UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Scienze Economiche “Marco Fanno”

COMPARING AND SELECTING PERFORMANCE MEASURES FOR RANKING ASSETS

MASSIMILIANO CAPORIN
University of Padova

FRANCESCO LISI
University of Padova

April 2009

“MARCO FANNO” WORKING PAPER N.99

Comparing and selecting performance measures for ranking assets*

Massimiliano Caporin†
Department of Economics and Management "Marco Fanno"
University of Padova
Italy

Francesco Lisi
Department of Statistical Sciences
University of Padova
Italy

April 22, 2009

Abstract

Within an asset allocation framework, when the number of assets is larger than the sample dimension, mean-variance approaches cannot be used due to the limited number of degrees of freedom. In such a situation, performance measures could be used to rank assets, and then select a subset of them for further analysis. However, the financial economics literature proposes dozens of measures, and there is thus a problem: which measures should be considered? Some authors already discussed this topic. We extend the current literature by enlarging the set of analyzed measures and also by exploiting the possible dynamic evolution of rank correlations. Our analysis is mainly empirical, based on the S&P 1500 constituents, and includes an example of the optimal combination of performance measures for allocating an equity portfolio.

Keywords: performance measurement, rank correlations, selecting performance measures, comparing performance measures, combining performance measures.

JEL Codes: C10, G11, C40

*Both authors acknowledge financial support from the University of Padova grant CPDA073598. Usual disclaimers apply.

†Corresponding author: Dipartimento di Scienze Economiche "Marco Fanno", Via del Santo, 33, 35123, Padova, Italy - tel. +39-049-827-4258, fax +39-049-827-4211 - email: massimiliano.caporin@unipd.it.

1 Introduction

Since the pioneering works of Sharpe (1964 and 1966) and Treynor (1965), the topic of performance measurement has attracted considerable interest in the financial economic literature. From a general viewpoint, we may identify two fundamental topics covered by performance measurement.The first considers the returns of financial assets, and aims to define and interpret ratios or indices, the performance measures or reward-to-risk ratios, for the purpose of determining the assets' risk/return trade-off. The second analyzes to the returns of managed portfolios and focuses on the introduction and use of models and approaches which make it possible to infer the choices made by investment managers. For examples on the second topic see Knight and Satchell (2002) and the references therein, the literature on style analysis (see Sharpe (1992), among others) and the contributions related to conditional CAPM approaches, including Ferson and Schadt (1996), Avramov and Chordia (2006), and Caporin and Lisi (2009).

This study belongs to the literature dealing with the first issue. We focus on the comparison of performance measures based on the returns of specific assets. The analysis proposed by this strand of the literature may be considered as tools for portfolio managers and agents facing investment decisions. Achieving this objective with performance measures could seem counterintuitive given that mean-variance based approaches, starting from the seminal paper by Markowitz, (1959), may provide appropriate answers. Unfortunately, their assumptions may not be valid, and even generalized mean-variance analysis may be unfeasible because the asset cross-sectional dimension (the number of tracked assets) may be larger than the temporal dimension (the number of periods where returns are available). In these cases, there are not enough degrees of freedom to estimate the full covariance matrix unless a limited factor structure is assumed. Some examples of the last approach are given by the contributions of Ledoit and Wolf (2003 and 2004), Briner and Connor (2008), and Caporin and Paruolo (2009).

Within this paper we focus on cases where the cross-sectional dimension is larger than the temporal one, and performance measures are used to select a small number of assets with given features (such as small drawdown or high return...) for a subsequent allocation using a generalization of the Markowitz approach. Alternatively, performance measures may be used to select a number of assets for the direct application of naïve portfolio allocation rules, such as the equally weighted one (see De Miguel et al. (2009)).

The financial economics literature proposes also to use performance measures as objective functions for determining the weights of an optimal portfolio. In other words, instead of choosing the portfolio with the highest Sharpe index, we may determine the optimal portfolio by maximizing, for instance, the Omega index of Shadwick and Keating (2002). We will not deal with this issue, but for an example of this approach see Farinelli et al. (2008, 2009). A related aspect, not included in the current paper, is the analysis of the optimality properties of risk measures with a focus on optimized portfolio allocation. Examples of this approach are given in Biglova et al. (2004), Ortobelli et al. (2005), Rachev et al. (2008), and Chen and Wang (2008).

In theory, the problem of a limited temporal dimension may be solved by increasing the observation frequency, say from monthly to daily. But pursuing this approach causes a number of associated problems to arise. From the statistical point of view, the returns have time-varying moments, they are far from the normality hypothesis and their prediction is even more complex. From the financial point of view, the analyzed returns should be consistent with the frequency of portfolio rotations or revisions which are generally monthly, given that the use of daily data makes it difficult to efficiently track a large number of companies, either with a quantitative or a qualitative focus. Therefore, we will assume in the following that the data frequency is monthly, and that the time dimension cannot be increased by changing the sampling frequency.

