

DESIRABLE PROPERTIES OF AN IDEAL RISK MEASURE IN PORTFOLIO THEORY

SVETLOZAR RACHEV

*University of California, Santa Barbara and
University of Karlsruhe, Germany*

SERGIO ORTOBELLI

University of Bergamo, Italy

STOYAN STOYANOV

FinAnalytica Inc., USA

FRANK J. FABOZZI*

*School of Management, Yale University,
135 Prospect Street, New Haven, CT 06520-8200, USA
frank.fabozzi@yale.edu*

ALMIRA BIGLOVA

University of Karlsruhe, Germany

Received 15 March 2006

Accepted 5 October 2007

This paper examines the properties that a risk measure should satisfy in order to characterize an investor's preferences. In particular, we propose some intuitive and realistic examples that describe several desirable features of an ideal risk measure. This analysis is the first step in understanding how to classify an investor's risk. Risk is an asymmetric, relative, heteroskedastic, multidimensional concept that has to take into account asymptotic behavior of returns, inter-temporal dependence, risk-time aggregation, and the impact of several economic phenomena that could influence an investor's preferences. In order to consider the financial impact of the several aspects of risk, we propose and analyze the relationship between distributional modeling and risk measures. Similar to the notion of ideal probability metric to a given approximation problem, we are in the search for an ideal risk measure or ideal performance ratio for a portfolio selection problem. We then emphasize the parallels between risk measures and probability metrics, underlying the computational advantage and disadvantage of different approaches.

Keywords: Risk aversion; portfolio choice; investment risk; reward measure; diversification.

*Corresponding author.

1. Introduction

In this paper, we describe the characteristics of risk. In doing so, we discuss and critically review some desirable properties of a risk measure in portfolio theory. We distinguish several observable financial phenomena such as the impact of aggregated risk, temporal horizon, propagation effect, risk aversion, transaction costs, and heteroskedasticity. In addition, we examine some properties that any risk measure has to take into account such as investment diversification, computational complexity, multi-parameter dependence, asymmetry, non-linearity, and incompleteness. Clearly, it is difficult to believe that a unique risk measure could capture all these characteristics and aspects of investor preferences. Thus, we propose some different ways to study the various aspects of risk.

Scientific methodology suggests first observing financial phenomena and then describing and characterizing it with respect to the tools and the information available. However, some studies on portfolio theory do not apply this approach to the investor problem. In fact, often proposals for risk measures for portfolio theory are just applications of measures found in the statistics literature. However, some of these proposed measures do not always take into account the range of investor attitudes towards risk. We believe that the main interest of investors is the consistency of a risk measure with their preferences.

Just defining the concept of “risk” and “uncertainty” is difficult. There is debate in the literature on the “right” definition of risk and uncertainty. Holton [16] proposes that a definition of risk has to take into account two essential components of observed phenomena: exposure and uncertainty. Moreover, all the admissible tools available to an investor to cope with risk can model only the risk that is perceived. Thus, in the finance literature, researchers can use only an operational definition of risk. That is, it is possible to operationally define only an investor’s perception of risk. This is in stark contrast with the definition of risk and of uncertainty proposed by Knight [18]: risk relates to objective probabilities and a probabilistic model can be given; uncertainty relates to subjective probabilities and no probabilistic model can be given.

In this paper we show with some simple examples many different aspects that could characterize the risk and the uncertainty of the portfolio choices made by investors. This is the first step to model the risk of investors. Even if we discuss and propose some methodologies to deal with problems related to risk, we cannot be exhaustive in providing a solution that takes into account all the aspects of risk. Once we define some desirable properties of an ideal risk measure to solve the portfolio choice problem, the second step is to appropriately model risk and uncertainty. While important, this is not a focus of our paper. Attempts at modeling have been proposed in the literature. For example, Ortobelli *et al.* [22] define and distinguish the uncertainty properties from the risk properties in order to describe the correct use of risk and uncertainty measures and their multi-dimensionality as it relates to the principal models known and employed in the portfolio literature. However, the models proposed by Ortobelli *et al.* [22] takes into account only

some aspects of perceived risk. Further analysis remains to solve all the related problems.

Attempts to quantify risk have led to the notion of a *risk measure*. A risk measure is a functional that assigns a numerical value to a random variable which is interpreted as a loss. Since risk is subjective because it is related to an investor's perception of exposure and uncertainty, risk measures are strongly related to utility functions. In particular, the link between expected utility theory and the risk of some admissible investments is generally represented by the consistency of the risk measure with a stochastic order, i.e., if X is preferred to Y by a given class of investors (non-satiable or non-satiable risk averse), then the risk of X is lower than the risk of Y from the perspective of that class of investors (see [25]). We shall not discuss here the details of this consistency. Nevertheless, it is important to realize that since risk measures associate a single number to a random variable, they cannot capture the entire information available in a stochastic order in which the cumulative distribution function of the loss is employed.

In portfolio theory, a risk measure has always been valued principally because of its capacity of ordering investor preferences. In particular, stochastic-order theory has provided some intuitive rules that are consistent with expected utility theory (see, among others, [3, 15, 31]). However, it is well recognized by expected utility specialists that von Neumann–Morgenstern utility functions cannot characterize all types of human behavior observed in financial markets. Several researchers have emphasized that the investors' choices are strictly dependent on the possible states of the returns (see, among others, [17]). Thus, investors have generally state-dependent utility functions. In order to take into account common attitudes and interests that characterize a decision maker's behavior, Karni [17], Schervish *et al.* [35], among others, have generalized the classical Von Neumann–Morgenstern approach to state-dependent utility functions. Moreover, it has been recently demonstrated that the state-dependent utility and the target-based approaches are equivalent (see [8, 9]). Therefore, when it is assumed that investors maximize their expected state-dependent utility functions, it is implicitly assumed that investors minimize the probability of the investment return falling below a specified risk benchmark. In particular, even if there are no apparent connections between the expected utility approach and a more appealing benchmark-based approach, expected utility can be reinterpreted in terms of the probability that the return is above a given benchmark (see, also, [10, 11]). These theoretical results justify many intuitive portfolio choice approaches based on the safety-first rules as a criterion for decision-making under uncertainty (see, among others, [3, 4, 24, 32, 43]).

It is well known that risk is an asymmetric concept related to downside outcomes, and any realistic way of measuring risk should consider upside and downside potential outcomes differently. Furthermore, a measure of uncertainty is not necessarily adequate in measuring risk. The standard deviation considers both positive and negative deviations from the mean as a potential risk. Thus, in this case, outperformance relative to the mean is penalized just as much as underperformance.

Balzer [5, 6] and Sortino and Satchell [38], among others, have proposed that investment risk might be measured by a functional of the difference between the investment return and a specified benchmark. In particular, the most celebrated and used benchmark approaches are based on coherent risk measures (see [41, 42]). As a matter of fact, the intuitive characteristics of investment risk, which are defined in a coherent risk measure, represent one of the most important aspects of the analysis by Artzner *et al.* [2]. However, even if a coherent risk measure “coherently” prices risk, it cannot consider exhaustively all investment characteristics. The benchmark might itself be a random variable, such as a liability benchmark (such as an insurance product or defined benefit pension fund liabilities), the inflation rate or possibly inflation plus some safety margin, the risk-free rate of return, the bottom percentile of return, a sector index return, a budgeted return or other alternative investments.

In practice, a benchmark is established by an investor and the risk benchmark is then communicated to the asset manager selected by the investor. The goal of the asset manager is not to underperform the benchmark. In contrast, minimizing the probability of being below a benchmark is equivalent to maximizing an expected state-dependent utility function (see [10, 11]). Thus, the benchmark approach is a generalization of the classic von Neumann–Morgenstern approach. In addition, the same investor could have multiple objectives and hence multiple benchmarks. Thus, risk is a *multidimensional* phenomenon. However, an appropriate choice of benchmarks is necessary in order to avoid incorrect evaluation of opportunities available to investors. For example, too often little recognition is given to liability targets. This is the major factor contributing to the underfunding of defined benefit pension plans in the United States (see [33]).

From this discussion, one must recognize that risk is a *relative* (to a given benchmark), *asymmetric*, and *multidimensional* concept. In particular, Rockafellar *et al.* [30] and Ortobelli *et al.* [22] have emphasized that risk cannot be assessed by measuring only the *uncertainty* of investments. In addition to asymmetry, relativity, and multidimensionality of risk, the discussion to follow justifies the following desirable features of investment risk: inter-temporal dependence, non-linearity; correlation, and diversification among different sources of risk. Moreover, risk measures have to take into account the impact of downside risk, aggregated risk, transaction costs, computational complexity, and risk aversion in investor’s choices. Clearly, we do not think that there exists a unique axiomatic definition of risk measure that summarizes all these phenomena related to risk. We only suggest in this paper the use of a scientific methodology to deal with these issues to detect the phenomena (typically using classic econometric tools, even if sometimes this could be done by a simple observation) and overcome singularly the related problems.

Section 2 summarizes some of the basic characteristics of risk emphasizing how some aspects of diversification motivate the use of reward-risk functionals. Section 3 describes the parallels between risk measures and probability metrics, while Secs. 4,

5 and 6 analyze several aspects of risk which impact the portfolio choices made by investors. We summarize our principal findings in Sec. 7.

2. Uncertainty and Risk: Temporal Dependence, Diversification and Reward-Risk Analysis

The most popular measure used as a proxy for risk is the standard deviation. However, as demonstrated in several papers, the standard deviation cannot always be utilized as a measure of risk because it is a measure of uncertainty. Nevertheless, the two notions of uncertainty and risk are related. Generally, we refer to a *generic risk measure* considering either a proper risk measure or a measure of uncertainty according to the definition in Ortobelli *et al.* [22]. Measures of uncertainty also known as *dispersion measures* can be introduced axiomatically (see [21]). We call *uncertainty measure* any increasing function of a positive functional D defined on the space of random variables satisfying the following properties:

Dev 1. $D(X + C) \leq D(X)$ for all X and constants $C \geq 0$.

Dev 2. $D(0) = 0$, and $D(aX) = aD(X)$ for all X and $a > 0$.

Dev 3. $D(X) \geq 0$ for all X , with $D(X) > 0$ for non-constant X .

According to these properties, positive additive shifts do not increase the uncertainty of the random variable X and the uncertainty measure D is equal to zero only if X is a constant. Therefore, we can say that the functional D measures the degree of uncertainty. One example of an uncertainty measure sensitive to additive shifts is the *colog*, defined as:

$$\text{colog}(X) = E(X \log X) - E(X)E(\log X).$$

This measure satisfies the above three properties and it is consistent with preferences of risk-averse investors, that is, if all risk-averse investors prefer the gross return X to Y , then $\text{colog}(X) \leq \text{colog}(Y)$ (see [14]). Particular uncertainty measures are the *deviation measures* (see [30]) that satisfy property Dev 1 as equality [i.e., $D(X + C) = D(X)$ for all X and constants C] properties Dev 2, Dev 3 and

Dev 4. $D(X + Y) \leq D(X) + D(Y)$ for all X and Y .

The deviation measure D depends only on the centered random variable $X - EX$ and is equal to zero only if $X - EX = 0$. Examples include the standard deviation and the mean absolute deviation. Further, the family of deviation measures does not include only symmetric representatives, i.e. the equality $D(X) = D(-X)$ is not guaranteed. The asymmetric representatives include, for instance, the semi-standard deviation. Moreover, next we show that the classical minimization problem that includes an uncertainty measure can be seen as a particular special case of the maximization of a reward-risk problem.

A systematic approach towards risk measures has been undertaken in Artzner *et al.* [2] where the family of *coherent* risk measures is introduced. A coherent risk

measure is any functional ρ defined on the space of random variables with finite variance satisfying the following properties:

- R1. $\rho(X + C) = \rho(X) - C$, for all X and constants C .
- R2. $\rho(0) = 0$, and $\rho(aX) = a\rho(X)$, for all X and all $a > 0$.
- R3. $\rho(X + Y) \leq \rho(X) + \rho(Y)$, for all X and Y .
- R4. $\rho(X) \leq \rho(Y)$ when $X \geq Y$.

Property R2 implies *positive homogeneity* of the functional. Property R3 implies *sub-additivity* and the combination of Properties R2 and R3 is *sub-linearity*, which implies *convexity*. If we relax the positive homogeneity assumption, we obtain the class of *convex risk measures*. That is, a risk measure is said to belong to the class of convex risk measures if it satisfies R1, R4, and the following convexity property:

- R5. $\rho(aX + (1 - a)Y) \leq a\rho(X) + (1 - a)\rho(Y)$, for all X, Y and $0 \leq a \leq 1$.

