

Desirable Properties of an Ideal Risk Measure in Portfolio Theory

S. T. Rachev^a, S. Ortobelli^{b*},
S. Stoyanov^c, F. Fabozzi^d, A. Biglova^e

^aUniversity of California, Santa Barbara and
University of Karlsruhe, Germany

^bUniversity of Bergamo, Italy

^cFinAnalytica Inc., USA

^dYale University, USA

^eUniversity of Karlsruhe, Germany

* *Corresponding author:* E-mail: sol@unibg.it. Ortobelli's research has been partially supported under MURST 60% 2004, 2005. S. Rachev's research was supported by grants from Division of Mathematical, Life and Physical Sciences, College of Letters and Science, University of California, Santa Barbara and the Deutschen Forschungsgemeinschaft.

Desirable Properties of an Ideal Risk Measure in Portfolio Theory

Abstract

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 to understand how to classify an investor's risk. The risk is an asymmetric, relative, heteroskedastic, multidimensional concept that has to take into account asymptotic behavior of returns, inter-temporal dependence, risk-time aggregation, further sources of risk, 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 the risk we propose and analyze the relationship between distributional modeling and risk measures. Similarly 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. Then we emphasize the parallels between risk measures and probability metrics underlying the computational advantage and disadvantage of different approaches.

Key words: risk aversion, portfolio choice, investment risk, reward measure, diversification.

JEL Classification: G11, G14, G15

1. INTRODUCTION

In portfolio theory, a risk measure has always been valued principally because of its capacity of ordering investor preferences. In particular, the stochastic-order theory has provided some intuitive rules consistent with expected utility theory (see, among others, Hanoch and Levy (1969), Rothshild and Stiglitz (1970), and Bawa (1976)). 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 researches have emphasized that the investors' choices are strictly dependent on the possible states of the returns (see, among others, Karni (1985)). 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 (1985), Schervisch, Seidenfeld and Kadane (1990), 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 Bordley and LiCalzi (2000), Castagnoli (2004)). 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, Castagnoli and LiCalzi (1996,1999)).

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, Roy (1952), Tesler (1955/6), and Bawa (1976, 1978) Ortobelli and Rachev (2001)). In particular, the most celebrated and used benchmark approaches are based on coherent risk measures (see Szegö (2004)). 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* (1999). However, even if a coherent risk measure “coherently” prices risk, it cannot consider exhaustively all the investment characteristics.

This paper discusses and reviews some desirable properties of a risk measure in portfolio theory, identifying some limits of previous studies on this topic. As a matter of fact, many studies proposed in portfolio theory deal with risk measures often used in the statistics literature (see Uryasev (2000), Szegö (2002), and (2004) and the references therein). However, most of the proposed measures do not take into account several investors' attitudes towards risk. In addition, if a risk measure has to be used in portfolio choice, then the main investor's interest is the consistency with his preferences. In particular, the knowledge of an investor's attitudes toward risk permits an investor to correctly employ any risk measure coherently with respect to his preferences (see, among others, Ortobelli et al (2005)). On the one hand, we do not believe that a unique risk measure could capture all aspects of an investor's preferences.

This paper distinguishes several observable financial phenomena such as the impact of aggregated risk, temporal horizon, propagation effect, risk aversion, transaction costs, and heteroskedasticity. In addition, it examines 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. Because we do not believe that a single risk measure can take into account all these characteristics, we propose some different ways to study the aspects of risk. For this reason, in Section 2 we summarize some of the basic characteristics of risk emphasizing the differences between measures of uncertainty and measures of risk. Section 3 shows how some aspects of diversification motivate the use of reward-risk ratios. Section 4 describes the parallels between risk measures and probability metrics, while Sections 5, 6 and 7 analyze several aspects of risk which impact an investors' choices. We summarize our principal findings in Section 8. .

2. BASIC CHARACTERISTICS OF INVESTMENT RISK: UNCERTAINTY AND RISK

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 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 (1990, 2001), Sortino and Satchell (2001), among others, have proposed that investment risk might be measured by a functional of the difference between the investment return and a specified benchmark. 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 to 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 Castagnoli and LiCalzi (1996,1999)). 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 U.S. pension funds (see Ryan and Fabozzi (2002)). From this simple discussion, one must recognize that risk is a *relative* (to a given benchmark), *asymmetric*, and *multidimensional* concept. In addition, Rockafellar *et al* (2005) and Ortobelli *et al* (2005) have emphasized that risk cannot be assessed by measuring only the *uncertainty* of investments.

There exist many examples in the portfolio selection literature illustrating that the standard deviation cannot always be utilized as a measure of risk because it is a measure of uncertainty. However, the two notions of uncertainty and risk are also related. Generally, measures of uncertainty, also known as *deviation measures*, can be introduced axiomatically (see Rockafellar *et al* (2005)). We call a deviation measure any positive functional D defined on the space of random variables with finite variance satisfying the following properties:

Dev 1: $D(X + C) = D(X)$ for all X and constants C

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

Dev 3: $D(X + Y) \leq D(X) + D(Y)$ for all X and Y

Dev 4: $D(X) \geq 0$ for all X , with $D(X) > 0$ for non-constant X

According to properties Dev 1 and Dev 4, the deviation measure D depends only on the centered random variable $X - EX$ and is equal to zero only if $X - EX = 0$. Therefore we can say that the functional D measures the degree of uncertainty in X . Particular examples include the standard deviation, the mean absolute deviation and so on. Further, the family of the 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.

