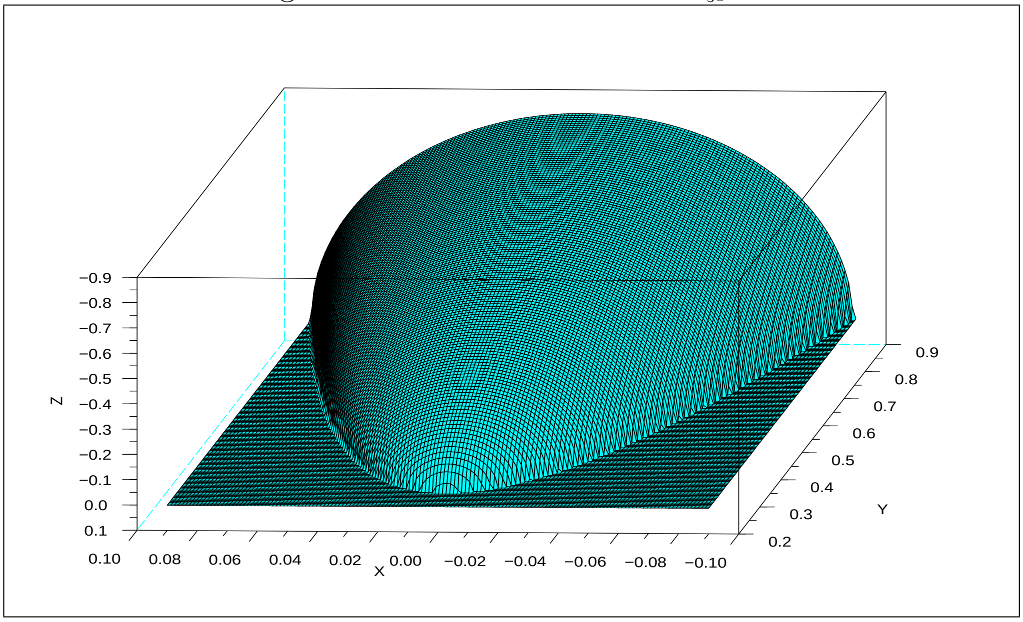
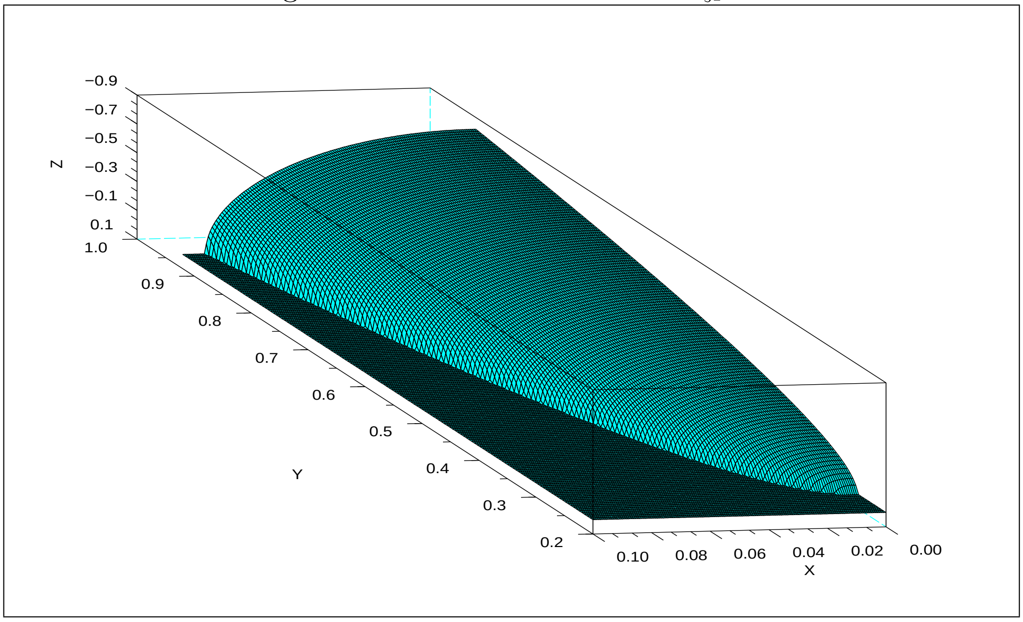
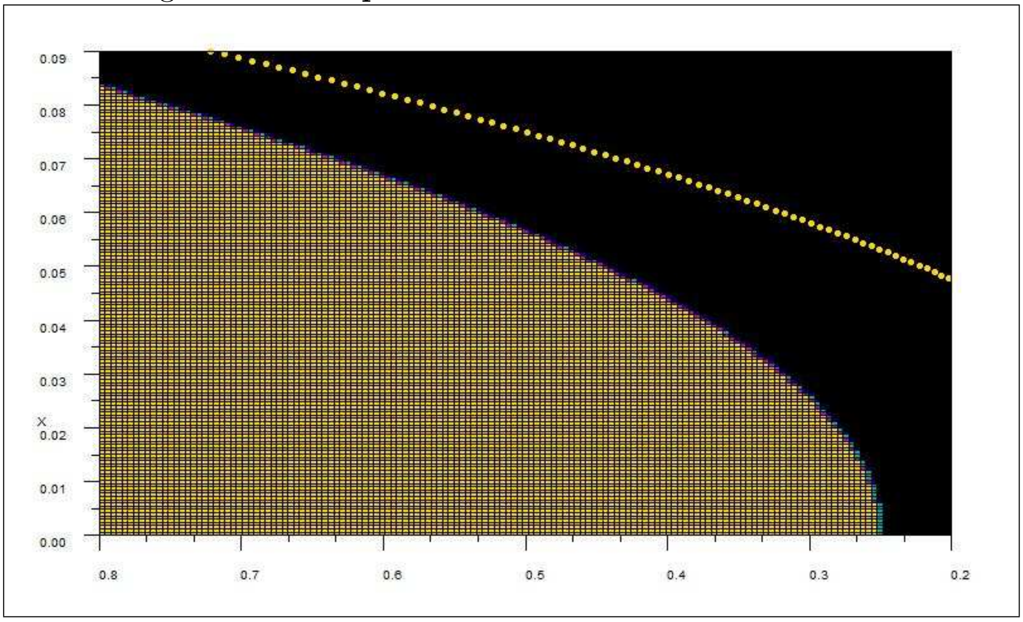