When the cross-sectional dimension is large compared to the temporal ones, performance indices could be considered as the easiest tool (and maybe the only quantitative tool) available for asset ranking and selection. However, a relevant open point remains: which performance measure should be used? In fact, many reward-to-risk ratios are available. Besides the well-known Sharpe, Sortino and Treynor indices, a number of alternative measures have been proposed, such as the Omega index (Shadwick and Keating, 2002), the Rachev ratio (Rachev et al., 2003), and the FT ratios (Farinelli and Tibiletti, 2003a,b), among others. Their number is increasing over time, and many authors propose indices designed in order to meet specific requirements, for example Pedersen and Rudholm-Alfvin (2003), or with the purpose of overcoming the limitations of the oldest measures. Some examples are given by the need of increasing the robustness of performance measures to deviations from normality, or of introducing measures more appropriate for agents characterized by loss aversion (Gemmill et al., 2006) or by aggressiveness (Farinelli and Tibiletti, 2003a,b).

Some authors have already considered the comparison of alternative performances, generally using rank correlations. In particular, we refer to Gemmill et al. (2006), Eling and Schuhmacher (2007), and Eling (2008). The previous contributions use a simple and effective approach for deciding which measures to use: in order to compare alternative indices, they verify whether they rank assets differently. Clearly, performance measures providing equivalent rankings are redundant and may thus be discarded. Following this method, we may identify a restricted set of performance measures carrying different information on the risk/return trade-off.

Building on the work of Eling and Schuhmacher (2007), this paper provides four main contributions. The first one extends the cited paper by broadening the set of performance measures to be compared. In particular, we extend Eling and Schuhmacher (2007) by including performance measures based on loss aversions (Gemmill et al., 2006) and on partial moments (Farinelli and Tibiletti (2003a,b) and Rachev (2003)). In addition, we base our analysis on equities, rather than on hedge funds as in Eling and Schuhmacher (2007). Since we reach different conclusions, we argue that the equivalence across performance measures may depend on the kind of assets considered. As a minor contribution, we also propose four new performance measures: the Expected Return over Range (a variation of the Sharpe ratio); the VaR-ratio (the ratio of upper and lower Value-at-Risk levels); and two measures generalizing the contribution of Gemmill et al. (2006).

The second contribution is associated with a different topic: the stability over time in the rankings induced by different performance measures. We will try to answer the question: "Are rank correlations dynamic?" To that end, we compare the rank correlations computed both over samples of different length, and over rolling windows. To the best of our knowledge, this issue has never been addressed.

Another important topic we consider is that of the redundancy of the performance measures in a dynamic context: given a set of N performances measures, how to reduce them in order to consider only those which really carry different information. In this paper we start, in the most general case, with 80 measures and, using a procedure based on the asymptotic distribution of the rank correlation coefficient, we conclude that 57 measures are redundant since they carry information similar to the 23 we select. In connection with the second outcome of this paper, we also infer that the set of performance measures carrying relevant information may be time-varying as well. This additional piece of information could be proven to be extremely relevant for periodic rotation or rebalancing of managed portfolios.

Finally, using the previous results, at each point in time we are able to build a set containing highly informative performance measures. However, a problem still remains unsolved: which measure should be used? Or, in other words, how could we jointly use the restricted set of selected measures? We propose to combine the "not equivalent" measures we identify following the approach previously described. This topic is not trivial and carries a number of additional problems. For this reason, in this paper we consider the simplest way to combine different performances measures: we sum up the ranks obtained from the selected performance measures to build a composite performance index. A more advanced and complex approach can be found, for example, in Hwang and Salmon (2002) that uses copula functions to combine the measures. The optimal combination of different reward-to-risk ratios is not the main purpose of the paper, and we leave this topic to future researches. In this work, we simply show that the use of our naïve combination can be effective in a simulated asset allocation strategy.

In the remainder of the paper, Section 2 reviews the performance measures that will be considered in our work. Section 3 describes the dataset and discusses some problems connected to bias selection. In Section 4 we report the results of the analysis concerning the correlations between different performance measures and we show how to reduce the initial set to the measures that are significantly different. Section 5 considers the combination of different performance measures and shows its effectiveness in a simulated portfolio allocation problem. Our final conclusions are presented in Section 6.

2 Performance measures: review and classification

We introduce here the set of performance measures that we will empirically compare in the following sections. From a general viewpoint, performance measures can be defined as ratios between a reward measure and a risk measure, and their value can be interpreted as the reward per unit of risk. Despite general agreement on what a performance measure is, a number of choices are available for the reward and risk measures to be considered, as well as the typed of variables to be used for their evaluation. For instance, many authors compute these quantities on nominal returns of financial instruments, or on excess returns from a risk-free investment, or on deviations from a benchmark investment. All three of these approaches have a proper relevance. The first allows comparisons of the performances of risky instruments without the introduction of a reference quantity. In the second and third, the users evaluate the performances in excess of what they could obtain from a benchmark investment (risky or risk-free). The third case, deviations from a risky benchmark, may also be used to evaluate the effectiveness of active management for managed portfolios.