Attempts to quantify risk have led to the notion of a *risk measure*. It 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 view of that class of investors (see Pflug (2000)). 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.

A systematic approach towards risk measures has been undertaken in Artzner *et al* (1999) 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. R3 implies *sub-additivity* and the combination of 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 additional 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 measure of risk is the conditional value-at-risk (CVaR), also known as expected shortfall or expected tail loss (ETL). It is defined as

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

where $VaR_{\alpha}(X) = -\inf_x \{x | P(X \leq x) \geq \alpha\}$ is the value-at-risk (VaR) of the random variable X and $ETL_{\alpha}(X) = E(X | X \leq VaR_{\alpha}(X))$ when we assume a continuous distribution for the distribution of X . 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. The ETL can be interpreted as the average loss beyond value-at-risk.

There is a relationship between the deviation measures and the coherent risk measures, see Rockafellar *et al* (2005). In particular, in a mean-risk world most of these measures are *equivalent* for risk averse investors (see Ortobelli *et al* (2005)). On the other hand, Biglova, *et al* (2004) 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.

3. TEMPORAL DEPENDENCE, DIVERSIFICATION AND REWARD-RISK ANALYSIS

Let us consider the following example. Figure 1 shows the S&P 500 daily return series from January 4, 1995 to January 30, 1998. As we can see, the dispersion around the mean

changes sensibly, in particular during the last period of our observations when the Asian market crisis began. These oscillations tell us that the process is *mean reverting* and that the *dispersion* changes over the time. Hence, in some periods there are big oscillations around zero and in other periods the oscillations are smaller. 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.

[INSERT HERE FIGURE 1]

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 the volatility* effect; that is, each observation influences the following 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. Typically this behavior is captured by ARMA-GARCH-type models.

One of the simplest models proposed in the literature is the exponentially weighted moving average model (EWMA) and considers exponential weights (see Longerstaey and Zangari (1996)). 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 Longerstaey and Zangari (1996)). 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 risk heteroskedasticity can be modeled assuming that $\sigma_{t+1/t} = \lambda\sigma_{t/t-1} + (1 - \lambda)f(r_t)$.

[INSERT HERE FIGURES 2 AND 3]

Figure 2 shows the S&P 500 daily return series from September 8, 1997 to January 30, 1998. Note that a positive return almost always follows a loss. The wavy behavior of returns also has a *propagation effect* on the other markets. This can be seen in Figure 3 which describes the DAX 30 daily return series (valued in US dollars) during the same period as the S&P 500 series. When we observe the highest peaks in the S&P 500 returns, there is an analogous peak in the DAX 30 series. This propagation effect is known as *cointegration of returns series* and is a consequence of the globalization because the risk of a country/sector is linked to the risk of the other countries/sectors.

Therefore, it could be important to limit the propagation effect by *diversifying risk*. As a matter of fact, there is considerable evidence that diversification, if opportunely modeled, 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 returns of the portfolio items. As a matter of fact, investors want to diversify the portfolio in order to minimize the risk. Therefore, the diversification makes sense only when there exists 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

$$\exists a \in (0,1) : \sigma_{aX+(1-a)Y} \leq \min(\sigma_X; \sigma_Y; a\sigma_X + (1-a)\sigma_Y), \quad (2)$$

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 of investors a diversification will appear convenient in a mean-risk plane if there exist some values $a \in (0,1)$, such that for a given 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 the ratio between the mean and the risk measure. In this case, there exists $a \in (0,1)$ such that

$$\frac{\partial \sigma_{aX+(1-a)Y}}{\partial a} E(aX + (1-a)Y) = \sigma_{aX+(1-a)Y} E(X - Y). \quad (3)$$

[INSERT HERE FIGURE 4]

Figure 4 provides an example where weak diversification holds but strong diversification does not. 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 $Y = bX + c$ (for opportune b and 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 (as the standard deviation) does not present weak diversification when two 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 the convexity holds but weak diversification does not, is shown in Figure 5.

These different definitions of diversification underline the importance of taking into account an investors's reward and not only risk in the portfolio selection problem. In particular, De Giorgi (2005) introduced the first axiomatic definition of reward measures. 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 interests satisfying the following intuitive property:

[INSERT HERE FIGURE 5]

a) if any risk averse non-satiable investor prefers X to Y then $v(X) \geq v(Y)$, i.e. the functional v is isotonic with respect to second stochastic order according to the definition given by De Giorgi (2005).

Considering that in portfolio theory we need only to 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 underline that when the wealth under risk is multiplied by a positive factor, then reward also must grow with the same proportionality. Thus in some cases we require that the functional v be positively homogeneous and $v(aX) = a v(X)$, for all X and all $a > 0$.