Figure 4.4: **Efficient Frontier:** $\xi_P = 2$



frontier are compared in the mean-variance space. The yellow line represents the mean-variance efficient frontier while the yellow area represents the projection of the E-V-S-K efficient frontier (with $\xi_P = 2$) on the mean-variance space. Note that, according to the proposed model, the variance of the E-V-S-K efficient portfolios is higher compared to the mean-variance efficient ones. As a consequence, E-V-S-K efficient portfolios are dominated in the mean-variance space by the portfolios belonging to the two moments efficient frontier.

Figure 4.5: **Comparison Between Efficient Frontiers**



4.4 Back Test

In this section the mean-variance optimal portfolio and some of the E-V-S-K optimal portfolios are compared through a back test. Let recall that x^*_{MV} (mean-variance) and x^*_{EVSK} are respectively equal to:

$$x^*_{MV} = D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix}\mu_p \\ 1\end{pmatrix}$$

$$x^*_{EVSK} = D^{-1}(\mu \quad \mathbf{1} \quad b)P_2^{-1}\beta + \sqrt{\frac{\sigma_P^2 - \sigma_{P_2}^2}{(s-H)}}\left[D^{-1}t - D^{-1}(\mu \quad \mathbf{1} \quad b)P_2^{-1}\begin{pmatrix}p \\ q \\ r\end{pmatrix}\right].$$

The back test is built in the following way: the data base is divided into two parts. The first part of the data set (training set) is used in order to estimate the model's parameters. The second part of the data base (validation set) is used in order to evaluate the performances of the portfolios. Portfolios x^*_{MV} and x^*_{EVSK} are chosen to have the same expected return. For x^*_{EVSK} some values for σ_p^2 such that $\sigma_p^2 > \sigma_{P_2}^2$ and for ξ_p are arbitrary chosen. Let underline that the reason of random choosing some portfolios on the E-V-S-K efficient frontier is not to define the agent utility function.

In figures 4.6, 4.7, 4.8 and 4.9 the returns of x^*_{MV} and x^*_{EVSK}, for different values of portfolio variance and skewness, are represented. In order to simplify the comparison between the portfolios, a Kernel smoothing (see Bowman and Azzalini (1997)) on returns distribution has been performed. For example, in figure 4.10 shows the comparison between the smoothed returns density functions of two of the selected portfolios. In figure 4.10 the blue and the red lines correspond to the distribution of returns respectively of $x^*_{EVSK_3}$ and x^*_{MV}. Note that, from figure 4.10, the returns distribution of portfolio x^*_{MV} presents tails that are fatter then the ones associated with