One example of a coherent risk measure is expected shortfall or expected tail loss (ETL) defined as

$$ETL_\alpha(X) = \frac{1}{\alpha} \int_0^\alpha VaR_q(X) dq, \quad (2.1)$$

where $VaR_\alpha(X) = -F_X^{-1}(\alpha) = -\inf\{x/P(X \leq x) \geq \alpha\}$ is the value-at-risk (VaR) of the random variable X . If we assume a continuous distribution for the distribution of X , then $ETL_\alpha(X) = -E(X/X \leq -VaR_\alpha(X))$, which is also known as conditional value-at-risk (CVaR). VaR itself is used as a risk measure and while it has an intuitive interpretation, examples can be constructed showing that it is not convex. ETL can be interpreted as the average loss beyond VaR. Both measures (VaR and ETL) consider the *downside risk* of portfolio returns (see, among others, [38]).

There is a relationship between deviation measures and coherent risk measures (see [30]). In particular, when the return distribution functions depend only on the mean and a risk measure, most of these measures are *equivalent* for risk-averse investors (see [22]). On the other hand, Biglova *et al.* [7] have shown (exploring the relationship between uncertainty measures and risk measures and how to employ them in order to obtain optimal choices) that one family cannot replace the other in portfolio selection problems.

2.1. Mean reversion, clustering of volatility, and cointegration

The previous axiomatic definitions of risk and uncertainty do not consider most of the observed phenomena that characterize typical financial series. Consider the following example. Figure 1 shows the MSCI World Index daily return series from January 4, 1993 to May 31, 2004. As we can see, the dispersion around the mean changes sensibly, in particular during the period after September 11, 2001 and when the oil/energy crises began. These oscillations suggest that the process is *mean reverting* and that the *dispersion* changes over the time. Hence, in some

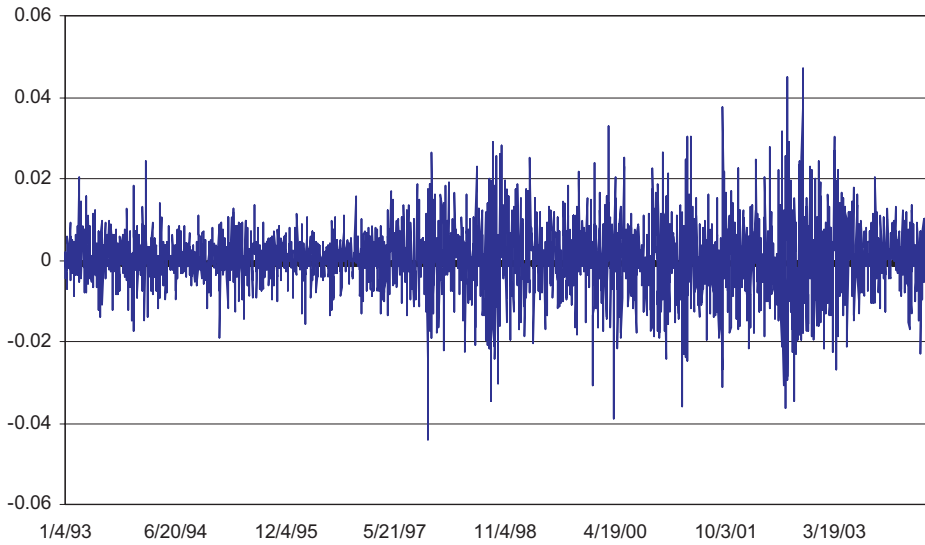


Fig. 1. MSCI World Index return time series from 1/4/1993 to 5/31/2004.

periods there are big oscillations around zero and in other periods the oscillations are small. Clearly, if the degree of uncertainty changes over time, the risk too has to change over time. In this case, the investment return process is not stationary; that is, we cannot assume that returns maintain their distribution unvaried in the course of time.

Under the assumption of stationary and independent realizations, the oldest observations have the same influence on our decisions as the most recent ones. Is this assumption realistic? Recent studies on investment return processes have shown that historical realizations are not independent and exhibit *autoregressive behavior*. Consequently, we observe the *clustering of volatility* effect; that is, each observation influences subsequent ones. In particular, the last observations have a greater impact on investment decisions than the oldest ones. Thus, any realistic measure of risk should change and evolve over time through a proper modeling of the behavior of financial variables.

One of the simplest models proposed in the literature for such modeling is the exponentially weighted moving average model (EWMA) and considers exponential weights (see [19]). Under the assumptions of the model, the risk measure follows a predictable process and at time t the observation r_k ($k < t$) is a possible outcome with probability $(1 - \lambda)\lambda^{t-k}$, where $\lambda \in [0, 1]$ is a decay factor that can be estimated with a root mean square error (RMSE) method (see [19]). Thus if the forecasted risk measure of return r_{t+1} is given by $\sigma_{t+1/t} = E_t(f(r_{t+1}))$ for an opportune function f , then the *conditional heteroskedasticity* of return series can be modeled assuming that $\sigma_{t+1/t} = \lambda\sigma_{t/t-1} + (1 - \lambda)f(r_t)$.

In contrast, the risk of a country/sector is linked to the risk of the other countries/sectors. For example, let us consider the United States and German daily return series as part of the MSCI World Index from January 4, 2000 to May 31, 2004 (see Figs. 2 and 3). These series present a correlation of 56.16% and we observe

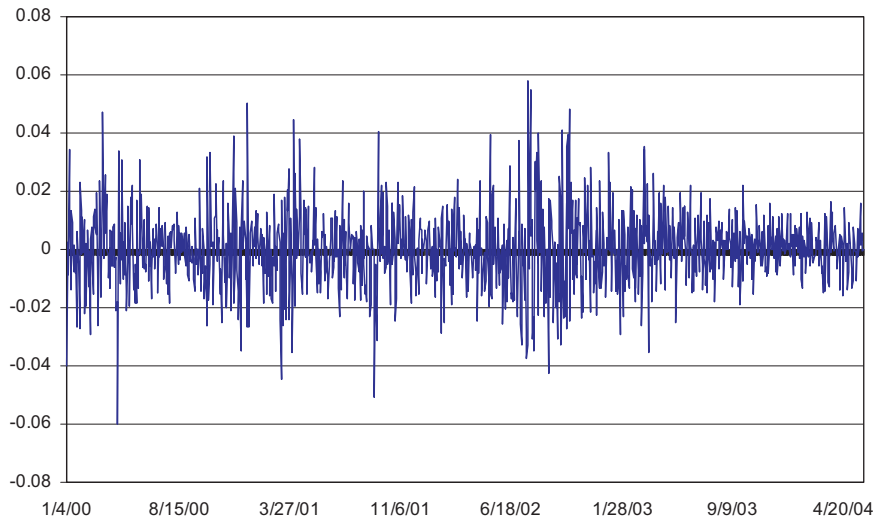


Fig. 2. US return time series (as part of MSCI World Index) from 1/4/2000 to 5/31/2004.

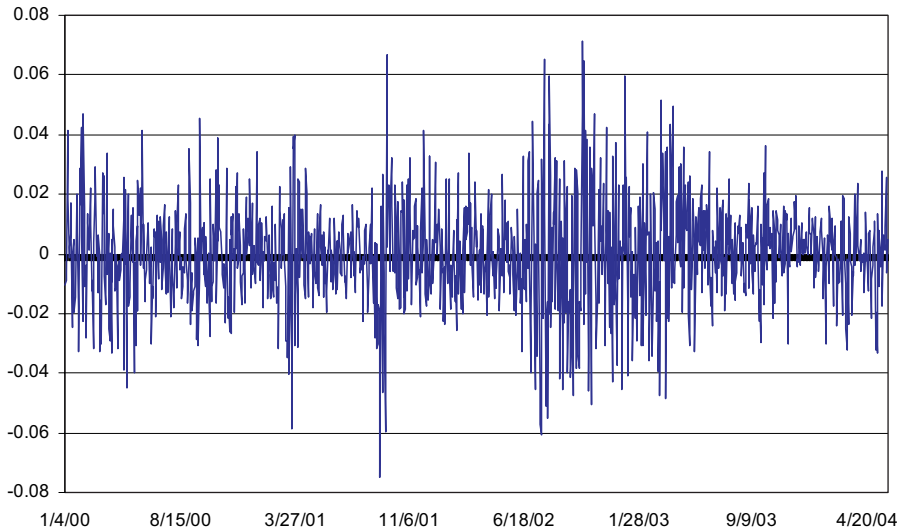


Fig. 3. German return time series (as part of MSCI World Index) from 1/4/2000 to 5/31/2004.

a greater dispersion of the German series than the US series (as suggested by their standard deviation: the US series standard deviation is 0.000174 and the German series standard deviation is 0.000323). As a matter of fact, for any peak in the US returns, there is an analogous higher peak in the German series. This propagation effect is known as *cointegration of returns series* and is a consequence of the globalization of markets. Clearly, Figs. 2 and 3 cannot provide a meaningful measure of cointegration of global capital markets and we need to test the financial data with standard econometric procedures to statistically assess the presence or absence of phenomena such as mean reversion, autoregressive behavior, and cointegration. Typically the behavior of simultaneous financial series is well captured by multivariate ARMA-GARCH type models (see, among others, [26]).

2.2. Risk diversification

From the above discussion we deduce that it could be important to limit the propagation effect by *diversifying risk*. As a matter of fact, there is considerable evidence that diversification diminishes the probability of big losses. Hence, an adequate risk measure values and accounts for the *dependence* among different investments, sectors, and markets. In particular, recall that in order to consider the diversification effect, it is required that a risk measure be a convex functional.

The convexity property only guarantees that diversification could take place once we construct a portfolio. In the optimal portfolio selection problem, this property alone is not sufficient to find a solution — we need an assumption about the multivariate distribution of portfolio items returns. It is through the multivariate modeling that we are able to describe the dependence between the gross returns of the portfolio items. As a matter of fact, investors want to diversify the portfolio in order to minimize the risk. Therefore, diversification makes sense only when there exist some values $a \in (0, 1)$, such that $\sigma_{aX+(1-a)Y} < \min(\sigma_X; \sigma_Y)$, i.e., when the risk of a portfolio is lower than the risk of the single investments. Thus, when

$$\exists a \in (0, 1) : \sigma_{aX+(1-a)Y} \leq \min(\sigma_X; \sigma_Y; a\sigma_X + (1-a)\sigma_Y), \quad (2.2)$$

we say that *strong diversification* holds and it is *convenient* to diversify the portfolio considering both X and Y . However, most of portfolio theory has been developed considering both the mean and the risk of the portfolios. Thus, for most investors diversification will appear convenient in a mean-risk plane if there exist some values $a \in (0, 1)$, such that for a given positive convex measure σ

$$\frac{E(aX + (1-a)Y)}{\sigma_{aX+(1-a)Y}} > \max\left(\frac{E(X)}{\sigma_X}; \frac{E(Y)}{\sigma_Y}\right).$$

In this case, we say that *weak diversification* is *convenient*. Weak diversification is generally guaranteed from the existence of some values $a \in (0, 1)$ that maximize

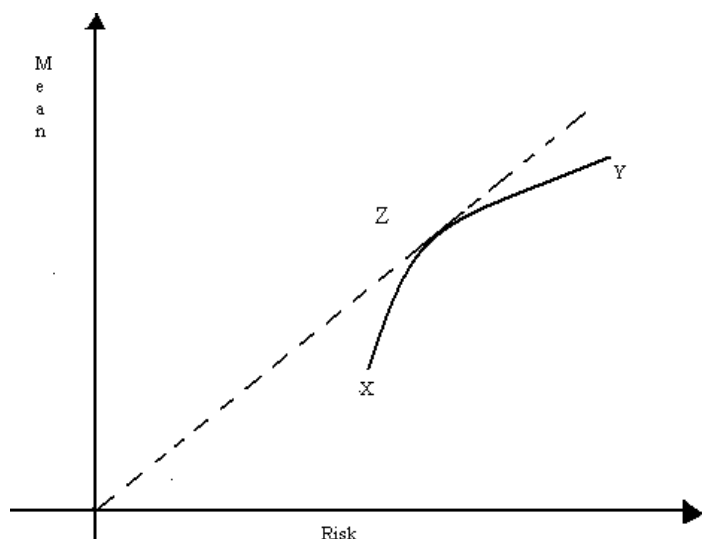


Fig. 4. Graphical example where weak diversification holds but strong diversification does not hold.

the ratio between the mean and the risk measure. In this case, there exists $a \in (0, 1)$ such that

$$\frac{\partial \sigma_{aX+(1-a)Y}}{\partial a} E(aX + (1-a)Y) = \sigma_{aX+(1-a)Y} E(X - Y). \quad (2.3)$$