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 :

$$\begin{aligned} \min_w & \rho(w^T r - r_b) \\ \text{s.t.} & \\ & v(w^T r - r_b) \geq R \\ & l \leq Aw \leq u \end{aligned} \quad (4)$$

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; that is, this portfolio is a solution to the optimal ratio problem

$$\begin{aligned} \max_w & \frac{v(w^T r - r_b)}{\rho(w^T r - r_b)} \\ \text{s.t.} & \\ & l \leq Aw \leq u \end{aligned} \quad (5)$$

In both problems (4) and (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 Sharpe (1994)). 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 Rachev et al (2003) and Biglova *et al* (2004) and the references therein):

- the STARR ratio: $E(w^T r - r_b)/ETL_\alpha(w^T r - r_b)$
- the 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 Rachev, Mittnik (2000)).
- the 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.
- the Sortino-Satchell ratio: $E(w^T r - r_b) / (E(\max(t - w^T r, 0)^\gamma))^{1/\gamma}$, $\gamma \geq 1$
- the Rachev ratio (R-ratio): $ETL_\alpha(r_b - w^T r)/ETL_\beta(w^T r - r_b)$
- the 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 $\gamma^* = \min(1, 1/\gamma)$.
- the VaR ratio: $VaR_\alpha(r_b - w^T r)/VaR_\beta(w^T r - r_b)$

In the R- and GR-ratio, the reward functional is non-linear. The R-ratio can be interpreted as the ratio between the average (active) profit exceeding a certain threshold and the average (active) loss below a certain level. The R-ratio and the GR-ratio have been proposed since there is empirical evidence that they are more appropriate for investment decisions in the case of heavy-tailed returns (see Biglova *et al* (2004)).

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. We can consider the following cases (for more details, see Stoyanov *et al* (2005)):

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:

$$X = \left\{ \begin{array}{l} \rho(w^T r - r_b) - t v(w^T r - r_b) \leq 0 \\ l \leq Aw \leq u \end{array} \right.$$

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 (2005). 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

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

are the optimal portfolios of the ratio problem (5).

Case 2: If, in addition to the conditions in Case 1, both functions are positively homogeneous, then the optimal ratio problem reduce to a convex programming problem. An example of an equivalent convex problem to (5) is

$$\begin{array}{ll} \min_{(x,t)} & \rho(x^T r - tr_b) \\ \text{s.t.} & \\ & v(x^T r - tr_b) \geq 1 \\ & lt \leq Ax \leq ut \\ & t \geq 0 \end{array} \quad (6)$$

where t is an additional variable. If (x_o, t_o) is a solution to (6), then $w_o = x_o/t_o$ is a solution to problem (5). There are other connections between problems (5) and (6). Let ρ_o be the value of the objective at the optimal point (x_o, t_o) in problem (6). Then $1/\rho_o$ is the value of the optimal ratio, i.e. the optimal value of the objective of problem (5). Moreover $1/t_o$ is the

reward of the optimal portfolio and ρ_o/t_o is the risk of the optimal portfolio if the reward constraint is satisfied as equality at the optimal solution.

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 (examples: the Sharpe ratio, the Stable ratio). An example of an equivalent problem to the Sharpe ratio problem is

$$\begin{aligned}
 & \min_{(x,t)} (-t, x) \Sigma_1 (-t, x)^T \\
 & s.t. \\
 & x^T Er - tEr_b = 1 \\
 & lt \leq Ax \leq ut \\
 & t \geq 0
 \end{aligned} \tag{7}$$

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 items returns, σ_b^2 is the variance of the benchmark portfolio returns, $\sigma_{br} = (\text{cov}(r_b, r_1), \text{cov}(r_b, r_2), \dots, \text{cov}(r_b, r_n))$ is a vector of covariances between the benchmark portfolio returns and the returns of the main portfolio items. Again, if (x_o, t_o) is a solution to (7), then $w_o = x_o/t_o$ is a solution to the version of problem (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 (7) and (5) are the same as the ones given in Case 2 above as problem (7) is just a particular version of problem (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 (6) by incorporating the linearizations:

$$\begin{aligned}
 & \min_{(x,t,d,\theta)} \theta + \frac{1}{N\alpha} \sum_{k=1}^N d_k \\
 & s.t. \\
 & x^T E r - t E r_b = 1 \\
 & -x^T r^k + t r_b^k - \theta \leq d_k, k = 1, 2, \dots, N \\
 & l t \leq A x \leq u t \\
 & t \geq 0, d_k \geq 0, k = 1, 2, \dots, N
 \end{aligned} \tag{8}$$

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 (8) and (5) are the same as the ones in Case 2. One should bear in mind that in the objective of problem (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 the linearization, see Rockafellar and Uryasev (2002).

Clearly, as we have noted, the dimension of the optimization problem increases as we simplify the problem structure. Certainly, if this trade-off suggests that the computational burden actually increases, 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 for all feasible portfolios. If there exists a feasible portfolio with a negative 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

$$v(w^T r - r_b) - \lambda \rho(w^T r - r_b) \quad (9)$$

where $\lambda > 0$ is a risk aversion parameter.