In order to provide a general setup that nests all the previous cases, we start by introducing some notation: t is the time index, and the available sample on which alternative performance measures are computed goes from 1 to T; returns on assets are determined as the log-price difference; we use $R_{i,t}$ to denote the (nominal) returns of asset i in period t; we assume that there exists a risk-free investment (over each period), denoted by $R_{f,t}$ (note the risk-free is time-varying since we consider it a pure risk-free investment within each period, that is, at the beginning of the period we know which will be the return we will get from that investment at the end of the period), and a benchmark index (or a market return index), called $R_{B,t}$.

The performance measures presented below will be defined over a variable $X_{i,t}$ that may take one of the following values

$$X_{i,t} = \left\{ \begin{array}{c} R_{i,t} \\ R_{i,t} - R_{f,t} \\ R_{i,t} - R_{B,t} \end{array} \right. . \tag{1}$$

These cases represent the three possible relevant dimensions for performance measurement, not necessarily mutually exclusive. In addition, we will denote by $X_{t=1}^{T}$ the sequence of observations of the variables X_t from time 1 to time T, by $E\left[X^p\right]$ the pth-order moment of X, by $\sigma\left[X\right]$ the volatility of X and by $E\left[X^p|Y\right]$ the conditional pth-order moment of X. The classification of performance measures we follow extends and generalizes the approach in Eling and Schuhmacher (2007).

We will not include in the following list the measures proposed by Favre and Galeano (2002), Gregoriou and Gueyie (2003), and Zakamouline and Koekebakker (2009). These authors modify standard performance measures to cope with deviations from normality. However, most of the indices we consider in the following are not assuming normality and could be considered for highly non-Gaussian assets. As a result, the inclusion of the measures proposed by the previously cited authors would need a deeper discussion of the links between the comparison of the performance measures and the true densities of the underlying assets. We believe this topic is outside the scope of the present paper, and it is left for future researches.

2.1 Traditional (and similar) performance measures

This first set of performance measures contains the most known and traditional indices. The starting point for performance evaluation of financial instruments is the Sharpe ratio, introduced by Sharpe (1966, 1994). The Sharpe ratio can be computed as the ratio between the expected returns and their standard deviations as

$$Sh\left(X_{i,t}\right) = \frac{E\left[X_{i,t}\right]}{\sigma\left[X_{i,t}\right]}. \tag{2}$$

Note that our representation nests the traditional one when computed on deviations from a time independent risk free return ($X_{i,t} = R_{i,t} - R_f$). The Sharpe ratio represents the compensation per unit of overall risk. A related measure is the Information ratio which is computed as the ratio between the mean and the standard deviation of a tracking error measure. In turn, tracking error could be defined as the deviation between the returns of a financial asset (security or fund) from a reference benchmark. In our representation, the Sharpe computed on excess returns is an Information ratio where the risk-free is a benchmark. In contrast, the Sharpe ratio computed on deviations from the benchmark is the traditional Information ratio ($X_{i,t} = R_{i,t} - R_{B,t}$).

Other standard performance measures could be derived from the empirical estimates of the CAPM model of Sharpe (1964), Lintner (1965) and Mossin (1969). In our setup, the CAPM equation could be estimated both on asset returns as well as on deviations from the risk-free investment, giving the two alternative equations

$$\begin{aligned} R_{i,t} &= \alpha_i + \beta_i R_{B,t} + \varepsilon_{i,t}, \\ R_{i,t} - R_{f,t} &= \alpha_i + \beta_i\left(R_{B,t} - R_{f,t}\right) + \varepsilon_{i,t}. \end{aligned} \tag{3}$$

Using the estimated parameters, we could compute the index proposed by Treynor (1965), for nominal returns and for excess returns as

$$Tr\left(X_{i,t}\right) = \frac{E\left[X_{i,t}\right]}{\beta_i}. \tag{4}$$

Note that when $X_{i,t} = R_{i,t} - R_{B,t}$ the Treynor index has no meaning. The intercept in equations (3) is the Jensen Alpha, a performance measure proposed by Jensen (1968). This index represents the extra-compensation provided by a financial instrument in excess of that predicted by the CAPM model. However, the Jensen Alpha is not consistent with the definition of performance measure which we consider, that is, reward to variability ratios. In our setup the Jensen Alpha is simply a reward measure. A different index, derived again from (3), is the Appraisal ratio, defined as

$$AR\left(X_{i,t}\right) = \frac{\alpha_i}{\sigma\left[\varepsilon_{i,t}\right]}. \tag{5}$$

The Sharpe and Treynor performance measures differ in the way they measure the risk of asset i. Additional performance measures are derived by using alternative risk indices. The Mean Absolute Deviation (MAD) was proposed as a risk measure in Konno (1990), while Konno and Yamazaki (1991) proposed as a performance measure the expected return over MAD ratio

$$ERMAD\left(X_{i,t}\right) = \frac{E\left[X_{i,t}\right]}{E\left[\left|X_{i,t} - E\left[X_{i,t}\right]\right|\right]}. \tag{6}$$

Differently, Young (1998) proposed the expected return over the MiniMax ratio

$$ERMM\left(X_{i,t}\right) = \frac{E\left[X_{i,t}\right]}{\max\left(\max X_{t=1}^{T}, -\min X_{t=1}^{T}\right)}. \tag{7}$$