Figure 4.6: **Back-Test: Returns of x^*_{MV} portfolio.**

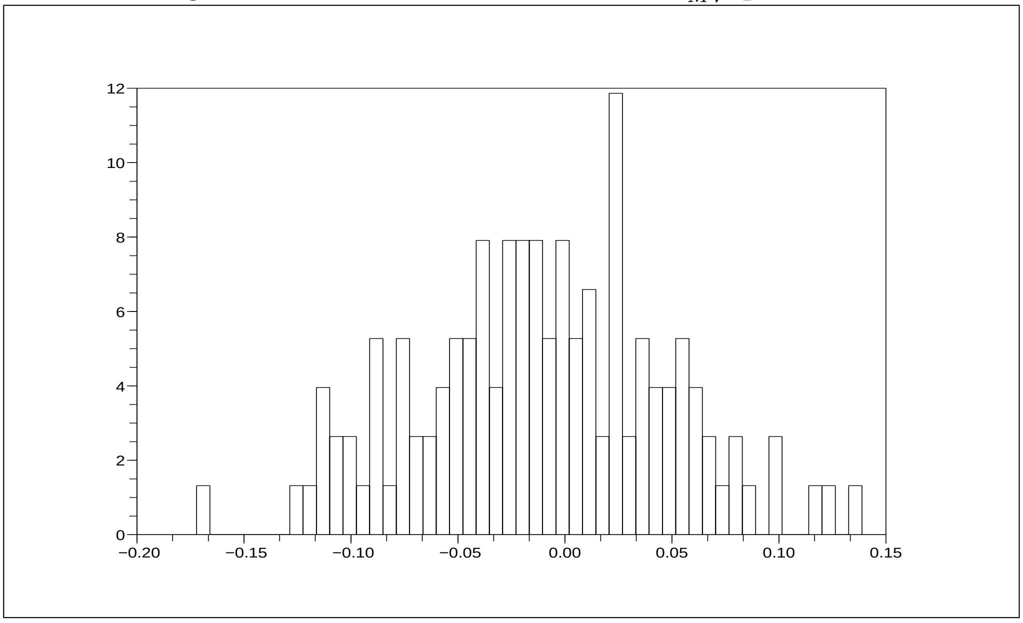
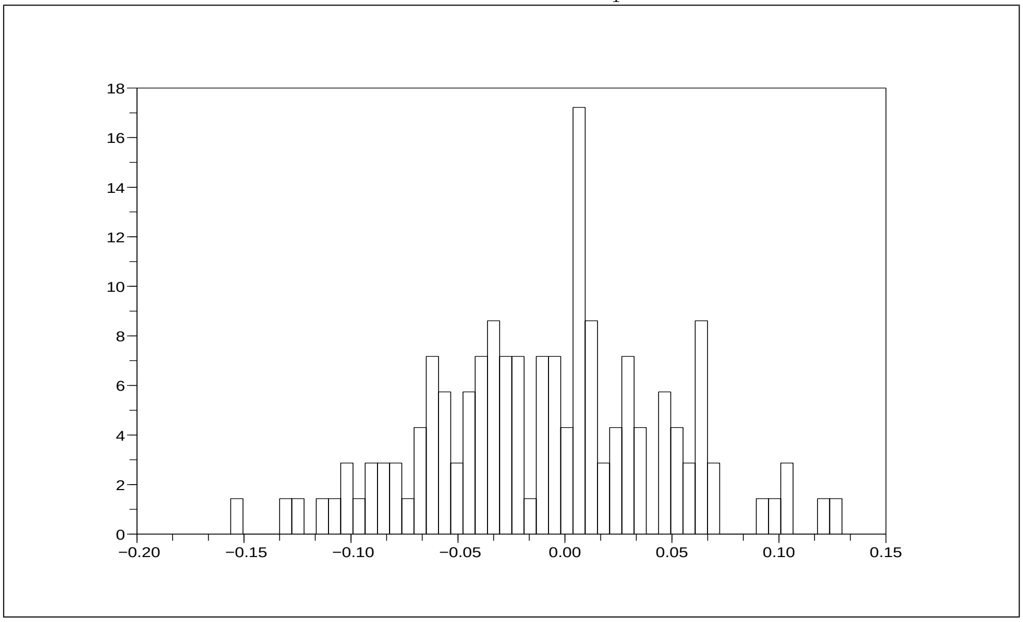


Figure 4.7: **Back-Test: Returns of $x^*_{EVSK_1}$ portfolio, $\xi_P = 2$, $\sigma^2_P = 0.5$.**



returns distribution of E-V-S-K optimal portfolio.

Let now analyze the moments of the returns distributions of the efficient

Figure 4.8: **Back-Test: Returns of $x^*_{EVSK_2}$ portfolio, $\xi_P = 1.5$, $\sigma^2_P = 0.6$.**