Figure 4 provides an example where weak diversification holds but strong diversification does not. In this figure the curve $X - Y$ represents the mean-risk representation of the portfolios $aX + (1-a)Y$ convex combinations of funds X and Y . Thus, strong diversification does not hold since $\sigma_X \leq \sigma_{aX+(1-a)Y}$ for any a belonging to the interval $(0, 1)$. However, there exists a portfolio $Z = bX + (1-b)Y$ with $b \in (0, 1)$ that presents the highest mean/risk ratio (i.e., $\frac{E(Z)}{\sigma_Z} > \max\left(\frac{E(X)}{\sigma_X}, \frac{E(Y)}{\sigma_Y}\right)$). Therefore in this case weak diversification is convenient. Observe that strong diversification implies weak diversification when the financial random variables are the gross returns and we assume no short sales plus the limited liability hypothesis (i.e., the final wealth is a positive random variable). These definitions of diversification serve only to identify when it makes sense to diversify the portfolio. When we assume only the convexity property, we do not know if it makes sense to diversify a portfolio between two investments and we cannot say anything about the optimal portfolio. For example, in some cases where we have a portfolio of two linearly dependent returns X and $Y = bX + c$, we do not need to diversify the portfolio because one return is redundant and in a frictionless market it should be replicated by the other one. Moreover, any deviation measure (such as the standard deviation) does not present weak diversification when two gross returns X and Y are strongly positive correlated, even if the convexity does not tell us anything about the opportunity of diversifying the portfolio. Another simple example where convexity holds but weak

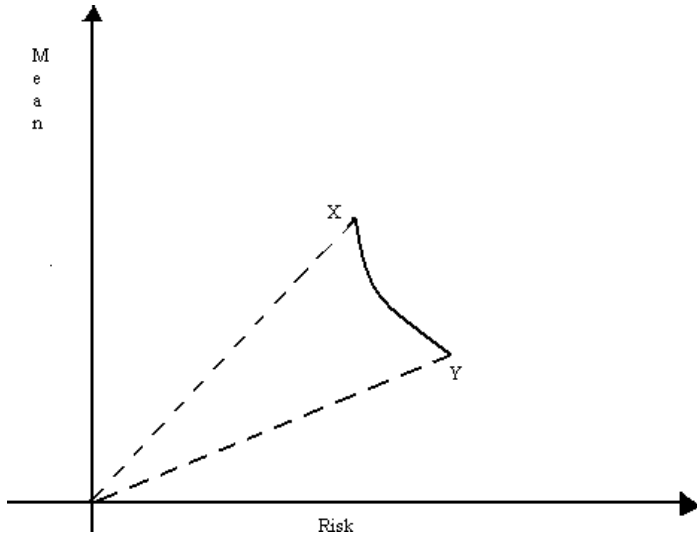


Fig. 5. Graphical example where convexity holds but the weak diversification does not hold.

diversification does not is shown in Fig. 5. As a matter of fact, in this example portfolio X presents the highest mean/risk ratio and then weak diversification is not convenient, even if $\sigma_{aX+(1-a)Y} \leq a\sigma_X + (1-a)\sigma_Y$ for any a belonging to $(0, 1)$.

2.3. Reward measures

These different definitions of diversification underline the importance of taking into account an investor's reward and not only risk in the portfolio selection problem. In particular, De Giorgi [12] introduced the first axiomatic definition of reward measures identifying the axioms that characterize uniquely the mean as a reward measure. In contrast to the highly restrictive definition proffered by De Giorgi, we assume a *reward measure* to be any functional v defined on the space of random variables of interest satisfying the following intuitive property:

Isotonicity with the market preferences: the functional v is isotonic with respect to the order of preference of the market, i.e., if any investor in the market prefers X to Y , then $v(X) \geq v(Y)$. In particular, when all the investors are non-satiable and risk-averse, we could consider functionals isotonic with the second-stochastic order.

Considering that in portfolio theory we need only order the choices for the investors' attitude towards risk, we do not need further axioms to express a choice. However, sometimes it could be important to underscore the coherency properties applied to reward measures. A *coherent reward measure* is any functional v defined on the space of random variables satisfying the following properties:

- M1. $v(X + C) = v(X) + C$, for all X and constants C .
- M2. $v(0)=0$, and $v(aX) = av(X)$, for all X and all $a > 0$.

M3. $v(X + Y) \geq v(X) + v(Y)$, for all X and Y .

M4. $v(X) \geq v(Y)$ when $X \geq Y$.

The combination of Properties M2 and M3 implies *concavity*. Similar to convex risk measures, a reward measure is said to belong to the class of *concave reward measures* if it satisfies M1, M4, and the following concavity property:

M5. $v(aX + (1 - a)Y) \geq av(X) + (1 - a)v(Y)$, for all X, Y and $0 \leq a \leq 1$.

From this definition we deduce that all the functionals opposite of coherent (convex) risk measure are coherent (concave) reward measures and vice versa. In practice, a coherent reward measure recognizes that: (1) wealth increases when we add an amount of riskless wealth; (2) when wealth is multiplied by a positive factor, then the reward must grow also with the same proportionality; (3) the aggregated reward of two investments is higher than the sum of the two associated single rewards; and (4) more wealth is preferred to less wealth.

The definition of a coherent risk (reward) measure takes into account that investors are generally non-satiable and risk-averse. However, in some circumstances it could be important to identify the most aggressive investments among several possible. In this case the investor should be non-satiable and a risk lover. The reward (risk) measure that considers the preferences of non-satiable and risk lover investors should satisfy the axioms M1, M2, M4 (R1, R2, R4), and the sub-additivity (super-additivity), i.e. $v(X + Y) \leq v(X) + v(Y)$ ($\rho(X + Y) \geq \rho(X) + \rho(Y)$). We will call these measures *aggressive reward (risk) measures*. A typical example of an aggressive reward measure is

$$v(X) = ETL_{\alpha}(-X).$$

When we assume a continuous distribution for X , then $ETL_{\alpha}(-X) = E(X/X \geq -VaR_{1-\alpha}(X))$. Therefore, the level of aggressiveness depends on α , the smaller α is, the more aggressive the investor is. When α is 1, an investor is maximizing the mean. As a matter of fact, this measure is isotonic with convex order (see [36]). Thus, if any risk-lover investor (with convex utility function) prefers X to Y , then $ETL_{\alpha}(-X) \geq ETL_{\alpha}(-Y)$ for any $\alpha \in (0, 1)$.

A reward measure v is aggressive if and only if it can be seen as a coherent risk measure ρ applied at the random variable multiplied by -1 (i.e., $v(X) = \rho(-X)$ for any random variable X). Similarly, a risk measure ρ is aggressive if and only if it can be seen as a coherent reward measure v applied at the random variable multiplied by -1 (i.e., $\rho(X) = v(-X)$ for any random variable X). On the other hand, De Giorgi [12] proved that the unique reward measure that satisfies axioms M1, M2, M4, and it is linear, i.e., $v(X + Y) = v(X) + v(Y)$, is the expected value, that is $v(X) = E(X)$.

2.4. Reward-risk ratios

The solution of the optimal portfolio problem is a portfolio that minimizes a given risk measure provided that the expected reward is constrained by some minimal

value R :

$$\begin{aligned} & \min_w \rho(w^T r - r_b) \\ & \text{s.t.} \\ & v(w^T r - r_b) \geq R \\ & l \leq Aw \leq u \end{aligned} \quad (2.4)$$

The set of all solutions, when varying the value of the constraint, is called the efficient frontier. Along the efficient frontier there is a portfolio that provides the maximum expected reward per unit of risk. In order to find this optimal portfolio, we have to minimize the ratio between the reward and the risk if the reward and risk measures are both negative for all portfolios and we have to maximize the ratio when both measures are positive. That is, this optimal portfolio is a stationary point of the ratio when reward and risk measures have the same sign. For simplicity, we assume that this optimal portfolio is the solution of the ratio problem

$$\begin{aligned} & \max_w \frac{v(w^T r - r_b)}{\rho(w^T r - r_b)} \\ & \text{s.t.} \\ & l \leq Aw \leq u \end{aligned} \quad (2.5)$$

In both problems (2.4) and (2.5), v is a functional measuring the expected reward, the vector notation $w^T r$ stands for the returns of a portfolio with composition $w = (w_1, w_2, \dots, w_n)$, l is a vector of lower bounds, A is a matrix, u is a vector of upper bounds, and r_b is some benchmark (which could be set equal to zero). The set comprised by the double linear inequalities in matrix notation $l \leq Aw \leq u$ includes all feasible portfolios. An example of a reward-risk ratio is the celebrated Sharpe ratio (see [37]). In this case, the reward measure v is a linear functional and is the expected active portfolio return $E(w^T r - r_b)$ and the risk measure ρ is represented by the standard deviation. Beside the Sharpe ratio, many more examples can be obtained by changing the risk and reward functional (see [7, 23, 27] and the references therein):

- STARR ratio: $E(w^T r - r_b)/ETL_\alpha(w^T r - r_b)$.
- Stable ratio: $E(w^T r - r_b)/\sigma_{rp}$, where σ_{rp} is the portfolio dispersion. Here it is assumed that the vector r follows a multivariate sub-Gaussian stable distribution and thus $\sigma_{rp} = (w^T Q w)^{1/2}$, where Q is the dispersion matrix (see [26]).
- Farinelli–Tibiletti ratio: $(E(\max(w^T r - t_1, 0)^\gamma))^{1/\gamma}/(E(\max(t_2 - w^T r, 0)^\delta))^{1/\delta}$, where t_1 and t_2 are some thresholds.
- Sortino–Satchell ratio: $E(w^T r - r_b)/(E(\max(t - w^T r, 0)^\gamma))^{1/\gamma}$, $\gamma \geq 1$.
- Rachev ratio (R-ratio): $ETL_\alpha(r_b - w^T r)/ETL_\beta(w^T r - r_b)$.
- Generalized Rachev ratio (GR-ratio): $ETL_{(\gamma, \alpha)}(r_b - w^T r)/ETL_{(\delta, \beta)}(w^T r - r_b)$, where $ETL_{(\gamma, \alpha)}(X) = (E((\max(-X, 0))^\gamma | -X > VaR_\alpha(X)))^{\gamma^*}$ and $\gamma^* = \min(1, 1/\gamma)$.
- VaR ratio: $VaR_\alpha(r_b - w^T r)/VaR_\beta(w^T r - r_b)$.

- Gini-type-ratio (GT-ratio): $GT_{(\beta,m)}(r_b - w^T r) / GT_{(\alpha,n)}(w^T r - r_b)m$, $n \geq 1$, $\beta \in (0,1)$ where the Gini-type measure $GT_{(\beta,v)}(X) = \frac{(v-1)v}{\beta^v} \int_0^\beta (\beta - u)^{v-2} \times u ETL_u(X) du$ is a linearizable coherent risk measure (see [23]).
- Spectral type ratio (ST-ratio) $M_{\phi 1}(r_b - w^T r) / M_{\phi 2}(w^T r - r_b)$ where $\phi 1, \phi 2$ are non-negative, decreasing and integrable functions such that $\int_0^1 \phi(u) du = 1$, and $M_\phi(X) = \int_0^1 \phi(u) VaR_u(X) du$ is a coherent risk measure identified by its risk spectrum ϕ (see [1]).

We say the reward-risk ratio is a *dispersion type ratio* if the risk measure is an uncertainty measure. Examples of dispersion type ratios are the Sharpe ratio and Stable ratio. As for the risk and reward measures we can define axiomatically coherent ratios. A *coherent ratio* is any functional G defined on the space of all admissible portfolios satisfying the following properties:

- A1. It admits the representation $G(X) = \frac{v(X)}{\rho(X)}$, where v is a reward measure, ρ is a generic risk measure that has the same sign of v for all admissible portfolios X .
- A2. The reward measure v must satisfy property M3 and the risk measure ρ must satisfy property R3.
- A3. If $X \geq Y$, then $G(X) \geq G(Y)$ provided that the reward and risk measures are both strictly positive, and $G(X) \leq G(Y)$ provided that the reward and risk measures are both strictly negative for all admissible portfolios.

We call *aggressive-coherent ratios* any functional that satisfies Properties A1, A3 and in the ratio representation both, reward and risk measures, satisfy either property M3 or property R3.

Clearly not all the above reward risk ratios are coherent, aggressive-coherent or dispersion type ratios. Typically, the ratio between a coherent reward measure v and a coherent risk measure ρ with the same sign of v for all admissible portfolios is a coherent ratio. For example, the STARR ratio and the ratio between the mean and a Gini-type measure are coherent ratios (as any ratio between the mean of gross returns and a positive coherent risk measure).

The ratio between an aggressive (coherent) reward measure v and a coherent (aggressive) risk measure ρ with the same sign of v , is an *aggressive-coherent ratio*. The R-ratio is a typical example and can be interpreted as the ratio between the average (active) profit exceeding a certain threshold and the average (active) loss below a certain level. Other examples of aggressive-coherent ratios are the GR-ratio, GT-ratio, and ST-ratio. In the R-ratio and GR-ratio, the reward functional is non-linear. The R-ratio and the GR-ratio have been proposed because there is empirical evidence that they are more appropriate for investment decisions in the case of heavy-tailed returns (see [7]).