We contribute to this group of indices by adding the expected return over the range ratio

$$ERR\left(X_{i,t}\right) = \frac{E\left[X_{i,t}\right]}{\max X_{t=1}^{T} - \min X_{t=1}^{T}}. \tag{8}$$

These indices differ from the Sharpe index only in the method used to measure the risk. A rather different measure was proposed by Modigliani and Modigliani (1997), the Risk Adjusted Performance (RAP) or M2 index. M2 is based on the excess return from a benchmark in the presence of a risk-free investment, and introduces a correction, taking into account the possible different risk levels of asset i and benchmark B. The index M2 is defined as

$$M2 = \left(E\left[R_{i,t}\right] - E\left[R_{B,t}\right]\right)\frac{\sigma\left[R_{B,t}\right]}{\sigma\left[R_{i,t}\right]} + E\left[R_{f,t}\right] - E\left[R_{B,t}\right]. \tag{9}$$

2.2 Measures based on Drawdown

Most of the previous performance measures are based on a risk index whose purpose is the evaluation of the overall risk of an asset. However, the risk measurement may follow alternative approaches (see Biglova et al. (2004), Ortobelli et al. (2005), and Rachev et al. (2008) for a survey). One possible alternative is to use as risk measure a quantity based on Drawdowns. Let us define Drawdown in a recursive way as

$$D_t\left(X_{i,t}\right) = \min\left(D_{t-1} + X_{i,t}, 0\right) \quad D_0 = 0. \tag{10}$$

Given the sample observations for $X_{i,t}$ $t = 0, 1, ...T$, the Drawdown $D_t\left(X_{i,t}\right)$ or simply D_t represents, at time t, the maximum loss an investor may have suffered from 0 to t. Risk measures are defined ordering drawdowns and computing quantities such as the maximum drawdown, $D_1(X_{i,t}) = \min D_{t=1}^{T}$, or the second largest drawdown $D_2(X_{i,t}) = \min\left(D_{t=1}^{T} - D_1(X_{i,t})\right)$, and so on. We also assume $D_1(X_{i,t}) < 0$. Three indices based on drawdowns have been proposed:

- the Calmar ratio, suggested by Young (1991), is the ratio between expected returns and the maximum drawdown:

$$CR(X_{i,t}) = \frac{E[X_{i,t}]}{-D_1(X_{i,t})}; \tag{11}$$

- the Sterling ratio, introduced by Kestner (1996), is the ratio between the expected returns and the N largest drawdowns:

$$SR(X_{i,t}; w) = \frac{E[X_{i,t}]}{-\frac{1}{w}\sum_{j=1}^{w} D_j(X_{i,t})}; \tag{12}$$

- the Burke ratio, due to Burke (1994), is the ratio between the expected returns and the second order non central moment of the N largest drawdowns:

$$BR(X_{i,t}; w) = \frac{E[X_{i,t}]}{\left(\frac{1}{w}\sum_{j=1}^{w}[D_j(X_{,t})]^2\right)^{\frac{1}{2}}}. \tag{13}$$

The Sterling and Burke ratios depend on a parameter, w, that identifies the number of values used in the computation of the risk index. While Eling and Schuhmacher (2007) fix the value to 5, we prefer linking the number of drawdowns to the sample dimension as $w = \left\{\left[\frac{T}{20}\right], \left[\frac{T}{10}\right]\right\}$ where $[a]$ denotes the nearest integer of a.

2.3 Measures based on partial moments

Drawdown-based indices measure the risk by assigning a larger weight to negative returns or cumulative losses, while the risk indices used in traditional performance measures consider the entire distribution of returns. In the spirit of drawdowns, Sortino and van der Meer (1991) define risk using a lower partial moment (LPM). In fact, LPMs could be preferred because they consider the downside deviations from the target return or minimum acceptable return while upside movements do not enter in the evaluation of risk but in the evaluation of returns only. Two examples of performance measures using partial moments are:

- the Sortino or Sortino-Satchell ratio, Sortino and Van der Meer (1991), Sortino (2000), Sortino and Satchell (2001), and Pedersen and Satchell (2002):

$$Sr(X_{i,t}) = \frac{E[X_{i,t}]}{E\left[(\min(X_{i,t}, 0))^2\right]^{\frac{1}{2}}}; \tag{14}$$

- the Kappa 3 measure of Kaplan and Knowles (2004):

$$K3(X_{i,t}) = \frac{E[X_{i,t}]}{E\left[(\min(X_{i,t}, 0))^3\right]^{\frac{1}{3}}}. \tag{15}$$