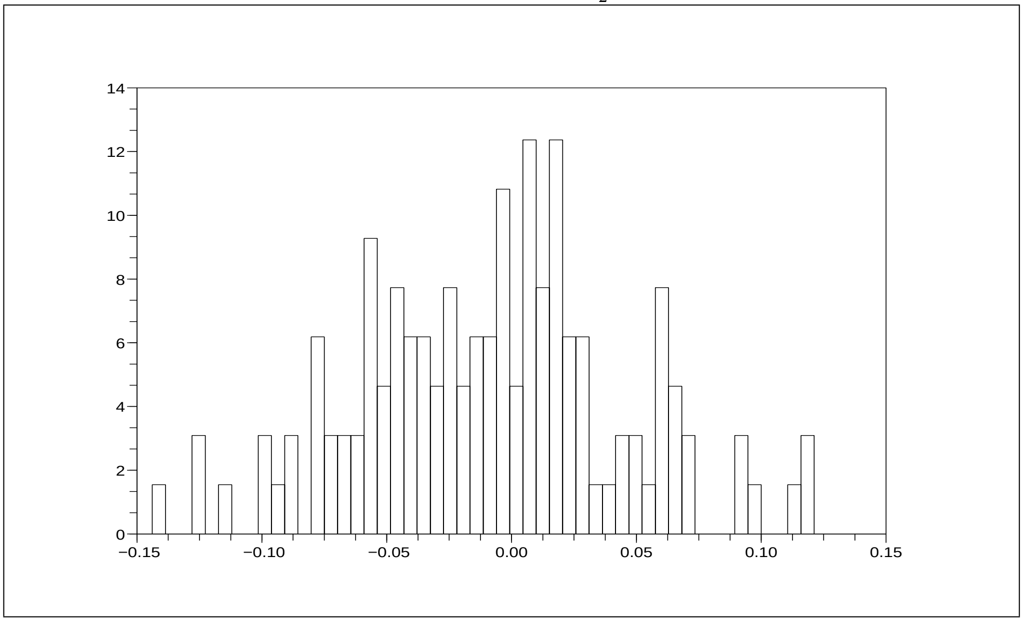
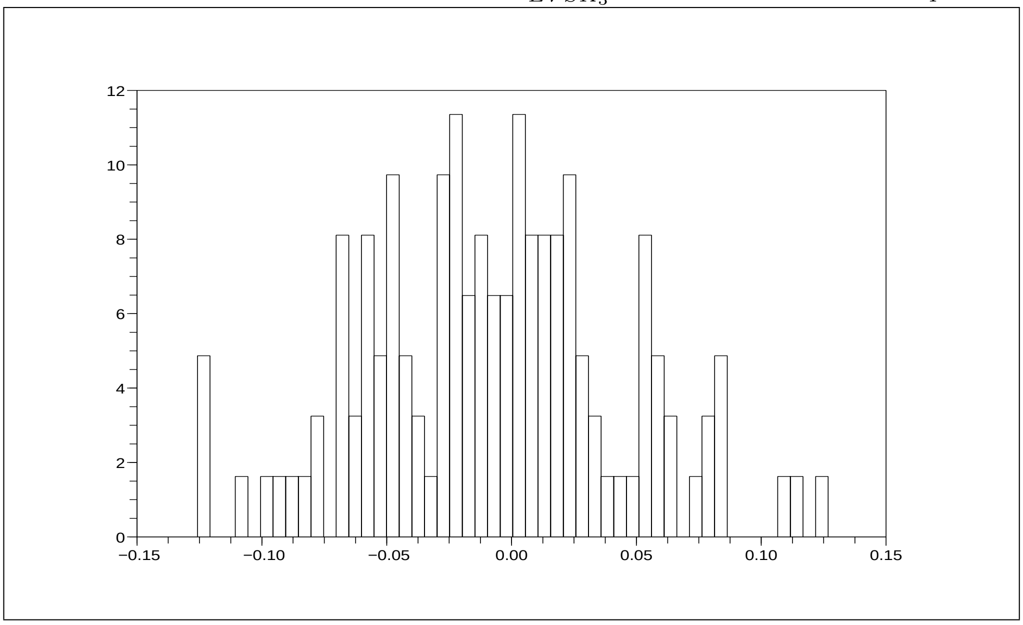


Figure 4.9: **Back-Test: Returns of $x^*_{EVSK_3}$ portfolio, $\xi_P = 1.2$, $\sigma^2_P = 0.7$.**



portfolios. Portfolio x^*_{MV} shows a mean and a variance respectively equal to -0.0105 and 0.0058. Note that, in this case, the portfolios selected on the