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 $\lambda = 1$, 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 (9) are deviation measures, i.e. they satisfy all axioms Dev 1 through Dev 4 (for more details and a proof, see Rockafellar *et al* (2005)). 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 .

Certainly an optimization problem in which we have a linearized reward-risk ratio in the objective with a pre-selected value for λ is equivalent to problem (4) with a suitable choice of the limit R . Objectives of type (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.

4. 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

- a) (identity) $\Pr(X = Y) = 1 \Leftrightarrow \mu(X, Y) = 0$
- b) (symmetry) $\mu(X, Y) = \mu(Y, X)$
- c) (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 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 \mathbb{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

- $\mu(X, Y) = \left[\int_R |F_X(t) - F_Y(t)|^p dt \right]^{1/p}$, $p \geq 1$, 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 (1991).

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

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

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 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.

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 $\gamma = \delta$

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

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.

5. DOWNSIDE RISK 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 international indexes (DAX30, DAX100, CAC 40, FTSE all share, FTSE 100, FTSE Actuaries 350, Nikkei 300 weighted average, Nikkei 300 simple average, Nikkei 500, Corn no.2, Coffee Brazilian, Dow Jones Industrial, and S&P 500) for the period 1/4/1995-7/18/1997 all converted into U.S. dollars. 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 the presence of a riskless asset with a daily return of $r_0 = 0.0002$ (six-month return $r_0 = 0.024$).

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 Rachev ratio

$$ETL_{30\%}(r_0 - w^T r) / ETL_{10\%}(w^T r - r_0) \quad (10)$$

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* (see, among others, Sortino and Satchell (2001) and Biglova *et al* (2004)). As a matter of fact, the risk measure expected shortfall $ETL_{10\%}(w^T r - r_0)$ considers the portfolio downside risk, and the reward measure $ETL_{30\%}(r_0 - w^T r)$ takes into consideration in a different way 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 six months of returns (120 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 Rachev ratio given by (10).

Figure 6 proposes an *ex-post* comparison during the period 2/3/1997-7/18/1997 (120 days) of the final wealth that an investor could obtain if on 1/31/1997 he/she invests in one of the four portfolios. 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 maximize the Rachev Ratio given by (10)), while Series monthminvar and monthmaxRR represent respectively the final wealth graph of the two six-month optimal 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 monthminvar) 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 six-month 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.

[INSERT HERE FIGURE 6]

[INSERT HERE TABLE 1]

This behavior is confirmed by the results reported in Table 1, 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) present a more conservative and risk-averse position than portfolios that maximize the Rachev ratio given by (10) (respectively portfolios 2 and 4). In contrast, the *ex-post* analysis as expected shows the differences between the first two and the last two portfolios. Therefore, Figure 6 and Table 1 not only describe strategies derived from different risk-averse positions, but they emphasize the importance of

aggregated risk in investors' preferences. As a matter of fact, the risk portfolio on one day is generally different from the risk portfolio based on six months, and the forecasting analysis has to take into account the *aggregated risk* possibly considering also its *heteroscedasticity*.

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 takes into account downside risk and offer 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.

6. DYNAMIC STRATEGIES, TRANSACTION COSTS, AND COMPUTATIONAL COMPLEXITY

Let us consider the portfolio selection among a risk-free asset with monthly return $r_0 = 0.004$ and the same 13 international indexes used in the previous example based on the period 1/4/1995-1/30/1998 all converted into U.S. dollars. We assume that investors recalibrate 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 (that we assume to be equal to 1 in data 2/3/1997). 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, it solves the problem

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

$$\text{where } X = \begin{cases} x_{(k)}'r - r_0 & \text{without transaction costs} \\ x_{(k)}'r - r_0 - 0.005 \sum_{i=1}^{13} \left| x_i^{(k)} - \frac{x_i^{(k-1)}(1+r_i^{(k-1)})}{\sum_{i=1}^n x_i^{(k-1)}(1+r_i^{(k-1)})} \right| & \text{with tr. costs} \end{cases} \quad \text{and } r^{(k)} = [r_1^{(k)}, \dots, r_{13}^{(k)}]' \text{ is}$$

the k -th ex-post monthly observation of return vector r . In addition, we assume that the variance follows the exponential weighted model $\sigma_{t+1/t} = \lambda \sigma_{t/t-1} + (1-\lambda)f(X_t)$ with $\lambda = 0.94$ (as suggested by RiskMetrics approach in Longerstaey, and Zangari (1996)). Then, the investor's wealth after the k months is given by

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

[INSERT HERE FIGURE 7]

Figure 7 shows the final wealth process with and without transaction costs during the period 2/3/1997-1/30/98 when the Asian market crisis began. After 12 recalibrations, the difference between the final wealth obtained without transaction costs (series 1) and with transaction costs (series 2) was about 2%. Therefore, it does not seem that the transaction costs have a significant impact in portfolio choice. However, the portfolio composition in some cases changes significantly and we have to expect that the impact of transaction costs still depends on the optimization problem we are solving. In addition, when the investor's attitude toward risk is riskier, transaction costs have generally a more relevant impact in the choices.