2.5. Computational complexity and reward-risk analysis

The *computational complexity* of the portfolio selection problem is another important aspect, particularly when we assess dynamic strategies. Thus the complexity of

the optimization problem could be much higher when we solve reward-risk problems with many assets and further simplifications are necessary to solve large portfolio problems.

Depending on what properties we assume for the reward and the risk measures, we can reduce the optimal ratio problem to a simpler one, under some regularity conditions, at the price of increasing the dimension. The regularity conditions are basically strict positivity of the risk measure in the feasible set and existence of a feasible portfolio with strictly positive reward measure. Similar considerations are still valid when we minimize the ratio for strictly negative risk and reward measures. In the following we consider the maximization ratio problem for which we discuss the following cases (for more details, see [39]):

Case 1. The reward functional v is concave and the risk functional ρ is convex. Then the ratio is a quasi-concave function and the optimal ratio problem can be solved through a sequence of convex feasibility problems. The sequence of feasibility problem can be obtained using the set:

$$X = \begin{cases} \rho(w^T r - r_b) - tv(w^T r - r_b) \leq 0 \\ l \leq Aw \leq u \end{cases},$$

where t is a fixed positive number. For a given t the set X is convex and therefore we have a convex feasibility problem. A simple algorithm based on bisection can be devised so that the smallest t is found, t_{\min} , for which the set X is non-empty, for more details, see Stoyanov *et al.* [2007a]. If t_{\min} is the solution of the feasibility problem, then $1/t_{\min}$ is the value of the optimal ratio and the portfolios in the set

$$X_{\min} = \begin{cases} \rho(w^T r - r_b) - t_{\min}v(w^T r - r_b) \leq 0 \\ l \leq Aw \leq u \end{cases},$$

are the optimal portfolios of the ratio problem (2.5).

Case 2. If, in addition to the conditions in Case 1, both functions are positively homogeneous, then the optimal ratio problem reduces to a convex programming problem. An example of an equivalent convex problem to (2.5) is

$$\begin{aligned} & \min_{(x,t)} \rho(x^T r - tr_b) \\ & \text{s.t.} \\ & v(x^T r - tr_b) \geq 1, \\ & lt \leq Ax \leq ut \\ & t \geq 0 \end{aligned} \tag{2.6}$$

where t is an additional variable. If (x_o, t_o) is a solution to (2.6), then $w_o = x_o/t_o$ is a solution to problem (2.5). There are other connections between problems (2.5) and (2.6). Let ρ_o be the value of the objective at the optimal point (x_o, t_o) in problem (2.6). Then $1/\rho_o$ is the value of the optimal ratio, i.e., the optimal value of

the objective of problem (2.5). Moreover, $1/t_o$ is the reward of the optimal portfolio and ρ_o/t_o is the risk of the optimal portfolio.

Case 3. If in addition to the conditions in Case 2, the reward function is linear (or linearizable), and the risk function is an increasing function of a quadratic form, then the optimal portfolio problem reduces to a quadratic programming problem. The Sharpe ratio and the Stable ratio are examples. An example of an equivalent problem to the Sharpe ratio problem is

$$\begin{aligned} & \min_{(x,t)} (-t, x) \Sigma_1 (-t, x)^T \\ & \text{s.t.} \\ & x^T E r - t E r_b = 1 \quad , \\ & l t \leq A x \leq u t \\ & t \geq 0 \end{aligned} \quad (2.7)$$

where Σ_1 is the covariance matrix

$$\Sigma_1 = \begin{pmatrix} \sigma_b^2 & \sigma_{br} \\ \sigma_{br}^T & \Sigma \end{pmatrix},$$

where Σ is the covariance matrix between portfolio asset returns, σ_b^2 is the variance of the benchmark portfolio returns and $\sigma_{br} = (\text{cov}(r_b, r_1), \text{cov}(r_b, r_2), \dots, \text{cov}(r_b, r_n))$ is a vector of covariances between the benchmark portfolio returns and the returns of the main portfolio asset returns. Again, if (x_o, t_o) is a solution to (2.7), then $w_o = x_o/t_o$ is a solution to the version of problem (2.5) in which the reward function is the mathematical expectation and the risk function is the standard deviation — the optimal Sharpe ratio problem. The connections between problems (2.7) and (2.5) are the same as the ones given in Case 2 above as problem (2.7) is just a particular version of problem (2.6).

Case 4. If the reward function is linear (or linearizable) and the risk function is linearizable, then the optimal ratio problem reduces to a linear programming problem. An example of such a problem in which we have the ETL as the risk measure, i.e., the STARR ratio problem, is readily obtained from the corresponding version of problem (2.6) by incorporating the linearization:

$$\begin{aligned} & \min_{(x,t,d,\theta)} \theta + \frac{1}{N\alpha} \sum_{k=1}^N d_k \\ & \text{s.t.} \\ & x^T E r - t E r_b = 1 \\ & -x^T r^k + t r_b^k - \theta \leq d_k, \quad k = 1, 2, \dots, N \\ & l t \leq A x \leq u t \\ & t \geq 0, \quad d_k \geq 0, \quad k = 1, 2, \dots, N, \end{aligned} \quad (2.8)$$

where r^k and r_b^k $k = 1, 2, \dots, N$ are scenarios for the vector of portfolio items returns and the benchmark portfolio returns accordingly, $d = (d_1, d_2, \dots, d_N)$ is a

vector of additional variables, θ and t are also additional variables. The relations between (2.8) and (2.5) are the same as the ones in Case 2. One should bear in mind that in the objective of problem (2.8) we have a linear approximation of the ETL function which is possible due to the available scenarios. Thus the objective at the optimal point is an approximation of the optimal ETL. For more details about linearization, see Rockafellar and Uryasev [29].

Clearly, as we have noted, the dimension of the optimization problem increases as we simplify the problem structure. If in practice the computational burden increases a lot, the reduction may not be considered. For instance the STARR ratio problem can be solved either as a linear programming problem or as a convex problem or as a sequence of convex feasibility problems depending on which is more practical. Unfortunately this classification is not complete in the sense that there are reward-risk ratios that are not quasi-concave, such as the R-ratio, the GR-ratio, and the Farinelli–Tibiletti ratio. One way to solve such a problem is to search for a local solution making use of quasi-Newton-type techniques.

It is very important for the optimal ratio problem that the risk measure be strictly positive (negative) for all feasible portfolios. If there exists a feasible portfolio with a negative (positive) risk measure in the interior of the feasible set, then the continuity of the risk function in the optimal ratio problem suggests that there will be a feasible portfolio for which the reward-risk ratio explodes. The risk function is continuous in an open set, since it is convex. Thus, sometimes it might be more appropriate to consider linearized versions of the reward-risk ratios, that is

$$av(w^T r - r_b) - \lambda \rho(w^T r - r_b), \quad (2.9)$$

where v is a reward measure, ρ is a generic risk measure, and $a \geq 0$, $\lambda \geq 0$ are risk aversion parameters that are not both equal to zero.

For example, the linearized version of the STARR ratio is

$$E(w^T r - r_b) - \lambda ETL_\alpha(w^T r - r_b).$$

In the special case of $a=\lambda=1$, if the reward functional is the mathematical expectation and the risk measure has the property R1 and is strictly expectation bounded, then expressions of type (2.9) are deviation measures, i.e., they satisfy all axioms of deviation measures (for more details and a proof, see [30]). Strict expectation boundedness means that the risk measure satisfies all properties R1, R2, and R3 (not necessarily R4) and also $\rho(X) > E(-X)$ for all non-constant X . Moreover, expression (2.9) includes most of the previous ones. As a matter of fact, we have just discussed the equivalence between optimal *reward-risk* ratios and the types of measures defined in (2.9). We obtain a generic risk measure when in formula (2.9) we consider $a = 0$, $\lambda = 1$. Similarly, we obtain reward measures when in formula (2.9) we consider $a = 1$, $\lambda = 0$.

Certainly an optimization problem in which we have a linearized reward-risk ratio in the objective with a pre-selected value for λ and a is equivalent to

problem (2.4) with a suitable choice of the limit R . Objectives of type (2.9) can also be regarded as utility functions with a special structure. The corresponding optimization problems are reducible to convex ones and it is not necessary to impose assumptions about the positivity of the risk function.

From the above discussion we deduce that functionals of type (2.9) provide the largest class of reward-risk portfolio selection problems investors should maximize (taking into account the complexity of the optimization problem). In particular, we could also consider the financial insight of the typical coherency axioms assuming the following new class of “utility measures” applied to all admissible portfolios.

Thus, we say the functional S is a *coherent reward-risk utility functional* if it satisfies the following properties:

- B1. It admits the decomposition $S(X) = av(X) - \lambda\rho(X)$ for all X , where v is a reward measure, ρ is a generic risk measure, $a \geq 0$, $\lambda \geq 0$ and they are not both equal to zero.
- B2. $S(0) = 0$, and $S(bX) = bS(X)$ for all X and $b > 0$.
- B3. $S(X + Y) \geq S(X) + S(Y)$ for all X and Y .
- B4. $S(X + C) \geq S(X)$ for all X and constants $C \geq 0$ and S is isotonic with the order of preference of the market, i.e., if all investors prefer X to Y then $S(X) \geq S(Y)$.

In addition, we say the functional S is an *aggressive-coherent reward-risk utility functional* if it satisfies Property B4 and it admits the decomposition B1 such that either

$$\begin{aligned} v(X + Y) &\geq v(X) + v(Y), & \rho(X + Y) &\geq \rho(X) + \rho(Y); & \text{ or} \\ v(X + Y) &\leq v(X) + v(Y), & \rho(X + Y) &\leq \rho(X) + \rho(Y). \end{aligned}$$

In the definition of coherent reward-risk utility functional we recognize the importance of risk aversion considering properties B2 and B3 that imply the concavity of the functional. At the same time, with property B4 we identify the most important aspect in portfolio theory that is the isotonicity with stochastic orders. Thus, in a market where investors are non-satiable and risk averse we could consider, v and ρ (in the decomposition of property B1) respectively, a coherent reward measure and a coherent risk measure for any positive a and λ . In addition, we could combine the aggressiveness of a reward measure with the coherency of a risk measure maximizing aggressive-coherent reward-risk utility functionals. A special case of aggressive-coherent reward-risk utility functional is the utility version of the R-ratio:

$$aETL_{\alpha}(r_b - w^T r) - \lambda ETL_{\beta}(w^T r - r_b).$$

Aggressive-coherent reward-risk utility functionals may be used to identify optimal portfolios selected by investors who are neither risk averters nor risk preferred — for example investors with Friedman–Savage [13] type utility functions. Therefore,

among the many measures proposed in portfolio literature, we noted that:

- (a) coherent risk, reward measures, coherent ratios and coherent reward–risk utility functionals serve to identify optimal choices of non-satiable risk averse investors;
- (b) aggressive risk, reward measures serve to identify optimal choices of non-satiable risk preferring individuals;
- (c) aggressive-coherent ratios and reward–risk utility functionals serve to identify optimal choices of non-satiable investors who are neither risk averters nor risk preferred.

Table 1 summarizes properties and examples of the main measures and functionals proposed in this section to solve portfolio selection problems.

3. A Parallel between Uncertainty Measures and the Theory of Probability Metrics

Let us consider the problem of benchmark tracking. A common formulation of the problem is

$$\min_{w \in X} \sigma(w^T r - r_b),$$

where σ is the standard deviation and X is the set of feasible portfolios. The measure shown in the objective function is called the *tracking error*. It is the standard deviation of active returns. In essence, by solving this problem we are trying to stay “close” to the benchmark while satisfying the constraints where the degree of proximity is calculated making use of the standard deviation. Here the benchmark r_b can be either stochastic or non-stochastic.

Certainly this problem can be reformulated using any uncertainty measure in the objective. Even more generally, this problem can be considered from the point of view of the theory of probability metrics, the rationale being that, under the most general conditions, the distance between two random variables can only be defined via a probability distance.

Let $\Lambda = \Lambda(R)$ be the set of all real-valued random variables on a given probability space (Ω, F, Pr) . A probability distance μ with a parameter K is a functional defined on the space of all joint probability distributions $Pr_{X,Y}$ generated by the pairs of random variables $X, Y \in \Lambda$ satisfying

- i. (identity) $Pr(X = Y) = 1 \Leftrightarrow \mu(X, Y) = 0$.
- ii. (symmetry) $\mu(X, Y) = \mu(Y, X)$.
- iii. (triangle inequality) $\mu(X, Z) \leq K(\mu(X, Y) + \mu(Y, Z))$ for all X, Y, Z in Λ .