Figure 4.10: **Kernel Smoothing for Optimal Portfolios Returns.**

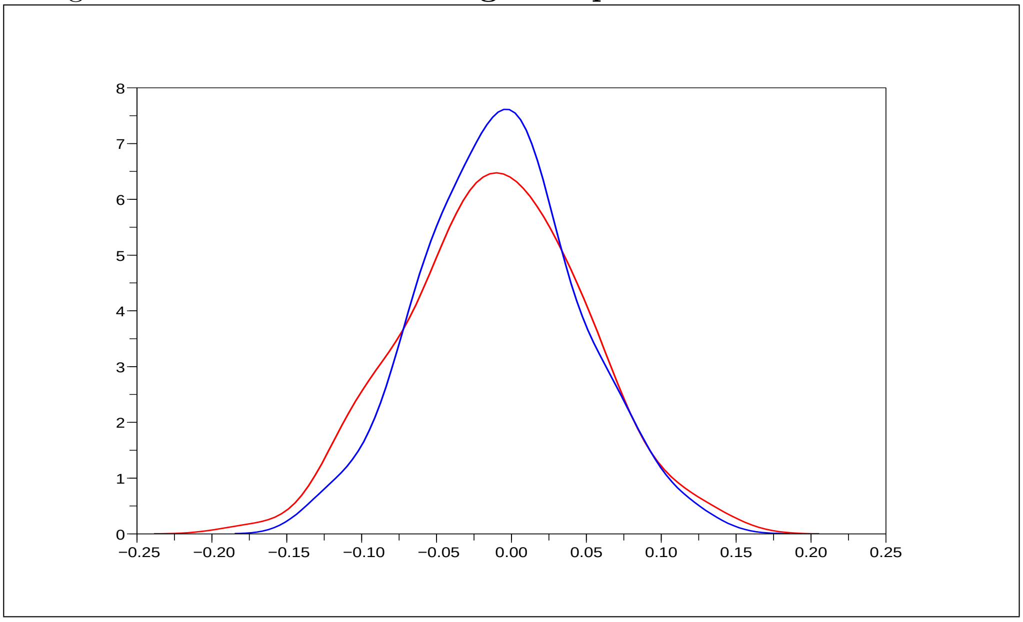


Table 4.5: Moments of the Optimal Portfolios

	$x^*_M V$	$x^*_{EVSK_1}$	$x^*_{EVSK_2}$	$x^*_{EVSK_3}$
Mean	-0.0105	-0.0093	-0.0084	-0.0071
Standard Deviation	0.058	0.055	0.053	0.051

four moments efficient frontier show a higher expected return and a lower standard deviation compared to the mean-variance portfolio, as shown in table 4.5. In other words, the mean-variance optimal portfolio is dominated in the sense of the mean-variance stochastic dominance by some portfolios chosen in the four moments framework. Of course, this is possible only in a out of sample context.

A first explanation of the result is that adding the third and the fourth moment to the classical mean-variance model permits to take into account more information with the consequence of a better asset allocation.

Let give a more intuitive explanation of the obtained result. As shown in the previous chapter, the optimal portfolio in E-V-S-K framework can be decomposed as follows:

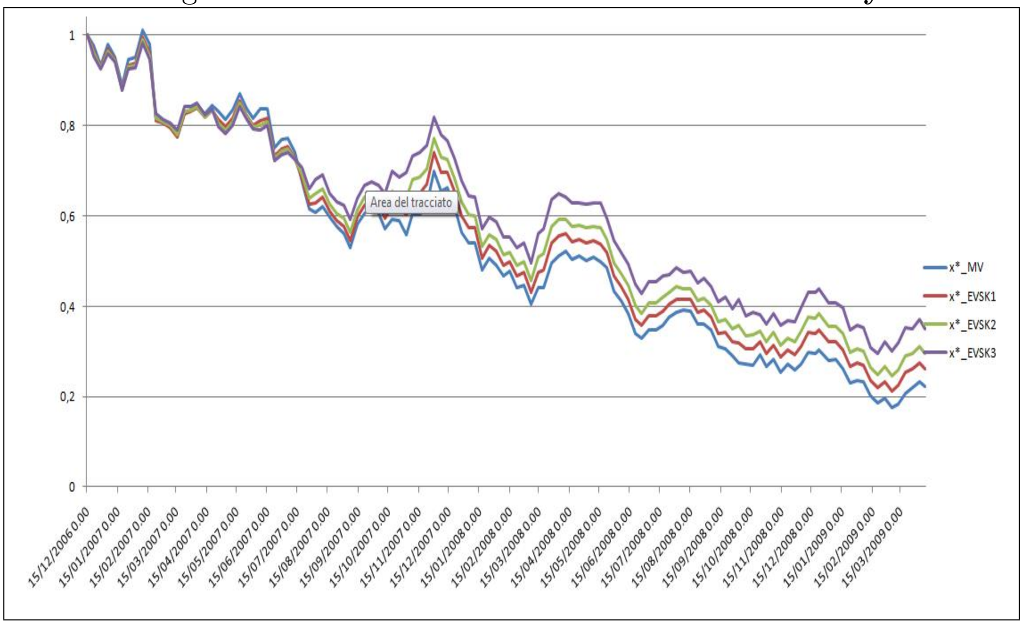
$$x_1^* = D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix} \mu_p \\ 1 \end{pmatrix}$$

$$x_2^* = \left[D^{-1}b + D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix} f \\ g \end{pmatrix} \right] \sqrt{\frac{\sigma_{P_2}^2 - \sigma_B^2}{e-h}}$$

$$x_3^* = \left[D^{-1}t + D^{-1}(\mu \quad \mathbf{1})A^{-1}\begin{pmatrix} p \\ q \end{pmatrix} \right] \sqrt{\frac{\sigma_P^2 - \sigma_{P_2}^2}{s-H}}.$$

Portfolio x_2^* is an arbitrage portfolio depending on skewness and it permits to change the allocation in portfolio x_1^* moving weight from the asset classes with lower values of skewness to the ones with higher values of skewness. Similarly, portfolio x_3^* permits to move weight to the asset classes with lower values of kurtosis.