Farinelli and Tibiletti (2003a,b) generalize the previous indices introducing partial moments also in the evaluation of rewards proposing the FT ratio:

$$FT\left(X_{i,t};b,p,q\right)=\frac{E\left[\left(\left(X_{i,t}-b\right)^{+}\right)^{p}\right]^{\frac{1}{p}}}{E\left[\left(\left(X_{i,t}-b\right)^{-}\right)^{q}\right]^{\frac{1}{q}}} \tag{16}$$

where $(X_{i,t}-b)^{+}=\max\left(X_{i,t}-b,0\right)$, $(X_{i,t}-b)^{-}=\max\left(b-X_{i,t},0\right)$. The quantities $E\left[\left(\left(X_{i,t}-b\right)^{-}\right)^{q}\right]^{\frac{1}{q}}$ and $E\left[\left(\left(X_{i,t}-b\right)^{+}\right)^{p}\right]^{\frac{1}{p}}$ are the LPM of order p and the Upper Partial Moment (UPM) of order q, respectively, and b is a threshold return level separating the returns entering in the upper and lower moments. The FT ratios are thus generalized ratios between an Upper Partial Moment and a Lower Partial Moment with respect to a given threshold or minimum acceptable return b. Following Farinelli and Tibiletti (2003a,b), we could consider the following combination of UPM and LPM orders, which are associated to investors' styles or preferences:

- $p=0.5$ and $q=2$ for a defensive investor;
- $p=1.5$ and $q=2$ for a conservative investor;
- $p=q=1$ for a moderate investor (note that this combination makes the $FT\left(X_{i,t};b,1,1\right)$ equivalent to the Omega index of Shadwick and Keating (2002), and Kazemi et al. (2003));
- $p=2$ and $q=1.5$ for a growth investor;
- $p=3$ and $q=0.5$ for an aggressive investor.

The previous indices will all be considered in the empirical part together with the parameter combination $p=1$, $q=2$ which defines the Upside Potential Ratio of Sortino et al. (1999). Finally, as b represents a minimum acceptable return, we could fix it to zero, simply distinguishing between positive and negative $X_{i,t}$, or to different values. In the empirical part we will consider the following cases $b=\{-0.02,0,0.02\}$, where the -2% and 2% values may represent the choices of a less risk averse and a more risk averse investor, respectively.

2.4 Measures based on quantiles

A class of performance measures similar to the previous one replaces partial moments with reward and variability measures based on quantiles (see Rachev et al. (2003), Biglova et al., 2004, among others). At first, we define the following quantities:

- the Value-at-Risk at the α level is the quantity $VaR\left(X_{i,t};\alpha\right)$ such that $P[X_{i,t}\leq VaR\left(X_{i,t};\alpha\right)]=\alpha$;
- the Expected Shortfall or Conditional Value at Risk (CVAR), or Expected Tail Loss (ETL)

$$ES\left(X_{i,t};\alpha\right)=E\left[X_{i,t}|X_{i,t}\leq VaR\left(X_{i,t};\alpha\right)\right]. \tag{17}$$

Using $VaR\left(X_{i,t};\alpha\right)$ and $ES\left(X_{i,t};\alpha\right)$, a number of indices could be considered:

- the Expected return over absolute VaR, of Dowd (2000), Favre and Galeano (2002), and Rachev et al. (2003):

$$VR\left(X_{i,t};\alpha\right)=\frac{E\left[X_{i,t}\right]}{\left|VaR\left(X_{i,t};\alpha\right)\right|}; \tag{18}$$

- the VaR ratio, defined as:

$$VaRR\left(X_{i,t};\alpha\right)=\frac{\left|VaR\left(-X_{i,t};\alpha\right)\right|}{\left|VaR\left(X_{i,t};\alpha\right)\right|}; \tag{19}$$

(to the best of our knowledge this index has never appeared in the literature - it could be also considered as a tail asymmetry index);

- the Expected return over absolute Expected Shortfall or Conditional Sharpe ratio or STARR ratio (Rachev et al., 2003, and Agarwal and Naik, 2004):

$$STARR\left(X_{i,t};\alpha\right)=\frac{E\left[X_{i,t}\right]}{\left|ES\left(X_{i,t};\alpha\right)\right|}; \tag{20}$$

- finally, the Generalized Rachev Ratios (Biglova et al., 2004):

$$GR\left(X_{i,t};\alpha,p,q\right)=\frac{E\left[\left|X_{i,t}\right|^{p}\left|X_{i,t}\geq -VaR\left(-X_{i,t};\alpha\right)\right]^{\frac{1}{p}}}{E\left[\left|X_{i,t}\right|^{q}\left|X_{i,t}\leq VaR\left(X_{i,t};\alpha\right)\right]^{\frac{1}{q}}},$$

where $p>0$ and $q>0$ are conditional moment orders.

In this last index, the combination $p=q=1$ gives the simple Rachev Ratio (Biglova et al., 2004) and the two orders could be combined as for the Farinelli and Tibiletti ratios, thus providing a moment design associated with investors' type. We note that the Generalized Rachev ratios are ratios of conditional upper and lower moments.

In the empirical part we will use 5% and 10% as the quantile levels. Note that, differently from other papers, we introduce the absolute value over Value-at-Risk and upper and lower conditional moments given that these quantities are defined over the returns. This allows obtaining positive values risk indices and avoids computing powers of upper negative quantities (in principle, upper VaR could be negative for an asset evaluated over an extremely negative period).