If the parameter K is equal to 1, then the probability distance is called a *probability metric* in line with the usual triangle inequality defining a metric.

Generally, there are three types of probability distances — *primary*, *simple* and *compound* — depending on certain modifications of the identity property — whether $\mu(X, Y) = 0$ implies that only certain moment characteristics of X and Y agree

Table 1. Properties and examples of the main reward, risk measures and functionals used in portfolio theory.

Measures used in portfolio theory	Properties satisfied for any couple of random variables X and Y	Examples
Generic risk measure	It is either an uncertainty measure or a measure consistent with first stochastic order	All the following risk measures
Uncertainty measure	Any increasing function of functional D that satisfies: (i) $D(X + C) \leq D(X) \ \forall \ C \geq 0$; (ii) $D(aX) = aD(X) \ \forall \ a \geq 0$; (iii) $D(X) \geq 0$ and $D(X) = 0$ only if X is constant.	$Colog(X)$ and all deviation measures
Deviation measure	(i) $D(X + C) = D(X) \ \forall \ \text{constant } C$; (ii) $D(aX) = aD(X) \ \forall \ a \geq 0$; (iii) $D(X + Y) \leq D(X) + D(Y)$; (iv) $D(X) \geq 0$ and $D(X) = 0$ only if X is constant.	Standard deviation; MAD; semi-stand. deviation
Expected bounded risk measure	(i) $\rho(X + C) = \rho(X) - C \ \forall \ \text{constant } C$; (ii) $\rho(aX) = a\rho(X) \ \forall \ a \geq 0$; (iii) $\rho(X + Y) \leq \rho(X) + \rho(Y)$; (iv) $\rho(X) > E(-X)$ for all non-constant X .	MiniMax and $ETL_{\alpha}(X)$
Coherent risk measure	(i) $\rho(X + C) = \rho(X) - C \ \forall \ \text{constant } C$; (ii) $\rho(aX) = a\rho(X) \ \forall \ a \geq 0$; (iii) $\rho(X + Y) \leq \rho(X) + \rho(Y)$; (iv) $\rho(X) \leq \rho(Y)$ if $X \geq Y$.	$ETL_{\alpha}(X)$; Spectral measures; Gini type measures
Aggressive risk measure	(i) $\rho(X + C) = \rho(X) - C \ \forall \ \text{constant } C$; (ii) $\rho(aX) = a\rho(X) \ \forall \ a \geq 0$; (iii) $\rho(X + Y) \geq \rho(X) + \rho(Y)$; (iv) $\rho(X) \leq \rho(Y)$ if $X \geq Y$.	$-ETL_{\alpha}(-X)$
Reward measure	Any measure isotonic to investors' preferences	Depends on the class of investors
Coherent reward measure	(i) $v(X + C) = v(X) + C \ \forall \ \text{constant } C$; (ii) $v(aX) = av(X) \ \forall \ a \geq 0$; (iii) $v(X + Y) \geq v(X) + v(Y)$; (iv) $v(X) \geq v(Y)$ if $X \geq Y$.	Opposite of coherent risk measure
Aggressive reward measure	(i) $v(X + C) = v(X) + C \ \forall \ \text{constant } C$; (ii) $v(aX) = av(X) \ \forall \ a \geq 0$; (iii) $v(X + Y) \leq v(X) + v(Y)$; (iv) $v(X) \geq v(Y)$ if $X \geq Y$.	Coherent risk measures applied to opposite of random variables
Reward-risk ratio	Any ratio between a reward measure v and a generic risk measure ρ with the same sign of v .	All the following reward risk ratios
Dispersion type reward-risk ratio	A reward-risk ratio where the risk measure is an uncertainty measure	Stable and Sharpe ratios

Table 1. (Continued)

Measures used in portfolio theory	Properties satisfied for any couple of random variables X and Y	Examples
Coherent reward-risk ratio	(i) $G(X) = v(X)/\rho(X)$ is a reward-risk ratio; (ii) If $X \geq Y$, then $G(X) \geq G(Y)$ provided that the reward and risk measures are both strictly positive, and $G(X) \leq G(Y)$ provided that the reward and risk measures are both strictly negative; (iii) $v(X + Y) \geq v(X) + v(Y)$, $\rho(X + Y) \leq \rho(X) + \rho(Y)$;	STARR ratio; the ratio between the mean and a coherent risk measure
Aggressive-coherent reward-risk ratio	(i) and (ii) as coherent reward-risk ratios and (iii) either $v(X + Y) \geq v(X) + v(Y)$, $\rho(X + Y) \geq \rho(X) + \rho(Y)$; or $v(X + Y) \leq v(X) + v(Y)$, $\rho(X + Y) \leq \rho(X) + \rho(Y)$.	R-; GR-, GT-and ST-ratios.
Reward-risk utility functional	$S(X) = av(X) - \lambda\rho(X)$ where v is a reward measure, ρ is a generic risk measure, $a \geq 0$, $\lambda \geq 0$ are not both equal to zero.	All the reward-risk utility functionals
Coherent reward-risk utility functional	(i) S is a reward-risk utility functional; (ii) $S(X + C) \geq S(X) \forall C \geq 0$ and S is isotonic with the order of preference of the market; (iii) $S(aX) = aS(X) \forall a \geq 0$; (iv) $S(X + Y) \geq S(X) + S(Y)$;	When v and ρ are respectively coherent reward and risk measure
Aggressive-coherent reward-risk utility functional	Property (i) and (ii) of coherent reward-risk utility functionals and (iii) either $v(X + Y) \geq v(X) + v(Y)$, $\rho(X + Y) \geq \rho(X) + \rho(Y)$; or $v(X + Y) \leq v(X) + v(Y)$, $\rho(X + Y) \leq \rho(X) + \rho(Y)$	One measure between v and ρ is aggressive and the other is coherent

or that only the cumulative distribution functions of X and Y coincide or that $Pr(X = Y) = 1$. For the purpose of restatement of the benchmark tracking problem, we shall first define the three types and give some examples.

In order to provide a formal definition of a primary probability distance, another notation is required. Let h be a mapping defined on Λ with values in R^J , that is we associate a vector of numbers with a random variable. The vector of numbers could be interpreted as a set of some characteristics of the random variable. An example of such a mapping is: $X \rightarrow (EX, \sigma_X)$ where the first element is the mathematical expectation and the second is the standard deviation. In particular, if the random variable is interpreted as investment returns, then the first element is the expected return and the second is a measure of the uncertainty. Similarly, we can extend the vector to include any (finite) number of characteristics among which we can have measures of risk, uncertainty, reward measures, etc.

Furthermore, the mapping h induces a partition of Λ into classes of equivalence. That is, two random variables X and Y are regarded as equivalent, $X \sim Y$, if their corresponding characteristics agree:

$$X \sim Y \Leftrightarrow h(X) = h(Y).$$

Since the probability distance is defined on the space of pairs of random variables, we have to translate the equivalence into the case of pairs of random variables. Two sets of pairs (X_1, Y_1) and (X_2, Y_2) are said to be equivalent if there is equivalence on an element-by-element basis, i.e., $h(X_1) = h(X_2)$ and $h(Y_1) = h(Y_2)$.

Let μ be a probability distance such that μ is constant on the equivalence classes induced by the mapping h :

$$(X_1, Y_1) \sim (X_2, Y_2) \Leftrightarrow \mu(X_1, Y_1) = \mu(X_2, Y_2).$$

Then μ is called *primary* probability distance. Examples of primary probability distances include:

- $\mu(X, Y) = |EX - EY|$, here h is the mapping $X \rightarrow EX$.
- $\mu(X, Y) = |(E|X|^p)^{1/p} - (E|Y|^p)^{1/p}|$, $p \geq 1$; here h is the mapping $X \rightarrow (E|X|^p)^{1/p}$.
- $\mu(X, Y) = |h_1(X) - h_1(Y)| + |h_2(X) - h_2(Y)|$, here h is the mapping $X \rightarrow (h_1(X), h_2(X))$.

As we have remarked, a simple probability distance is such that $\mu(X, Y) = 0$ implies that $F_X(t) = F_Y(t)$ where $F_X(t) = P(X < t)$ is the cumulative probability distribution function. Examples of simple probability distances include

- $\mu(X, Y) = \sup_t |F_X(t) - F_Y(t)|$ which is also known as the uniform (or Kolmogorov) metric.
- $\mu(X, Y) = \int_R |F_X(t) - F_Y(t)| dt$, which is also known as the Kantorovich metric.
- $\mu(X, Y) = [\int_R |F_X(t) - F_Y(t)|^p dt]^{1/p}$, $p \geq 1$, which is also known as the class of L_p metrics.

A compound probability distance is such that $\mu(X, Y) = 0$ implies $Pr(X = Y) = 1$. Examples include:

- $\mu(X, Y) = (E|X - Y|^p)^{1/p}$.
- $\mu(X, Y) = \inf\{\varepsilon > 0 : Pr(|X - Y| > \varepsilon) < \varepsilon\}$ and $\mu(X, Y) = E \frac{|X - Y|}{1 + |X - Y|}$, both are also known as the Ky Fan metrics.

For many more examples of the various types of probability distances and approaches to construct them, see Rachev [28] and Stoyanov *et al.* [40].

The benchmark tracking problem that we started with can be reformulated in the following way:

$$\min_{w \in X} \mu(w^T r, r_b), \quad (3.1)$$

where μ is some probability distance. The second argument in the probability distance does not change with w ; hence in solving the problem we intend to “approach” the benchmark, but changing the type the probability distance changes the perspective. If we would like only certain characteristics of our portfolio to be as close as possible to the corresponding characteristics of the benchmark, we can use a primary probability distance. When the objective for our portfolio is to mimic the entire distribution of the benchmark, not just some characteristics of it, then we should use a simple probability distance. Finally, if we would like to replicate the benchmark exactly, then we should use a compound probability distance. In particular, in Table 2 we summarize the main properties and examples of tracking error type measures.

In the initial benchmark tracking problem, we have a compound probability distance as the objective function because the standard deviation is just one example of an L_p metric in the space of random variables with finite variance. Therefore, relating the benchmark tracking problem to the theory of probability distances represents a significant extension of the initial problem.

Table 2. Properties and examples of the main tracking error type measures.

Measures	Properties of tracking error type measures	Examples
Primary probability distance (metric)	(i) $\mu(X, Y) = 0 \Leftrightarrow$ only certain characteristics of X and Y agree (ii) $\mu(X, Y) = \mu(Y, X)$ (iii) $\mu(X, Z) \leq K(\mu(X, Y) + \mu(Y, Z))$ for all real random variables X, Y, Z defined on the probability space (primary probability metric if $K = 1$)	$\mu(X, Y) = EX - EY ,$ $\mu(X, Y) = (E X ^p)^{1/p} - (E Y ^p)^{1/p} $
Simple probability distance (metric)	(i) $\mu(X, Y) = 0 \Leftrightarrow X$ and Y have the same distribution (ii) $\mu(X, Y) = \mu(Y, X)$ (iii) $\mu(X, Z) \leq K(\mu(X, Y) + \mu(Y, Z))$ for all real random variables X, Y, Z defined on the probability space (simple probability metric if $K = 1$)	$\mu(X, Y) = \sup_t F_X(t) - F_Y(t) $ $\mu(X, Y) = \int_R F_X(t) - F_Y(t) dt$
Compound probability distance (metric)	(i) $\mu(X, Y) = 0 \Leftrightarrow (Pr(X = Y) = 1)$ (ii) $\mu(X, Y) = \mu(Y, X)$ (iii) $\mu(X, Z) \leq K(\mu(X, Y) + \mu(Y, Z))$ for all real random variables X, Y, Z defined on the probability space (compound probability metric if $K = 1$)	$\mu(X, Y) = (E X - Y ^p)^{1/p}$ $\mu(X, Y) = E \frac{ X - Y }{1 + X - Y }$

In addition, it should be noted that, some risk measures and reward-risk ratios have properties similar to some probability metrics. For example, let us consider a version of the GR-ratio in which $\gamma = \delta$

$$GR_\gamma = ETL_{(\gamma,\alpha)}(r_b - w^T r) / ETL_{(\gamma,\beta)}(w^T r - r_b), \quad (3.2)$$

where $ETL_{(\gamma,\alpha)}(X) = (E((\max(-X, 0))^\gamma | -X > VaR_\alpha(X)))^{\gamma^*}$ and $\gamma^* = \min(1, 1/\gamma)$. Letting γ approach zero and infinity, at the limit we obtain expressions close to the corresponding expressions of the L_p metric. That is, as $\gamma \rightarrow \infty$ we obtain

$$GR_\infty = ETL_{(\infty,\alpha)}(r_b - w^T r) / ETL_{(\infty,\beta)}(w^T r - r_b),$$

where $ETL_{(\infty,\alpha)}(X) = \text{ess sup}(\max(-X, 0) | -X > VaR_\alpha(X))$. Here *ess sup* stands for the essential supremum. At the other limit, as $\gamma \rightarrow 0$,

$$GR_0 = ETL_{(0,\alpha)}(r_b - w^T r) / ETL_{(0,\beta)}(w^T r - r_b)$$

where $ETL_{(0,\alpha)}(X) = Pr(\{-X > 0\} \cap \{-X > VaR_\alpha(X)\})$, i.e., if $VaR_\alpha(X) > 0$, then $ETL_{(0,\alpha)}(X) = \alpha$ and if $VaR_\alpha(X) \leq 0$, then $ETL_{(0,\alpha)}(X) = Pr(-X > 0)$.