As we have emphasized in Section 3, the *computational complexity* of the problem is another important aspect. In particular, this is the case if 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.

As proposed by Balzer (2001), 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.

7. NON-LINEARITY AND DISTRIBUTIONAL MODELING

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 Balzer (2001)). 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 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, one arises from a high probability of some small shortfalls (the first investment), and the other, a low probability of a very large shortfall (the second investment). Considering that the two investments have the same expected shortfall $ETL_{5\%}$, investors that assume this risk measure will be indifferent between the two investments. However, evidence reported by Olsen (1997) 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.

The previous example also underline 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 the kurtosis. Typically, a measure of an investment's skewness is introduced to take into account the an investor's preferences. Generally skewness is parameterized with a non-linear measure that partially overcomes and solves the empirical misspecification of some linear factor models. For example, let us consider the evolution of a unit of random wealth, considering two admissible gross returns F and G (see Figure 8).

[INSERT HERE FIGURE 8]

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 1.0001 (see Rachev and Mittnik (2000) and Samorodnitsky and Taqqu (1994) 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 Figure 8, 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 (see, among others, Ortobelli *et al* (2005)) and (2) investors generally prefer positive skewness. In addition, many other distributional parameters could have an important impact in the 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 Rachev and Mittnik (2000)). In particular, stable modeling of financial variables permit the correct identification of investor behavior. It is well known that daily asset returns r have distributions whose tails are heavier than the Gaussian law, that is, for large x

$$P(|r| > x) \sim x^{-\alpha} L(x) \quad (11)$$

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

$\{r^{(i)}\}_{i \in \mathbb{N}}$ of independent and identically distributed (i.i.d.) observations on r , then, there exist a sequence of positive real values $\{d_i\}_{i \in \mathbb{N}}$ and a sequence of real values $\{a_i\}_{i \in \mathbb{N}}$ such that,

as $n \rightarrow +\infty$

$$\frac{1}{d_n} \sum_{i=1}^n r^{(i)} + a_n \xrightarrow{d} X \quad (12)$$

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 Samorodnitsky and Taqqu (1994) and Rachev and Mittnik (2000)) and it is the main justification of stable modeling in finance and econometrics. In particular, 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 $X \approx S_\alpha(\gamma, \beta, \delta)$ is a stable standardized distribution, then, when x tends to infinity,

$$P(\pm X > x) \sim C_\alpha (1 \pm \beta) \gamma^\alpha x^{-\alpha} \quad (13)$$

where $C_\alpha = \frac{\Gamma(\alpha)}{\pi} \sin\left(\frac{\pi\alpha}{2}\right)$. Thus, returning back to our previous example, we observe that for large positive x :

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

This relation provides theoretical justification of what intuition suggested in the previous example. As a matter of fact, gross return G presents 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.

8. 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. In particular, the examples presented justify the following desirable features of investment risk:

- Asymmetry of risk
- Relativity of risk

- Multidimensionality of risk
- Inter-temporal dependence.
- Non-linearity and distributional modeling

In addition, we demonstrate 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
- 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 assets returns distribution. 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 be better studied in future research as they might imply interesting connections between classes of risk measures or propose an approach to select an ideal risk measure or performance ratio for a given portfolio choice problem.

REFERENCES

- Artzner, P., Delbaen, F., Heath, J-M., Eber, D., 1999. Coherent measures of risk. *Mathematical Finance* 9, 203-228.
- Bawa, V. S., 1976. Admissible portfolio for all individuals. *Journal of Finance* 31, 1169-1183.
- Bawa, V. S., 1978. Safety-first stochastic dominance and optimal portfolio choice. *Journal of Financial and Quantitative Analysis*, 255-271.
- Balzer, L.A., 1990. How to measure risk. Conference on Investment Performance, AIC Conferences, Sydney, Australia.
- Balzer, L.A., 2001. Investment risk: a unified approach to upside and downside returns, in: Sortino F. and Satchell (Eds.) *Managing downside risk in financial markets: theory practice and implementation*, Butterworth-Heinemann, Oxford, pp. 103-156.
- Biglova, A., Ortobelli, S., Rachev, S., Stoyanov, S., 2004. Different Approaches to Risk Estimation in Portfolio Theory. *Journal of Portfolio Management* 31, 103-112.
- Bordley, R., LiCalzi, M., 2000. Decision analysis using targets instead of utility functions. *Decisions in Economics and Finance* 23, 53-74.
- Castagnoli, E., 2004. Benchmarking under risk. *Atti convegno AMASES 2004*, Modena.
- Castagnoli, E., LiCalzi, M., 1996. Expected utility without utility. *Theory and Decision* 41, 281-301.
- Castagnoli, E., LiCalzi, M., 1999. Non-expected utility theories and benchmarking under risk. *SZIGMA* 29, 199-211.
- De Giorgi, E., 2005. Reward-risk portfolio selection and stochastic dominance. *Journal of Banking & Finance* 29 (4), 895-926.
- Hanoch, G., Levy, H., 1969. The efficiency analysis of choices involving risk. *Review of Economic Studies* 36, 35-46.
- Karni, E., 1985. *Decision making under uncertainty: the case of state-dependent preferences*. Harvard University Press, Cambridge.
- Longerstaey J., Zangari, P., 1996. *RiskMetrics - Technical Document*. J.P. Morgan, Fourth edition, New York.
- Olsen, R.A., 1997. Investment risk: the experts' perspective. *Financial Analysts Journal*,

March/April.