2.5 Measures derived from utility functions

Some performance measures deviate from the general structure of reward-to-variability ratios. A relevant example is given by quantities derived from utility functions, and allowing the computation of risk-adjusted returns. In the mutual fund industry such a measure is used by Morningstar. The Morningstar Risk-Adjusted Return, MRAR (Morningstar, 2007), is derived from a power-utility function and defined as the expected value of the certainty equivalent annualized geometric return. We rearrange the Morningstar formulae and represent MRAR over $X_{i,t}$ as

$$MRAR\left(X_{i,t};\lambda\right)=\begin{cases} E\left[\left(1+X_{i,t}\right)^{-\lambda}\right]^{-\frac{12}{\lambda}} & \lambda>-1,\lambda\neq 0 \\ e^{E\left[\ln\left(1+X_{i,t}\right)\right]} & \lambda=0 \end{cases}, \tag{21}$$

where λ identifies the risk aversion coefficient (which is set equal to 2 by Morningstar).

The MRAR is also known as AIRAP, Alternative Investments Risk Adjusted Performance, Sharma (2004). Sharma (2004) suggests a risk aversion coefficient of 3, while Ait-Sahalia et al. (2001) estimates a coefficient of 2.2 for ultra high net worth individuals. In the empirical part, we will fix the risk aversion coefficient to the values 2, 10 and 50.

Gemmill et al. (2005) introduced a set of performance measures derived within a behavioral finance framework. Following the prospect theory of Kahnemann and Tversky (1979), the utility function is replaced by a value function displaying loss aversion, and focusing on gains and losses at time t with respect to the beginning of period wealth W_{t-1}. The following equation defines the value function

$$V_t\left(X_{i,t}\right)=\begin{cases} \left[W_{t-1}X_{i,t}\right]^p & X_{i,t}\geq 0 \\ -\lambda\left[-W_{t-1}X_{i,t}\right]^q & X_{i,t}<0 \end{cases}, \tag{22}$$

where p, q and λ are positive parameters, and loss-aversion is included if $\lambda>1$. Note that the value function is defined over gains or losses while the wealth evolves according to $W_t=W_{t-1}\left(1+R_{i,t}\right)$. In the studies of Tversky and Kahnemann (1992), Bernatzi and Thaler (1995) and Gemmill et al. (2005), λ was set to 2.25. In contrast, p and q were set to 0.75 and 0.95, respectively, by Gemmill et al. (2005), and both to 1 by Barberis et al. (2001). The Value Function in (22) displays a 'House-Money' effect, as defined in Barberis et al. (2001), if the loss aversion coefficient depends on previous gains and losses, thus becoming time varying

$$\lambda_t=\lambda_0+\lambda_1 W_{t-2}X_{i,t-1}. \tag{23}$$

Following Gemmill et al. (2005) we define a set of performance measures accounting for loss aversion. Like the cited authors, we first rewrite the value function as

$$V_t\left(X_{i,t}\right)=\left[W_{t-1}X_{i,t}\right]^p I\left(X_{i,t}\geq 0\right)-\lambda\left[-W_{t-1}X_{i,t}\right]^q I\left(X_{i,t}<0\right), \tag{24}$$

where the first component identifies gains while the second identifies losses. The expectation of the ratio between the two quantities is a performance measure

$$\frac{E\left[\left[W_{t-1}X_{i,t}\right]^p I\left(X_{i,t}\geq 0\right)\right]}{E\left[\lambda\left[-W_{t-1}X_{i,t}\right]^q I\left(X_{i,t}<0\right)\right]}=\frac{E\left[\left(W_{t-1}X_{i,t}\right)^p|X_{i,t}\geq 0\right]}{E\left[\lambda\left(-W_{t-1}X_{i,t}\right)^q|X_{i,t}<0\right]}, \tag{25}$$

as it can be considered a reward to variability quantity. Similar equivalences can be considered with the introduction of the House money effect.

Gemmill et al. (2005) suggest as performance measures the following ratios

$$LAP^S = \frac{E\left[(X_{i,t})^p \left| X_{i,t} \geq 0\right.\right] P\left(X_{i,t} \geq 0\right)}{E\left[(-X_{i,t})^q \left| X_{i,t} < 0\right.\right]\left(1 - P\left(X_{i,t} \geq 0\right)\right)}, \tag{26}$$

$$LAP^H = \frac{E\left[(X_{i,t})^p \left| X_{i,t} \geq 0\right.\right] P\left(X_{i,t} \geq 0\right)}{\lambda_t E\left[(-X_{i,t})^q \left| X_{i,t} < 0\right.\right]\left(1 - P\left(X_{i,t} \geq 0\right)\right)}, \tag{27}$$

where $P\left(X_{i,t} \geq\right)$ is the probability of having returns above zero. Given that the wealth evolves according to $W_{t-1} = W_{t-2}\left(1 + R_{i,t-1}\right)$ the previous indices are valid only if the expectations are conditional to time $t-1$ information set. In fact, under this restriction the ratios involving the wealth are known at time t. In this case the ratio in (25) becomes

$$\frac{W_{t-1}^p}{\lambda W_{t-1}^q} \frac{E\left[(X_{i,t})^p \left| X_{i,t} \geq 0\right.\right]}{E\left[(-X_{i,t})^q \left| X_{i,t} < 0\right.\right]}. \tag{28}$$