The parallels considered suggest that there is an interesting relationship between the well-developed theory of probability metrics and the theory of optimal portfolio choice. This interplay might throw more light on the relationship between different classes of risk measures and/or uncertainty measures. Moreover, it might suggest an approach to select an ideal risk measure or an ideal performance ratio for a particular portfolio choice problem just as there is an ideal probability metric for a given approximation problem in probability theory. For this reason, we think that the established relationship should be extended and better studied in future research.

4. Temporal Horizon and Aggregated Risk

We will use the following empirical example to evaluate and qualify several other aspects of investor preferences. Let us consider the portfolio selection among 13 developed country stock market indices: Australia, Canada, France, Germany, Hong Kong, Italy, Japan, Netherlands, Singapore, South Africa Gold, Sweden, United Kingdom, and United States. The stock indices are part of the MSCI World Index for the period January 4, 1993 to May 31, 2004. Part of this historical data (during the period 1/4/1993–1/1/2002) is used to estimate the different parameters of the models. Then over a period of two years and five months (1/1/2002–5/31/2004), we verify the behavior of the models. We share the period of analysis in this way in order to have enough observations to get robust statistics. Thus, the vector of returns is given by $r' = [r_1, \dots, r_{13}]$ and vector of wealth is $w' = [w_1, \dots, w_{13}]$. In addition, we assume a risk-free asset proxied by 30-day Eurodollar CD (and offering one-month Libor) that on 1/1/2002 was $r_0 = 1.87\%$.

Considering daily data for this period, we first value two “optimal portfolios” when no short sales are allowed (i.e., $w_i \geq 0$) and it is not possible to invest more

than 25% (i.e., $w_i \leq 0.25$) of the initial capital (that we assume to be equal to 1) in a single asset:

- (a) The first portfolio is the global minimum variance portfolio (that is a strong risk-averse choice).
- (b) The second portfolio maximizes the R-ratio

$$ETL_{40\%}(r_0 - w^T r) / ETL_{1\%}(w^T r - r_0). \quad (4.1)$$

Thus, the first portfolio minimizes the uncertainty and, as intuition suggests, it presents the highest level of *risk aversion*. With the second portfolio we do not minimize the uncertainty, but we take into account downside risk. As a matter of fact, the risk measure expected shortfall $ETL_{1\%}(w^T r - r_0)$ considers the portfolio downside risk, and the reward measure $ETL_{40\%}(r_0 - w^T r)$ takes into consideration the possible profits. Therefore, the second portfolio maximizes the excess return considering the greatest profits and at the same time minimizing and controlling the biggest losses.

Alternatively, we compute two other portfolios with the same restrictions of the previous ones (i.e., $0 \leq w_i \leq 0.25$), but assuming yearly returns (261 days of returns) with daily frequency for the same time period. The two portfolios are again the global minimum variance portfolio and the portfolio that maximizes the R-ratio given by (4.1).

Figure 6 proposes an *ex-post* comparison of the final wealth during the period 1/1/2002–5/31/2004. That is, we consider the *ex-post* final wealth of the four optimal portfolios assuming a unitary wealth is invested on 1/1/2002. In particular, series dayminvar and daymaxRR describe respectively, the final wealth behavior of the two daily optimal return portfolios (global minimum variance and the portfolio that maximizes the R-ratio given by (4.1)), while series yearminvar and yearmaxRR represents respectively, the final wealth graph of the two optimal yearly return portfolios. Figure 6 indicates and emphasizes the differences among the four portfolios which are coherent with the different choices made. As a matter of fact, the minimum variance portfolios (series dayminvar and yearminvar) generally present a lower final wealth than the maximum ratio portfolios. In addition, we observe that the final wealth obtained with optimal portfolios based on yearly returns is generally higher than that obtained with optimal portfolios valued on daily returns. Therefore the investor's *temporal horizon* and the relative aggregated risk influence his/her future choices.

This behavior is confirmed by the results reported in Table 3, which shows the *ex-ante* and *ex-post* VaR and ETL (for two confidence levels, 99% and 95%) based on daily returns of the four optimal portfolios. The *ex-ante* analysis clearly indicates that the minimum variance portfolios (portfolios 1 and 3) provide a lower dispersion (standard deviation) and a higher risk of big losses (VaR and ETL) than portfolios that maximize the R-ratio given by (4.1) (respectively portfolios 2 and 4). Thus the *ex-ante* analysis suggests that the more conservative minimum variance portfolios

(portfolios 1 and 3) do not always take into account the possibility of big losses. In contrast, the *ex-post* analysis shows the differences between the first two and the last two portfolios. Therefore, Fig. 6 and Table 3 emphasize how investor's *temporal horizon* could influence the agents' future choices.

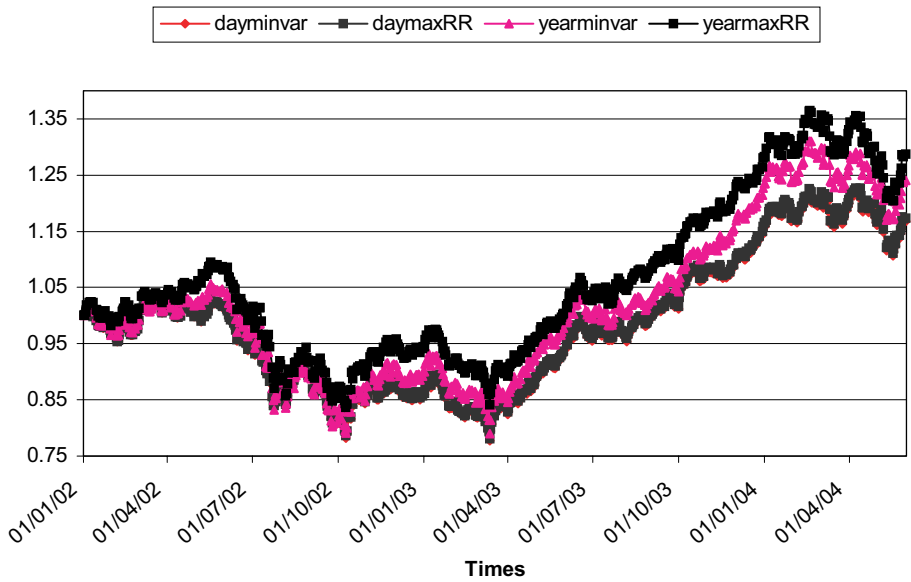


Fig. 6. *Ex-post* comparison among the final wealth of the four optimal portfolios obtained with daily returns (series dayminvar and daymaxRR) and six months returns and daily frequency (series yearminvar and yearmaxRR).

Table 3. *Ex-ante* and *ex-post* valuation of the four portfolio risk positions.

	Optimal portfolios on daily data		Optimal portfolios on yearly data and daily frequency	
	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4
<i>EX-POST</i> Analysis				
Final Wealth	1.1678741	1.1738809	1.24117601	1.287114
VaR _{1%}	0.0219778	0.0221615	0.02891808	0.024173
VaR _{5%}	0.0144643	0.0141017	0.01736294	0.015798
ETL _{1%}	0.026436	0.0261711	0.03146571	0.028343
ETL _{5%}	0.0189137	0.0187883	0.02423259	0.021512
Standard Deviation	0.0084353	0.008361	0.01011825	0.009037
<i>EX-ANTE</i> Analysis				
VaR _{1%}	0.0183934	0.0180712	0.02297311	0.022305
VaR _{5%}	0.0114121	0.0113635	0.0128053	0.013089
ETL _{1%}	0.0235661	0.0235116	0.02910487	0.030808
ETL _{5%}	0.0160334	0.0159598	0.01898697	0.019467
Standard Deviation	0.0069074	0.0069197	0.00798875	0.008152

As a matter of fact, the risk portfolio on one day is generally different from the risk portfolio based on one year, and the forecasting analysis has to take into account the *aggregated risk* possibly considering also the conditional heteroskedasticity of portfolio series.

Moreover, the following three questions are raised by this example:

- (1) What are the best risk measures: the best ratios and the best reward measures?
- (2) Which reward-risk measure takes into account downside risk and offers flexibility with respect to risk aversion?
- (3) What different roles are covered by reward measures and ratios?

In addition, we want to better understand the impact of transaction costs on portfolio dynamic strategies.

5. Dynamic Strategies and Transaction Costs

In order to value the impact of transaction costs and of different reward-risk ratios in a dynamic setting, we propose two empirical comparisons.

5.1. Impact of proportional transaction costs

Let us consider the portfolio selection among the one-month Libor risk-free asset (that was $r_0 = 1.32\%$ on 5/30/2003) and the same 13 international indexes used in the previous example based on the period 1/3/2000–5/31/2004. We assume that on 5/30/2003 the agent invests a unitary capital and recalibrates the portfolio monthly considering that no short sales are allowed and it is not possible to invest in a single asset more than 25% of the initial capital. Hence, we consider dynamic strategies with and without constant and proportional transaction costs of 0.5%. Then, we compare dynamic portfolio strategies with and without constant proportional transaction costs. In particular we assume that after k months the investor chooses the portfolio composition $x_{(k)} = [x_{(k),1}, \dots, x_{(k),13}]'$ that maximizes the Sharpe ratio. That is, the investor solves the problem

$$\begin{aligned} & \max_{x_{(k)}} \frac{E(X)}{\sqrt{E((X - E(X))^2)}} \\ & 0 \leq x_{(k),i} \leq 0.25 \\ & \text{subject to } \sum_{i=1}^{13} x_{(k),i} = 1, \end{aligned} \quad (5.1)$$

where

$$X = \begin{cases} x'_{(k)} r - r_0 & \text{without transaction costs} \\ x'_{(k)} r - r_0 - 0.005 \sum_{i=1}^{13} \left| x_i^{(k)} - \frac{x_i^{(k-1)}(1 + r_i^{(k-1)})}{\sum_{i=1}^n x_i^{(k-1)}(1 + r_i^{(k-1)})} \right| & \text{with tr. costs} \end{cases}$$

and $r^{(k)} = [r_1^{(k)}, \dots, r_{13}^{(k)}]'$ is the k th *ex-post* monthly observation of return vector r . In addition, we assume that the variance follows the exponential weighted model $\sigma_{t+1/t}^2 = \lambda \sigma_{t/t-1}^2 + (1 - \lambda) X_t^2$ with $\lambda = 0.94$ (as suggested by the RiskMetrics approach described in [19]). Then, the investor's wealth after k months is given by

$$W_k = \begin{cases} W_{k-1}(1 + x'_{(k)} r^{(k)}) & \text{without transaction costs} \\ W_{k-1} \left(1 + x'_{(k)} r^{(k)} - 0.005 \sum_{i=1}^{13} \left| x_i^{(k)} - \frac{x_i^{(k-1)}(1 + r_i^{(k-1)})}{\sum_{i=1}^n x_i^{(k-1)}(1 + r_i^{(k-1)})} \right| \right) & \text{with tr. costs} \end{cases} \quad (5.2)$$

Figure 7 shows the final wealth process with and without transaction costs during the period 5/30/2003–5/31/2004. After 12 recalibrations, the difference between the final wealth obtained without transaction costs and with transaction costs was about 1% (of the initial wealth). In particular, when we consider transaction costs, we observe that the portfolio composition does not change during the monthly recalibrations. Therefore, in this case the transaction costs do not have a material impact on the investor choices, but they could have, when the investor is more sensitive toward risk.