Figure 4.11: **Back Test: Portfolios Value History.**



In general, it seems not be intuitive to find portfolios able to dominate the optimal mean-variance portfolio in the sense of the mean-variance dominance. On the other hand, this result give strong empirical evidence of the needs of considering higher order moments in asset allocation problems. Moreover, it's interesting to underline that the portfolios chosen in the four moments framework are sub-optimal in the mean-variance model. A rational economical is going to choose those portfolios only through the definition of the four moments efficient frontier.

In figure 4.11 the values of x^*_{MV}, $x^*_{EVSK_1}$, $x^*_{EVSK_2}$ and $x^*_{EVSK_3}$ are compared. Again, it is clear that the portfolios with higher moments result as a better asset allocation compared to the classical mean-variance one.

Conclusions

In this dissertation the consequences of non-normality of financial returns have been investigated. In chapter 3 a novel model for asset allocation that consider mean, variance, skewness and kurtosis has been proposed. This new model is based on the assumption that asset returns can be seen as the sum of a spherical random variable and two non-spherical variables that describe skewness and kurtosis. Together with the definition of the E-V-S-K stochastic dominance rule, the assumption on asset returns permit to write the optimization problem in a suitable way as the minimization of portfolio kurtosis, under budget, skewness, variance and expected return constraints.

The optimization problem has been solved analytically and a closed form solution for the optimal portfolio has been derived. The expression of the optimal kurtosis as a function of skewness, variance and expected return has been calculated. As first result, it has been shown that the optimal portfolio, solution of the proposed asset allocation model, is a generalization of classical asset allocation models. In fact, Markowitz and Gamba and Rossi optimal portfolios can be found as special cases for particular values of model's parameters, as shown in chapter 3.

Moreover, the returns of optimal portfolio have been analyzed. The optimal portfolio results as the sum of the mean-variance optimal portfolio plus two arbitrage portfolios that describe respectively investors preferences for skewness and kurtosis. In other words, the variance of the four moments optimal portfolio is higher than the variance of the mean-variance optimal one and the extra variance is counterbalanced by the investors preferences for higher order moments.

As last result, in chapter 3 a four funds separation theorem has been proved. This property is very useful and it permits to define four vectors that span the E-V-S-K space. Let underline that the separation property directly follows from the assumption on asset returns structure. The four spanning funds are directly comparable to the ones obtained in the classical asset allocation models.

In chapter 4, in order to test the model an empirical application on real financial data has been provided. In this section a technique for the estimation of skewness and kurtosis parameters has been proposed. A comparison between the mean-variance optimal portfolio and some E-V-S-K optimal portfolios is proposed. The results are really interesting. The back test shows that some four moments optimal portfolios beat in term of expected return and variance the two moments optimal one in an out of sample framework. Therefore it seems clear, also from the empirical results, that higher moments have to be taken into account for asset allocation problems.

No conditional moments are considered in the proposed framework. Despite of this, the proposed model is still valid also with the introduction of conditional moments. The choice not to consider conditional moments in this dissertation deals with the argument that the better is the estimation of the moments the better are the results of the asset allocation. As a consequence, with conditional moments the forecasting model becomes central. The introduction of conditional model in the proposed scheme will be the subject of further research.

The proposed model results as a natural extension of classical asset allocation models. The appeal of the model is to give a closed form solution to the optimization problem and obtain a functional form for the efficient frontier in the E-V-S-K framework. Through the decomposition of the optimal portfolio the role of skewness and kurtosis is clearly pointed out.

Bibliography

[1] Ang A., Bekaert G. (2002), International Asset Allocation with Regime Shifts, Review of Financial Studies 15(4), 1137-1187.

[2] Arditti F. D. (1967), Risk and the Required Return on Equity, Journal of Finance 22(1), 19-36.

[3] Athayde G. M., Flores Jr R. G. (2004), Finding a Maximum Skewness Portfolio - A General Solution to Three-moments Portfolio Choice, Journal of Economic Dynamics and Control 28, 1335-1352.

[4] Bauwens L., Laurent S. (2002), A New Class of Multivariate Skew Densities with Application to GARCH Models, CORE Discussion Paper 2002/20.