Then, if the purpose is the selection of the best instruments for an investment with a one-period horizon, today's wealth can be fixed at 1 and the first ratio in (28) is irrelevant in the comparison of alternative assets. Notably, the ratios proposed by Gemmill et al. (2005) add a Loss Aversion interpretation to the FT ratios when the minimum acceptable return b is set to zero (and House money is not included). However, in the ex-post evaluation of an investment, the indices in (26) loose the loss aversion interpretation because they do not maintain the effective perception of reward and risk faced by the investors. In fact, this is given by the wealth evolution, the risk aversion coefficients, and, when present, by the House Money effect. In order to make the indices more suitable for the ex-post performance evaluation we introduce two other measures:

$$\begin{aligned} LAP^{WS} &= \frac{\sum_{t=1}^T \left(W_{t-1} X_{i,t}\right)^p}{\sum_{t=1}^T I\left(X_{i,t} \geq 0\right)} \left(\frac{\sum_{t=1}^T \left(-W_{t-1} X_{i,t}\right)^q}{\sum_{t=1}^T I\left(X_{i,t} < 0\right)}\right)^{-1}, \\ LAP^{WH} &= \frac{\sum_{t=1}^T \left(W_{t-1} X_{i,t}\right)^p}{\sum_{t=1}^T I\left(X_{i,t} \geq 0\right)} \left(\frac{\sum_{t=1}^T \lambda_t \left(-W_{t-1} X_{i,t}\right)^q}{\sum_{t=1}^T I\left(X_{i,t} < 0\right)}\right)^{-1}. \end{aligned}$$

The two indices we propose do not include the expectation symbol since they are not unconditional expectations. LAP^{WS} is the ratio of gains to losses realized over the sample without the loss aversion effect, while LAP^{WH} is the ratio between the cumulative sums of value functions for gain with respect to that for losses. Note that the LAP^S index is equivalent to the Omega index of Shadwick and Keating (2002) if $p = q = 1$. In the empirical part we will fix the p and q orders to the values suggested by Gemmill et al. (2005) and to the combinations proposed by Farinelli and Tibiletti (2003a,b), thus allowing for different investor types. In all cases, the threshold for gains and losses is set to 0. The case with orders set to 1 is not considered since it provides quantities equivalent to the Omega index.

3 Data and performance measures

In the following section, we compare a set of performance measures over a dataset that contains the stocks included in the S&P 1500 index. The index covers from large-cap to small-cap stocks and is thus heterogeneous with respect to the company market value. We retrieved from Datastream the monthly returns of the S&P 1500 components for the period January 1990 - October 2008. For these assets, the S&P 1500 index represents the appropriate equity benchmark, and the US 1-month bond index, provided by Citigroup, is our proxy for the risk-free asset. For each asset, we consider logarithmic returns and excess returns over the risk-free asset and over the benchmark.

Note that the index composition changes over time. Since our dataset includes the 1500 assets belonging to the S&P 1500 index at the end of October 2008, not all of them are available for the whole considered period (for example, in January 1990 only 754 assets out of 1500 were available). To deal with this problem, we followed two different strategies. At first, we focused on the last 120 observations of the sample, corresponding to the period November 1998 - October 2008: there are 1236 assets always included in the index over this range. On this reduced set of stocks, we performed a static analysis of the rank correlation using three different evaluation windows: November 1998 - October 2008 (120 monthly returns); November 2003 - October 2008 (60 monthly returns); November 2005 - October 2008 (36 monthly returns). This study allows a comparison of performance measures over time, avoiding possible effects due to a changing cross-sectional dimension. We also obtain some preliminary results on the window size effect and on the time-varying nature of the rankings.

In a second step, we focus on the entire sample (January 1990 - October 2008) and use a rolling approach to evaluate the stability of rankings over time. At this stage, the rank correlations are measured over a rolling window of 60 months for assets always available in the window. The number of assets is 754 in the first window and 1404 in the last one. This different approach allows a comparison of rank correlations over a number of periods.

The use of an increasing number of assets over time could be questionable. However, using only the 754 available for the entire sample period would have induced a sample selection bias in the analysis. Clearly, the optimal solution would have been to use the entire track record of all the S&P 1500 components, including dead or de-listed companies, but unfortunately this piece of information was not available to us. This second approach also permits us to study portfolio allocation rules based on performance rankings.

[TABLE 1]

Table 1 lists the performance measures considered in the empirical analysis. In brackets we report the number of cases considered for each performance measure, deriving it from the parameter combinations presented in the previous section. For instance, the Burke and Sterling ratios have two different cases associated with the two values of the parameter w. Similarly, the Farinelli-Tibiletti ratios are included in eighteen different forms combining the six cases for the moment order pairs and the three thresholds. Note that the LAP measures include 19 cases, obtained by combining the 4 performance measures described in the previous section, and 6 parameter combinations mimicking Farinelli and Tibiletti (2003a,b) and Gemmill et al. (2005). The 5 cases with LAP^S and the FT parameter combinations are not considered since these are equivalent to Omega measures (which are separately included in the list).