- Ortobelli, S., Rachev, S., Stoyanov, S., Fabozzi, F., Biglova, A., 2005. The proper use of risk measures in the Portfolio Theory. to appear in International Journal of Theoretical and Applied Finance.
- Ortobelli, S., Rachev, S.T., 2001. Safety first analysis and stable Paretian approach. Mathematical and Computer Modelling 34, 1037-1072.
- Pflug, G., 2000. Some remarks on the value at risk and the conditional value at risk, in: Uryasev S. (Ed.) Probabilistic Constrained Optimization: Methodology and Application. Kluwer, Dordrecht.
- Rachev, S., Mittnik, S., 2000. Stable Paretian model in finance. Wiley & Sons, Chichester.
- Rachev, S., Martin, D., Siboulet, F., 2003. Phi-alpha optimal portfolios and extreme risk management. Wilmott Magazine of Finance, November, pp. 70-83.
- Rachev, S., 1991. Probability metrics and the stability of stochastic models. Wiley & Son, Chichester.
- Rockafellar, R. T., Uryasev, S., 2002. Conditional value-at-risk for general loss distributions. Journal of Banking and Finance 26 (7), 1443–1471.
- Rockafellar, R. T., Uryasev, S., Zabarankin, M., 2005. Generalized Deviations in Risk Analysis. Finance and Stochastics 9.
- Rothschild, M., Stiglitz, J., 1970. Increasing risk: I. definition. Journal of Economic Theory 2, 225-243.
- Roy, A.D., 1952. Safety-first and the holding of assets. Econometrica 20, 431-449.
- Ryan, R., Fabozzi, F., 2002. Rethinking Pension Liabilities and Asset Allocation. Journal of Portfolio Management 28 (4), 7-15.
- Samorodnitsky, G., Taqqu, M.S., 1994. Stable non Gaussian random processes: stochastic models with infinite variance. Chapman and Hall, New York
- Schervisch, M.J., Seidenfeld, T., Kadane, J.B., 1990 State-dependent utilities. Journal of the American Statistical Association 85, 840-847.
- Sharpe, W.F., 1994. The Sharpe Ratio. Journal of Portfolio Management. Fall, 45-58.
- Sortino, F.A., Satchell, S., 2001. Managing downside risk in financial markets: theory, practice and implementation. Butterworth Heinemann, Oxford.

- Stoyanov, S., Rachev, S., Fabozzi, F., 2005. Optimal financial portfolios. *Working Paper*, Department of Probability and Applied Statistics, University of California, Santa Barbara, USA
- Szegö, G., 2002. Measures of risk. *Journal of Banking & Finance* 26 (7), 1253-1272.
- Szegö, G., 2004. Risk measures for the 21st century. Wiley & Son, Chichester.
- Tesler, L.G., 1955/6. Safety first and hedging. *Review of Economic Studies* 23, 1-16.
- Uryasev, S.P., 2000. Probabilistic Constrained Optimization: Methodology and Applications. Kluwer Academic Publishers, Dordrecht.

EX-POST ANALYSIS				
	Optimal portfolios on daily data		Optimal portfolios on six months data	
	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4
final wealth	1.1754827	1.25018179	1.230095955	1.26919731
VaR_{1%}	0.013750795	0.012791431	0.014087966	0.015202632
VaR_{5%}	0.00981716	0.008574332	0.008925009	0.00946738
ETL_{1%}	0.014123377	0.01401933	0.014186846	0.01761125
ETL_{5%}	0.012115733	0.01167852	0.012251441	0.01431338
EX-ANTE ANALYSIS				
	Optimal portfolios on daily data		Optimal portfolios on six months data	
	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4
VaR_{1%}	0.011551536	0.01455467	0.013398256	0.013954627
VaR_{5%}	0.006307627	0.00848358	0.009508602	0.009709126
ETL_{1%}	0.014491413	0.015143569	0.016963572	0.017068158
ETL_{5%}	0.009442397	0.01085136	0.012251953	0.012705347

Table 1. Ex-ante and ex-post valuation of the four portfolio risk positions.