[5] Benishay H. (1992), The Pratt-Arrow Requirement in a Fourth Degree Polynomial Utility Function, Journal of Accounting, Auditing and Finance 7, 97-115.

[6] Berenyi Z. (2001), Performance of Leveraged Asset Funds, working paper, Munich School of Management.

[7] Bowman A. W., Azzalini A. (1997), Applied Smoothing Techniques for Data Analysis, Oxford University Press, New York.

[8] Brockett P. L., Garven J. R. (1998), A Reexamination of the Relationship Between Preferences and Moment Orderings by Rational Risk Averse Investors, Geneva Papers on Risk and Insurance Theory 23, 127-137.

[9] Brockett P. L., Kahane Y. (1992), Risk, Return, Skewness and Preference, Management Science 38, 854-860.

[10] Brown C., Robinson D. (2002), Skewness and Kurtosis Implied by Option Prices: a Correction, Journal of Financial Research 25(2), 279-282.

[11] Campbell H., Liechty J., Liechty W., Muller P. (2004), Portfolio Selection with Higher Moments.

[12] Campbell H., Siddique A. (1999), Autoregressive Conditional Skewness, Journal of Financial and Quantitative Analysis.

[13] Campbell J. Y., Viceira L. M. (1999), Consumption and Portfolio Decisions When Expected Returns Are Time Varying, Quarterly Journal of Economics, 114(2), 433-495.

[14] Chamberlain G. (1983), A Characterization of the Distributions that Imply Mean-variance Utility Functions, Journal of Economic Theory 29, 185-201.

[15] Chunhachinda P., Dandapani K., Hamid S., Prakash A. J. (1997), Portfolio Selection and Skewness: Evidence from International Stock Markets, Journal of Banking and Finance 21(2), 143-167.

[16] Das S. R., Uppal R. (2004), Systematic Risk and International Portfolio Choice, Journal of Finance 59, 2809-2834.

[17] Davies R., Kat H., Lu S. (2003), Fund of Hedge Funds Portfolio Selection: A Multiple Objective Approach.

[18] Dittmar R. F. (2002), Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross Section of Equity Returns, Journal of Finance 57, 369-403.

[19] Ederington L. (1986), Mean-variance as an Approximationto Expected Utility Maximation, Working Paper 86-5, Washington University.

[20] Engle R. F. (1982), Auto-regressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation, Econometrica 50, 987-1007.

[21] Engle R. F., Sheppard K. (2002), Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH, working paper.

[22] Fang H., Lai T. Y. (1997), Co-Kurtosis and Capital Asset Pricing, Financial Review 32(2), 293-307.

[23] Fishburn P. C. (1980), Stochastic Dominance and Moments of Distributions, Mathematics of Operations Research 5(1), 94-100.

[24] Gamba A., Rossi F. A., (1998), A Three-Moment Based Portfolio Selection Model, Rivista di Matematica per le Scienze Economiche e Sociali.

[25] Guidolin M., Timmermann A. (2003), Optimal Portfolio Choice Under Regime Switching, Skew and Kurtosis Preferences, Working Paper, University of California San Diego.

[26] Hanoch G., Levy H. (1969), The Efficiency Analysis of Choices Involving Risk, Review of Economic Studies 36, 335-346.

[27] Hansen B. E. (1994), Autoregressive Conditional Density Estimation, International Economic Review 35(3), 705-730.

[28] Harvey C. R., Siddique A. (1999), Autoregressive Conditional Skewness, Journal of Financial and Quantitative Analysis 34(4), 465-487.

[29] Harvey C. R., Siddique A. (2000), Conditional Skewness in Asset Pricing Tests, Journal of Finance 55(3), 1263-1295.

[30] Hlawitschka W. (1994), The empirical nature of Taylor series approximations to expected utility, American Economic Review 84, 713-719.

[31] Horvath P., Scott R. (1980), On the Direction of Preference for Moments of Higher Moments than the variance, Journal of Finance 35, 915-919.

[32] Huang C., Litzemberger R. H. (1988), Foundation for Financial Economics, North Holland, New York.

[33] Hwang S., Satchell S. E. (1999), Modelling Emerging Market Risk Premia Using Higher Moments, Journal of Finance and Economics 4(4), 271-296.

[34] Ingersoll J. E., (1987), Theory of Financial Decision Making, Rowman & Littlefield.

[35] Jondeau E., Rockinger M. (2009), The Impact of Shocks on Higher Moments, Journal of Financial Econometrics.

[36] Jondeau E., Rockinger M. (2006), Optimal Portfolio Allocation Under Higher Moments, European Financial Management.

[37] Jondeau E., Poon S. H., Rockinger M. (2007), Financial Modeling Under Non-Gaussian Distributions, Springer.