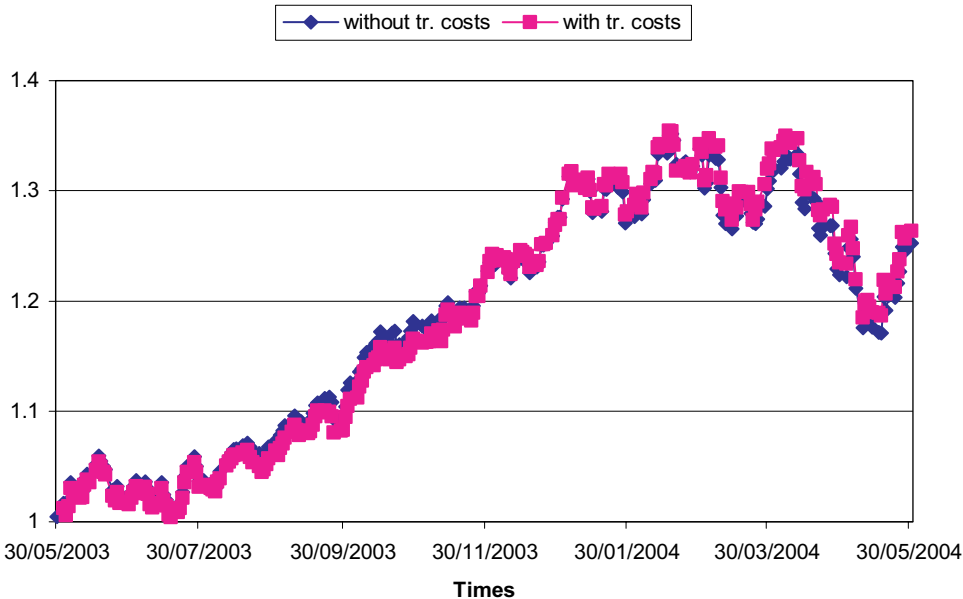


Fig. 7. *Ex-post* comparison between the final wealth of the optimal dynamic strategies obtained by maximizing monthly Sharpe ratio with and without transaction costs.

5.2. Impact of different reward risk ratios

Let us compare the *ex-post* performance of three different reward-risk strategies. We consider the portfolio selection among the one-month Libor risk-free asset r_0 and the same 13 international indexes used in the previous example based on the period 1/4/1993–5/31/2004. We assume an agent invests an unitary capital on 11/1/96 (i.e., $W_0 = 1$) and he/she recalibrates the portfolio daily when no short sales are allowed (i.e., $w_i \geq 0$). In particular, we compare the optimal dynamic strategies of investors that maximize either a dispersion type ratio (Sharpe ratio), a coherent ratio (STARR ratio $E(w^T r - r_0)/ETL_{1\%}(w^T r - r_0)$) and an aggressive-coherent ratio (R-ratio $ETL_{1\%}(r_0 - w^T r)/ETL_{1\%}(w^T r - r_0)$) during the period 11/1/96–5/31/2004. Every day we solve the optimization problem using observations from the prior 1,000 days. Once we determine the optimal portfolio $x_{(k)}$, the investor's wealth after k days is given by

$$W_k = W_{k-1}(1 + x'_{(k)}r^{(k)}),$$

where $r^{(k)}$ is the k th *ex-post* daily observation of return vector r . We solved the optimization problem every day from date 11/1/96 till 5/31/2004 (a total of 1,974 times).

Table 4 reports the average of the 1,974 optimal portfolio weights obtained from the different strategies and the half absolute differences among any couple of portfolio composition vectors. That is, we compute the distance $\frac{1}{2} \sum_{i=1}^{13} |x_i^{(STARR)} - x_i^{(Sharpe)}|$, $\frac{1}{2} \sum_{i=1}^{13} |x_i^{(R)} - x_i^{(Sharpe)}|$, $\frac{1}{2} \sum_{i=1}^{13} |x_i^{(STARR)} - x_i^{(R)}|$,

Table 4. Averages and differences among optimal weights obtained under different dynamic strategies (based on the STARR, R, and Sharpe ratios).

Indexes of the MSCI World Index	Structure of optimal portfolios (average over 1974 optimal portfolio weights)		
	STARR ratio	R-Ratio	Sharpe ratio
Australia	0,0515	0,0284	0,0722
Canada	0,0576	0,0529	0,0801
France	0,1081	0,0015	0,1160
Germany	0,0005	0,0654	0,0104
Hong Kong	0,0211	0,0310	0,0208
Italy	0,0528	0,1181	0,0380
Japan	0,0000	0,3289	0,0000
Netherlands	0,0498	0,0121	0,0639
Singapore	0,0546	0,1593	0,0619
South African Gold	0,1383	0,0349	0,0959
Sweden	0,2071	0,0226	0,1015
UK	0,0049	0,0199	0,0251
USA	0,2537	0,1250	0,3142

$\frac{1}{2} \sum_{i=1}^{13} |x_i^{(STARR)} - x_i^{(Sharpe)}| = 16.31\%; \frac{1}{2} \sum_{i=1}^{13} |x_i^{(STARR)} - x_i^{(R)}| = 58.87\%;$
 $\frac{1}{2} \sum_{i=1}^{13} |x_i^{(R)} - x_i^{(Sharpe)}| = 57.16\%$

where $x_i^{(\text{STARR})}$, $x_i^{(\text{Sharpe})}$, $x_i^{(R)}$ are the i th components of the average vectors of weights obtained maximizing respectively the STARR, Sharpe, and R-ratios. These values give an idea on how the average portfolio composition changes, adopting the different strategies. The difference in the average of the portfolio compositions is much higher (more than 57%) when we consider the optimal weights obtained with the R-ratio and the other strategies than the difference between the Sharpe and STARR strategies (about 16.31%). In particular, we observe that the Japan index gives the largest weight (on average about the 32.89%) in the optimal portfolio composition of the strategy based on the R-ratio. However, the Japan index does not enter into the composition of the optimal portfolios obtained maximizing the STARR and Sharpe ratios. Figure 8 shows the *ex-post* final wealth of the three strategies during the period under consideration. This figure shows that the optimal portfolios obtained by applying the R-ratio outperform the optimal portfolios obtained using alternative ratios over the entire period. In particular, on 5/31/2004 the final wealth of the three different strategies based on R-, Sharpe, and STARR ratios is respectively 1.76, 1.07, and 0.91. Therefore, as we expected, we find that the strategy based on the maximization of the STARR ratio provides the most conservative behavior while the strategy based on the R-ratio permits the final wealth to increase much more than the others. Thus, even in periods with big market crises (Asian and Russian crises 1997–2000 and crises during and subsequent to

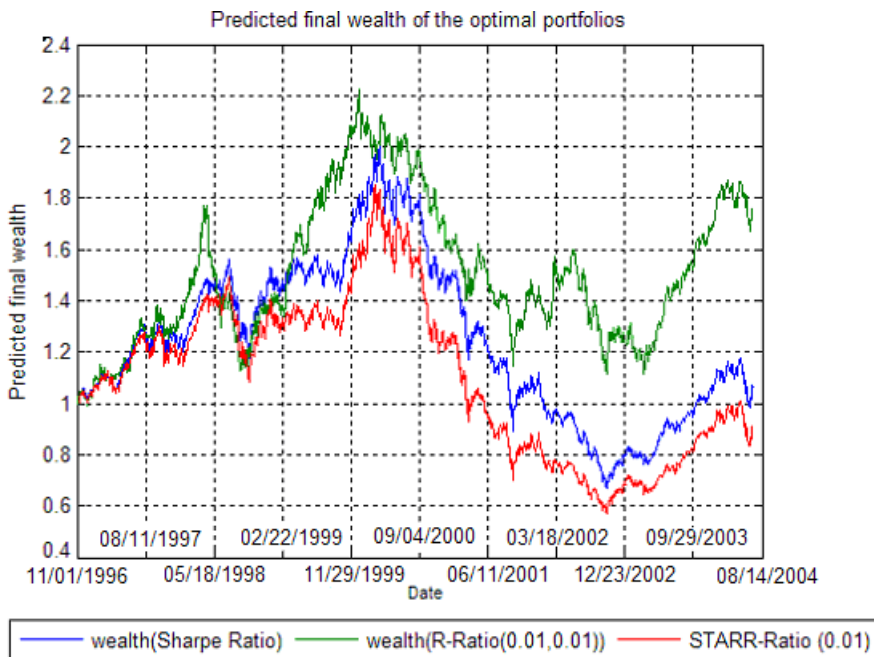


Fig. 8. *Ex-post* comparison of the optimal dynamic strategies (without transaction costs) obtained by maximizing daily the R-, STARR, and Sharpe ratio.

September 11, 2001), the “aggressive-coherent” behavior of the non-satiable investors who are neither risk averse nor risk loving (those that maximize the R-ratio) permits the final wealth to increase much more than by adopting “conservative strategies”.

6. Non-Linearity and Distributional Modeling

As proposed by Balzer [6], we could consider other desirable properties of a risk measure, such as non-linearity and distributional modeling of risk. Next, we briefly summarize these properties.

6.1. Non-linearity of risk

According to Balzer’s definition, the non-linearity of risk is related to an investor’s attitude, which is generally considered non-linear with respect to different sources of risk (see [6]). For example, suppose that investors employ the expected shortfall as a risk measure. Let us consider two investments, the returns of which have equal expected shortfall $ETL_{5\%}$. For example, suppose that for the first investment the future losses below the VaR at 95% confidence level are 2% with a probability 0.025 and 1% with a probability 0.025. For the second investment, the future losses are 30% with a probability 0.002 and 0.3125% with a probability 0.048. Thus, the extreme losses from the first investment are medium in magnitude and are equally probable, and the losses from the second investment are more dispersed — there is a very large loss with a small probability and a very small loss with a much higher probability. Considering that the two investments have the same expected shortfall $ETL_{5\%}$, investors that use this risk measure will be indifferent between the two investments. However, evidence reported by Olsen [20] suggests that most investors perceive a low probability of a large loss to be far more risky than a high probability of a small loss. Therefore, investors perceive risk to be non-linear. This simple counter-example shows that a unique risk measure (even if coherent) cannot be sufficient to describe investors’ behavioral tendencies.

6.2. Asymptotic distributional modeling

The previous example also suggests that a risk measure does not summarize all the information relevant to the risk. In order to overcome this incompleteness of risk measures, further parameters that characterize the investor’s attitude towards risk are used and analyzed, such as skewness and kurtosis. Typically, a measure of an investment’s skewness is introduced to take into account the investors’ preferences. Generally skewness is parameterized with a non-linear measure that partially overcomes and solves the empirical misspecification of some linear factor models. On the other hand, Ortobelli *et al.* [22] have shown that skewness and further distributional parameters could have an important impact on portfolio choices. However, that approach was based on a classical definition of skewness and in many situations

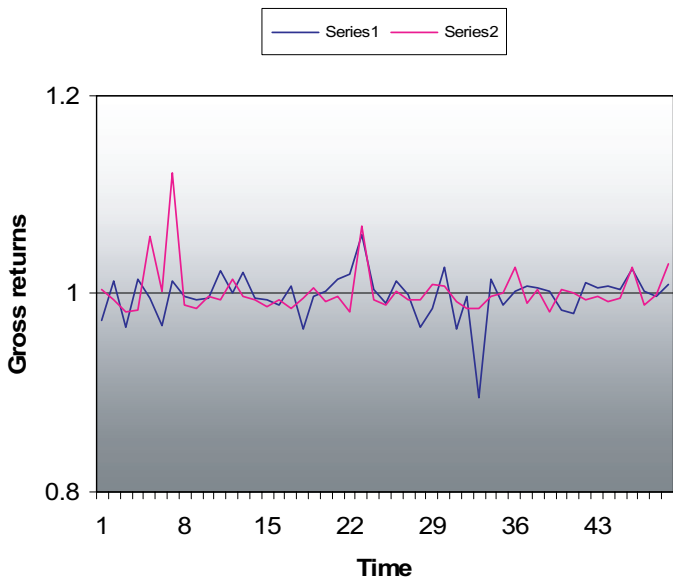


Fig. 9. Comparison among simulated data with stable asymmetric distributions.

the data behavior suggests utilizing precise assumptions on the asymmetric return distribution based on a different definition of skewness. For example, let us consider the evolution of a unit of random wealth, considering two admissible gross returns F and G (see Fig. 9).

The gross return $F \approx S_{\alpha}(\gamma, \beta_1, \delta)$ (series 1) is drawn from an α -stable distribution with index of stability $\alpha = 1.5$, dispersion 0.008, skewness parameter $\beta_1 = -1$, and daily mean equal to $\delta = 1.0001$ (see [26, 34] about stable modeling of asset returns). The gross return $G \approx S_{\alpha}(\gamma, \beta_2, \delta)$ (series 2) is α -stable distributed with the same parameters except for skewness, that is $\beta_2 = 1$. From Fig. 9, intuition suggests that gross return G is preferable to F even if the two gross returns present the same mean, dispersion, and index of stability (these three parameters could be used to characterize the behavior of symmetric returns).

This example makes clear that (1) a reward measure and a risk measure are still insufficient to describe the complexity of investor's choices; and (2) investors generally prefer positive skewness. In addition, many other distributional parameters could have an important impact in investor choices.