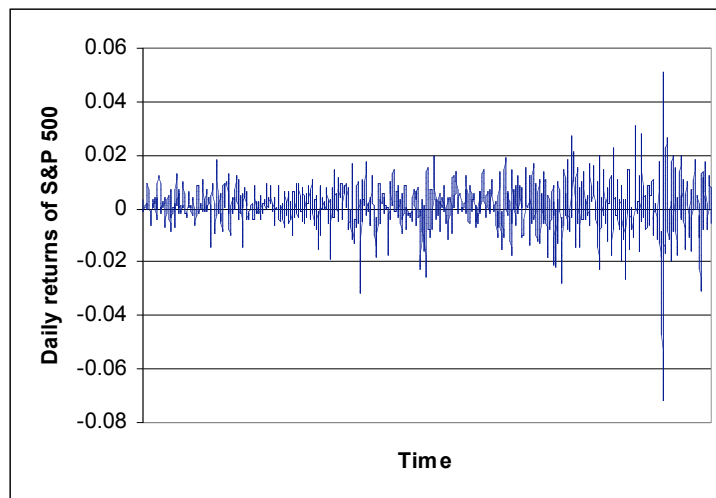


Figure 1 S&P 500 return time series from 1/4/1995 to 1/30/1998.

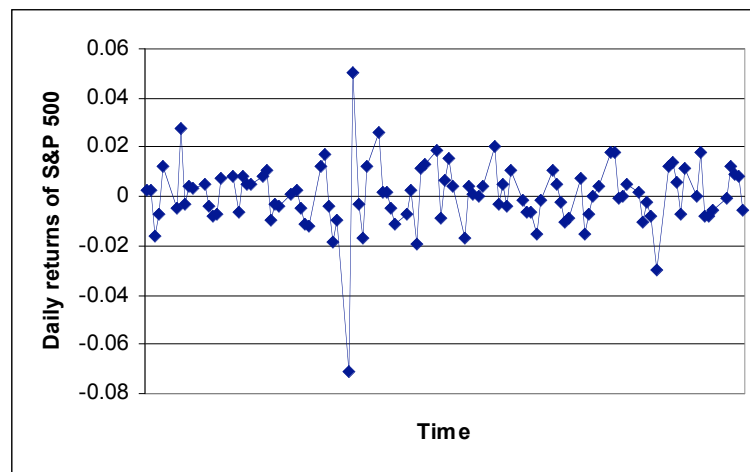


Figure 2 S&P 500 return time series from 9/8/1997 to 1/30/1998.

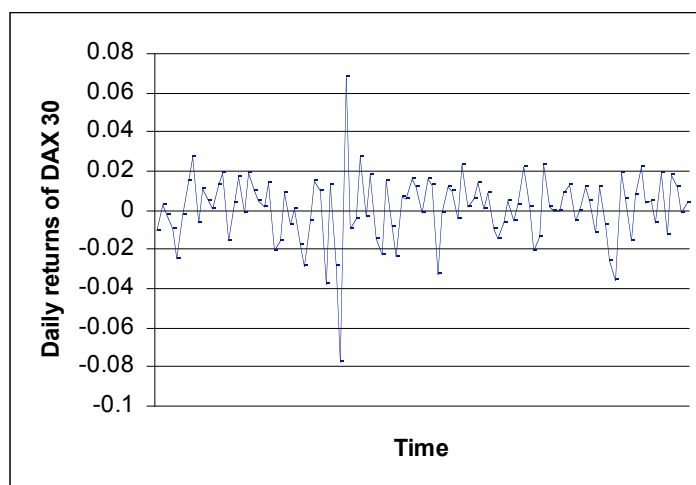


Figure 3 DAX 30 return time series from 9/8/1997 to 1/30/1998.

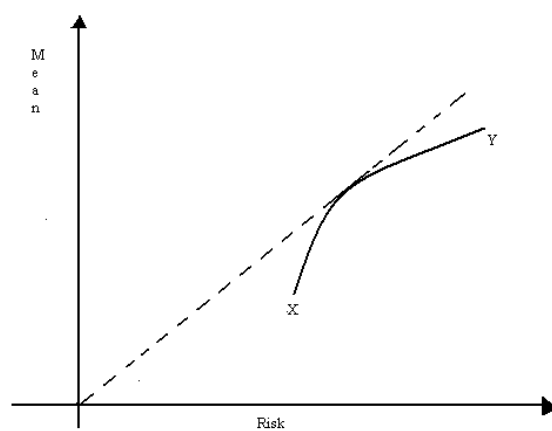


Figure 4. Graphical example where weak diversification holds but the strong diversification does not hold.

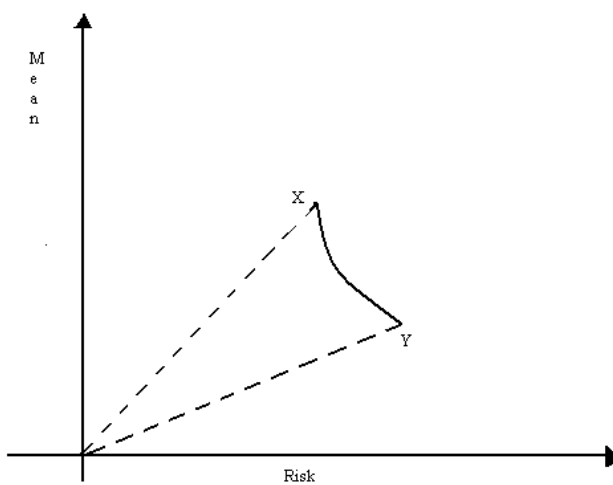


Figure 5. Graphical example where convexity holds but the weak diversification does not holds.