[38] Jondeau E., Rockinger M. (2003), Conditional Volatility Skewness and Kurtosis: Existence, Persistence and Comovements, Journal of Economic Dynamics and Control.

[39] Jondeau E., Rockinger M. (2003), How Higher Moments Affect the Allocation of Assets, Finance Letters.

[40] Jurczenko E., Maillet B. (2001), The 3-CAPM: Theoretical Foundations and an Asset Pricing Model Comparison in a Unified Framework, in C. Dunis, A. Timmermann, and J. Moody (eds.), Developments in Forecast Combination and Portfolio Choice, John Wiley & Sons.

[41] Jurczenko E., Maillet B. (2006), Multi-Moment Asset Allocation and Pricing Models, John Wiley & Sons, New York.

[42] Kimball M. S. (1993), Standard Risk Aversion, Econometrica 61, 589-611.

[43] Kraus A., Litzemberger R. H. (1976), Skewness Preference and the Valuation of Risky Assets, The Journal of Finance, 1985-1100.

[44] Kroll Y., Levy H., Markowitz H. M. (1984), Mean-variance Versus Direct Utility Maximization, Journal of Finance 36, 47-61.

[45] Lai T. Y. (1991), Portfolio Selection with Skewness: A Multiple-objective Approach, Review of Quantitative Finance and Accounting 1, 293-305.

[46] Levy H. (1969), A Utility Function Depending on the First Three Moments, Journal of Finance 24, 715-719.

[47] Levy H., Markowitz H. M. (1979), Approximating Expected Utility by a Function of Mean and Variance, American Economic Review 69, 308-317.

[48] Lhabitant F. (1997), On the (Ab)use of Expected Utility Approximations for Portfolio Selection and Portfolio Performance, Working Paper, University of Lausanne.

[49] Lintner J. (1965), The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolio and Capital Budgets, Review of Economics and Statistics 47, 13-37.

[50] Loistl O. (1967), The Erroneous Approximation of Expected Utility by Means of a Taylor's Series Expansion: Analytic and Computational Results, American Economic Review 66, 904-910.

[51] Markowitz H., (1952), Portfolio Selection, Journal of Finance.

[52] Markowitz H., (1959), Portfolio Selection: Efficient Diversification of Investments, Wiley, New York.

[53] Markowitz H., (1987), Mean-Variance Analysis in Portfolio Choice and Capital Markets, Blackwell.

[54] Mossin J. (1966), Equilibrium in a Capital Market, Econometrica 34, 768-83.

[55] Prakash A., Chang C., Pactwa T. (2003), Selecting a Portfolio with Skewness: Recent Evidence from US, European, and Latin American Equity Markets, Journal of Banking and Finance 27, 1375-1390.

[56] Pratt J. W., Zeckhauser R. J. (1987), Proper Risk Aversion, Econometrica 55, 143-154.

[57] Pulley L. B. (1981), A General Mean-variance Approximation to Expected Utility for Short Holding Period, Journal of Financial and Quantitative Analysis 16, 361-373.

[58] Rohan C. D., Chaudhry M. (2000), Coskewness and Cokurtosis in Future Markets, Journal of Financial Economics.

[59] Roll R. (1977), A Critique of the Asset Pricing Theory's Tests Part I: On Past and Potential Testability of the theory, Journal of Financial Economics.

[60] Samuelson P. (1970), The Fundamental Approximation Theorem of Portfolio Analysis in Terms of Means, Variances and Higher Moments, Review of Economic Studies 37, 537-542.

[61] Scott R. C., Horvath P. A. (1980), On the Direction of Preference for Moments of Higher Order than the Variance, The Journal of Finance 35, 915-919.

[62] Sharpe W. (1964), Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, Journal of Finance 19, 425-442.

[63] Simaan Y., (1993), Portfolio Selection and Asset Pricing: Three-Parameters Framework, Management Science, 39, 568-577.

[64] Simaan Y., (1986), Portfolio Selection and Capital Asset Pricing for a Class of Non-spherical Distributions , Dissertation, Baruch College, NY.

[65] Simaan Y., (1990), What is the Opportunity Cost of the Mean-Variance Investment Strategy?, Working Paper, Fordham University at Lincoln Center, NY.

[66] Tobin J., (1958),Liquidity Preference as Behavior Toward Risk, Review of Economical Studies 25, 65-86.

[67] Von Neumann J., Morgenstern O. (1947), Theory of Games and Economic Behavior, 2nd Edition, Princeton University Press, Princeton.

[68] Zieman V. (2004), Conditional Asset Pricing with Higher Moments, PhD Dissertation, University of Berlin.

[69] Ziemba W. T., (1974), Choosing Investment Portfolios When Returns Have Stable Distributions, Mathematical Programming in Theory and Practice, North Holland, Amsterdam.