In order to consider the best approximation of historical return series, many statistical studies have emphasized the advantage of an asymptotic approximation (see [26]). In particular, stable modeling of financial variables permits the correct identification of investor behavior. It is well known that daily returns r often have distributions whose tails are significantly heavier than the Gaussian law, that is, for large x

$$P(|r| > x) \approx x^{-\alpha} L(x), \tag{6.1}$$

where $0 < \alpha < 2$ and $L(x)$ is a slowly varying function at infinity. This tail condition implies that the returns are in the domain of attraction of a stable law. That is, given a sequence $\{r^{(i)}\}_{i \in N}$ of independent and identically distributed (i.i.d.) observations on r , then, there exist a sequence of positive real values $\{d_i\}_{i \in N}$ and a sequence of real values $\{a_i\}_{i \in N}$ such that, as $n \rightarrow +\infty$

$$\frac{1}{d_n} \sum_{i=1}^n r^{(i)} + a_n \xrightarrow{d} X \quad (6.2)$$

where “ \xrightarrow{d} ” points out the convergence in the distribution, $X \approx S_\alpha(\gamma, \beta, \delta)$ is an α -stable random variable. This convergence result is a consequence of the Stable Central Limit Theorem (SCLT) for normalized sums of i.i.d. random variables (see [26, 34]) and it is the main justification of stable modeling in finance and econometrics. In particular, the SCLT permits one to characterize the skewness and kurtosis of investment returns in a statistically proper way. Moreover, using the maximum likelihood method to estimate the stable parameters, we could also obtain appropriate confidence intervals of these parameters. In addition, if $X \approx S_\alpha(\gamma, \beta, 0)$ is a stable centered distribution, then, when x tends to infinity,

$$P(\pm X > x) \approx \begin{cases} \exp\left(-(\alpha-1)\left(\frac{x}{\alpha\gamma B_\alpha}\right)^{\alpha/(\alpha-1)}\right) \frac{(\frac{x}{\alpha\gamma B_\alpha})^{-\alpha/(2\alpha-2)}}{\sqrt{2\pi\alpha(\alpha-1)}} & \text{if } (\beta = \mp 1 \wedge \alpha \in (1, 2)) \\ \exp(-0.5((\pi/2\gamma)x - 1) - \exp((\pi/2\gamma)x - 1))/\sqrt{2\pi} & \text{if } (\beta = \mp 1 \wedge \alpha = 1) \\ C_\alpha(1 \pm \beta)\gamma^\alpha x^{-\alpha} & \text{otherwise} \end{cases}, \quad (6.3)$$

where $B_\alpha = (\cos \frac{\pi}{2}(2-\alpha))^{-1/\alpha}$ and $C_\alpha = \frac{\Gamma(\alpha)}{\pi} \sin(\frac{\pi\alpha}{2})$. Thus, returning to our previous example, we observe that for large positive x :

$$P(\text{return}(G) < -x) \leq P(\text{return}(F) < -x) \quad \text{and} \\ P(\text{return}(G) > x) \geq P(\text{return}(F) > x).$$

This relation provides theoretical justification of what intuition suggested in the previous example. As a matter of fact, gross return G presents a lower probability of big losses and a larger probability of great earnings than gross return F , even if the two alternative returns present the same dispersion, mean, and index of stability. This is another typical way to consider multi-parameter dependence of portfolio choices. Thus, modern portfolio theory has to answer many more questions regarding risk measures.

7. Concluding Remarks

Using several examples, in this paper we have described some intuitive characteristics of risk measures. Although, we could not claim that this is an exhaustive

analysis, the principal focus of the paper is identifying the intrinsic properties of risk that all investors have to take into account. Thus, even if in order to study some problems we propose some risk measures/ratios and utility functionals, we do not claim that these measures or functionals could solve all the problems related to the theory of portfolio choice. However, we believe that they could serve to deal with those particular problems. These basic considerations suggest that we cannot use a unique axiomatic definition of risk that has to be valid for any measurement. As a matter of fact, celebrated coherent risk measures, such as shortfall risk, do not take into account the non-linearity and multi-dimensionality nature of risk. In contrast, it is reasonable to expect that a single measure cannot be relied upon to characterize uniquely investor choices. Since a unique risk measure cannot be specified, we propose to model risk of return series taking into account their specific characteristics which could depend on the period of analysis, the country/sector, or economic conditions. In particular, the examples presented justify the following desirable features of investment risk:

- Asymmetry, relativity, and multidimensionality of risk.
- Inter-temporal dependence.
- Non-linearity and distributional modeling.

In addition, we discuss how any investment choice has to take into account:

- Stochastic dominance order, correlation, and diversification among different sources of risk.
- Differences and common features between risk and uncertainty and their proper use.
- Impact of downside risk, aggregated risk, reward measures, proper risk measures, and risk aversion in investor's choices.
- Impact of dynamic strategies, transaction costs, and computational complexity.

In summary, some aspects of risk are embedded in the risk measure and other aspects are incorporated through proper modeling of the asset return distributions. For example, the axioms behind the coherent risk measures demonstrate how some characteristics of risk can be implanted in the definition of a risk measure, while other aspects such as diversification, are possible to account for only if the return distribution of the assets is modeled in a realistic way. Therefore it is the combination of a risk measure and a stochastic model that investment decisions should be based on.

In addition, we considered the benchmark tracking problem and modified the deviation measure in the objective function to be a probability distance. The new problem is more flexible and contains the traditional problem as a special example and is a significant extension. We also mention a parallel between reward-risk ratios and probability metrics, suggesting that such relationships should be studied in future research as they might imply interesting connections between classes of risk

measures or suggest an approach to select an ideal risk measure or performance ratio for a given portfolio choice problem.

Acknowledgments

Ortobelli's research has been partially supported under MURST 60% 2005, 2006, 2007. Rachev's research was supported by grants from the Division of Mathematical, Life and Physical Sciences, College of Letters and Science, University of California, Santa Barbara and the Deutschen Forschungsgemeinschaft.

References

- [1] C. Acerbi, Spectral measures of risk: A coherent representation of subjective risk aversion, *Journal of Banking and Finance* **26** (2002) 1505–1518.
- [2] P. Artzner, F. Delbaen, J.-M. Heath and D. Eber, Coherent measures of risk, *Mathematical Finance* **9** (1999) 203–228.
- [3] V. S. Bawa, Admissible portfolio for all individuals, *Journal of Finance* **31** (1976) 1169–1183.
- [4] V. S. Bawa, Safety-first stochastic dominance and optimal portfolio choice, *Journal of Financial and Quantitative Analysis* **13** (1978) 255–271.
- [5] L. A. Balzer, How to measure risk, Conference on Investment Performance, AIC Conferences, Sydney, Australia (1990).
- [6] L. A. Balzer, Investment risk: A unified approach to upside and downside returns, in F. Sortino and S. Satchell (Eds.), *Managing Downside Risk in Financial Markets: Theory Practice and Implementation* (Butterworth-Heinemann, Oxford, 2001), pp. 103–156.
- [7] A. Biglova, S. Ortobelli, S. T. Rachev and S. Stoyanov, Different approaches to risk estimation in portfolio theory, *Journal of Portfolio Management* **31** (2004) 103–112.
- [8] R. Bordley and M. LiCalzi, Decision analysis using targets instead of utility functions, *Decisions in Economics and Finance* **23** (2000) 53–74.
- [9] E. Castagnoli, Benchmarking under risk, Atti convegno AMASES 2004, Modena.
- [10] E. Castagnoli and M. LiCalzi, Expected utility without utility, *Theory and Decision* **41** (1996) 281–301.
- [11] E. Castagnoli and M. LiCalzi, Non-expected utility theories and benchmarking under risk, *SZIGMA* **29** (1999) 199–211.
- [12] E. De Giorgi, Reward-risk portfolio selection and stochastic dominance, *Journal of Banking and Finance* **29**(4) (2005) 895–926.
- [13] M. Friedman and L. J. Savage, The utility analysis of choices involving risk, *Journal of Political Economy* **56** (1948) 279–304.
- [14] R. Giacometti and S. Ortobelli, A comparison between dispersion measures for the asset allocation problem, Technical Report 7, University of Bergamo (2001).
- [15] G. Hanoch and H. Levy, The efficiency analysis of choices involving risk, *Review of Economic Studies* **36** (1969) 35–46.
- [16] G. A. Holton, Defining risk, *Financial Analysts Journal* **60** (2004) 19–25.
- [17] E. Karni, *Decision Making Under Uncertainty: The Case of State-Dependent Preferences* (Harvard University Press, Cambridge, 1985).
- [18] F. H. Knight, *Risk, Uncertainty and Profit* (Hart, Schaffner and Marx, New York, 1921).
- [19] J. Longerstaeey and P. Zangari, *RiskMetrics — Technical Document*, 4th Ed. (J.P. Morgan, New York, 1996).

- [20] R. A. Olsen, Investment risk: The experts' perspective, *Financial Analysts Journal* **53**(2) (1997) 62–66.
- [21] S. Ortobelli, The classification of parametric choices under uncertainty: Analysis of the portfolio choice problem, *Theory and Decision* **51** (2001) 297–327.
- [22] S. Ortobelli, S. T. Rachev, S. Stoyanov, F. J. Fabozzi and A. Biglova, The proper use of risk measures in the portfolio theory, *International Journal of Theoretical and Applied Finance* **8** (2005) 1107–1133.
- [23] S. Ortobelli, S. T. Rachev, H. Shalit and F. J. Fabozzi, Risk probability functionals and probability metrics applied to portfolio theory, Working Paper, Department of Probability and Applied Statistics, University of California, Santa Barbara, USA (2006).
- [24] S. Ortobelli and S. T. Rachev, Safety first analysis and stable Paretian approach, *Mathematical and Computer Modelling* **34** (2001) 1037–1072.
- [25] G. Pflug, Some remarks on the value at risk and the conditional value at risk, in S. Uryasev (Ed.), *Probabilistic Constrained Optimization: Methodology and Application* (Kluwer, Dordrecht, 2000).
- [26] S. T. Rachev and S. Mitnik, *Stable Paretian Models in Finance* (Wiley & Sons, Chichester, 2007).
- [27] S. Rachev, D. Martin, B. Racheva-Iotova and S. Stoyanov, Stable ETL optimal portfolios and extreme risk management. Forthcoming in *Decisions in Banking and Finance* (2007).
- [28] S. Rachev, *Probability Metrics and the Stability of Stochastic Models* (Wiley & Sons, Chichester, 1991).
- [29] R. T. Rockafellar and S. Uryasev, Conditional value-at-risk for general loss distributions, *Journal of Banking and Finance* **26**(7) (2002) 1443–1471.
- [30] R. T. Rockafellar, S. Uryasev and M. Zabarankin, Generalized deviations in risk analysis, *Finance and Stochastics* **10** (2006) 51–74.
- [31] M. Rothschild and J. Stiglitz, Increasing risk: I, definition, *Journal of Economic Theory* **2** (1970) 225–243.
- [32] A. D. Roy, Safety-first and the holding of assets, *Econometrica* **20** (1952) 431–449.
- [33] R. Ryan and F. J. Fabozzi, Rethinking pension liabilities and asset allocation, *Journal of Portfolio Management* **28** (2002)(4) 7–15.
- [34] G. Samorodnitsky and M. S. Taqqu, *Stable Non Gaussian Random Processes: Stochastic Models with Infinite Variance* (Chapman and Hall, New York, 1994).
- [35] M. J. Schervish, T. Seidenfeld and J. B. Kadane, State-dependent utilities, *Journal of the American Statistical Association* **85** (1990) 840–847.
- [36] M. Shaked and G. Shanthikumar, *Stochastic Orders and Their Applications* (Academic Press Inc. Harcourt Brace & Company, New York, 1994).
- [37] W. F. Sharpe, The Sharpe ratio, *Journal of Portfolio Management* Fall (1994) 45–58.
- [38] F. A. Sortino and S. Satchell, *Managing Downside Risk in Financial Markets: Theory, Practice and Implementation* (Butterworth Heinemann, Oxford, 2001).
- [39] S. Stoyanov, S. T. Rachev and F. J. Fabozzi, Optimal financial portfolios, *Applied Mathematical Finance* **14** (2007) 402–436.
- [40] S. Stoyanov, S. T. Rachev, S. Ortobelli and F. J. Fabozzi, Relative deviation metrics and the problem of strategy replication, to appear in *Journal of Banking and Finance*.
- [41] G. Szegö, Measures of risk, *Journal of Banking and Finance* **26**(7) (2002) 1253–1272.
- [42] G. Szegö (Ed.), *Risk Measures for the 21st Century* (Wiley & Sons, Chichester, 2004).
- [43] L. G. Tesler, Safety first and hedging, *Review of Economic Studies* **23** (1955/6) 1–16.