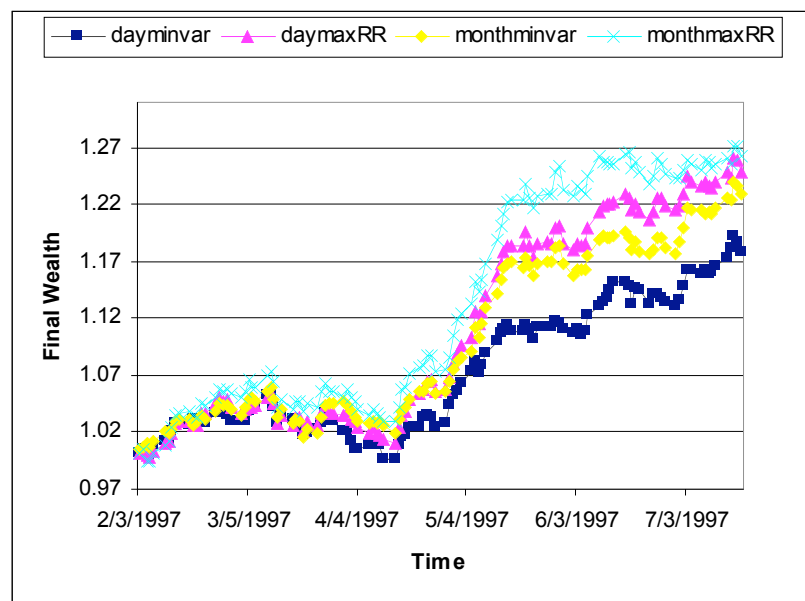


Figure 6 Ex-post comparison among the final wealth of the four optimal portfolios obtained with daily returns (Series dayminvar and daymaxRR) and six months returns and daily frequency (Series monthminvar and monthmaxRR).

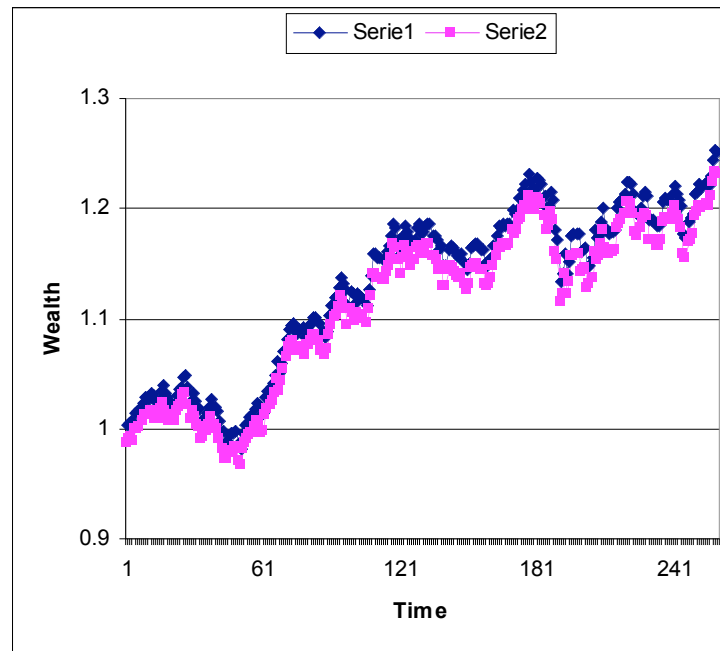


Figure 7. Ex-post comparison between the final wealth of the optimal dynamic strategies obtained maximizing monthly Sharpe ratio with (Series 2) and without transaction costs (Series 1).

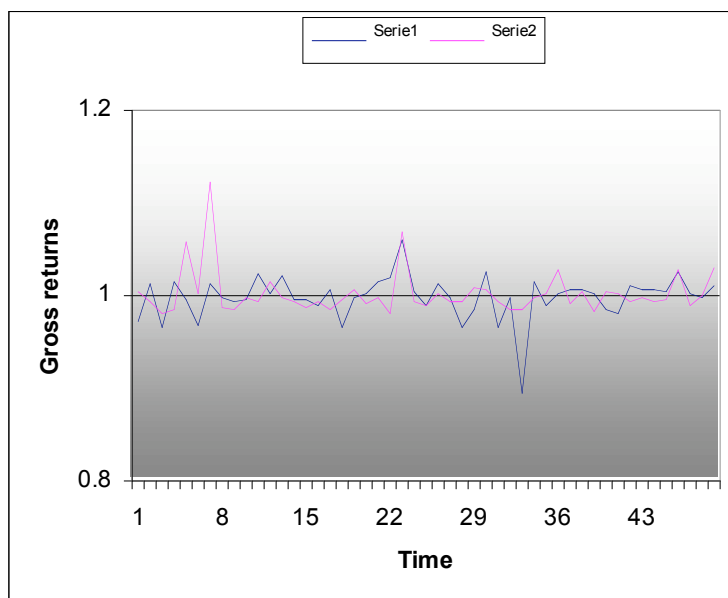


Figure 8. Comparison among simulated data with stable asymmetric distributions.