4 Empirical analysis

We compute the previously listed performance measures over the S&P 1500 constituents and compare them using the Spearman rank correlation (R_S). This choice makes our results comparable with those in Eling and Schuhmacher (2007) and Eling (2008). After the Z-transformation of Fisher (1915), the Spearman rank correlation has an asymptotic density which could be used to test the null hypothesis of independence between two variables. However, our purpose is not to test independence, but rather to study the degree of correlation between performance measure based ranks and, in particular, to detect measures that are highly correlated or concordant. Eling and Schuhmacher (2007) tested the null hypothesis $R_S \leq p$, and determined the value of p associated with a rejection of the null for all assets. They found that for $p = 0.917$ the null hypothesis was always rejected. Note that the test cannot be applied under the null of unit correlation, i.e. perfect agreement, because, as claimed also by Eling and Schuhmacher (2007), in this case there is no discrepancy between the rankings induced by the performance measures and thus no variability.

In this work, we follow an approach similar to that of Eling and Schuhmacher (2007), but differing in the kind and in the number of assets used to compute performance measures. In fact, the database of Eling and Schuhmacher (2007) includes both surviving and dead hedge funds, implying by construction that the series have different lengths across funds and that the sample periods are not the same. The inclusion of data referring to different periods may strongly affect the results about performance evaluation. In contrast, we focus on equities and we always compute performance measures, and the associated rank correlations, across assets available over a common period. In addition, our results suggest that the threshold level p may depend on the asset type as well as on the sample dimension.

Another important issue not considered by Eling and Schuhmacher (2007), is that of defining the decision rule that specifies when two performance measures carry different pieces of information. Since Eling and Schuhmacher (2007) found only very high correlations between performance measures, they did not face the problem of defining what is a "low" rank correlation. Instead, for our data, we do find evidence of "low" correlation and, in order to develop a decision rule, we define as "low" a rank correlation lower than 0.8. Since we only know the value of the sample rank correlation, $\hat{R}_S$, to define a precise threshold, we considered the asymptotic distribution of R_S. We thus considered the critical value, at 1% significance level, of the test $H_0 : R_S \leq 0.8$ against $H_1 : R_S > 0.8$. In detail, if we denote by ρ the Fisher transformation of R_S, $\rho = \frac{1}{2}\ln(\frac{1+\hat{R}_S}{1-\hat{R}_S})$ and by $\hat{\rho}$ the corresponding sample quantity, asymptotically we have

$$\sqrt{N-2}\hat{\rho} \sim \mathcal{N}\left(\rho, 1\right).$$

This allows us to define the required threshold for R_S as

$$R_S^*\left(\alpha\right) = \frac{\left(\exp\left(\ln\left(\frac{1+R_S}{1-R_S}\right) + 2Z_{1-\alpha}\sqrt{\frac{1}{N-2}}\right)\right) - 1}{\left(\exp\left(\ln\left(\frac{1+R_S}{1-R_S}\right) + 2Z_{1-\alpha}\sqrt{\frac{1}{N-2}}\right)\right) + 1},$$

where $Z_{1-\alpha}$ is the $(1-\alpha)-$th quantile of a standard normal distribution. In our analysis, with $N = 1236$ in the static case, and $\alpha = 1\%$, the threshold defining the low correlation is 0.822.

4.1 Within group analysis

In this section we report, analyze and comment on the rank correlation between performance measures that differ only for the parameters included in their definition. The purpose of this section is to provide a first reduction of the number of performance measures included in Table 1.

The first group we consider includes some measures based on Drawdowns: the Sterling and Burke indices. These two quantities depend on the number of returns used for their computation. In the previous section we suggest the use of at least two values associated with 5% and 10% of the sample dimension. Given these two sets of performance measures, we evaluate whether the number of points used in the computation of the indices provides a different ranking across the assets. The results are reported in the first and second row of Table 2. The rank correlations show evidence of equivalent informative content of the performance measures with respect to the number of returns used for the evaluation of the Burke and Sterling indices. Result do not change with respect to the sample dimension or to the return used for the evaluation. We conclude that there is no need to consider the Sterling and Burke indices computed over different numbers of drawdowns.

The second set of performance measures we analyze includes the quantile based measures, with the exclusion of the Generalized Rachev ratios. Table 2 reports the rank correlations between the VR index, the VaR ratios and the Conditional Sharpe index at the 5% and 10% quantile levels. Results show that the VR index and the Conditional Sharpe ratios should be considered with a single quantile level (rank correlation is always higher than 0.985) while the VaR ratio should be considered with both the 5% and 10% quantiles, given that the rank correlation is lower than 0.822 in all cases and also reaches a minimum close to 0.6 with a 10 years sample dimension (irrespective of the return used).

Table 3 reports the rank correlations across the Generalized Rachev ratios. We recall that we computed 10 different GR ratios combining five parameter combinations (Aggressive, Growth, Moderate, Conservative and Defensive) with two quantile levels (5% and 10%). We distinguished two groups, separating the effect of the Aggressive indices. Our analysis points out that this last parameter combination is the most sensible to the sample dimension, providing results different from the other GR ratios when the sample used is medium to small (3 or 5 years). The difference tends to be canceled with the sample set to 10 years, with the exclusion of the case of the evaluation of deviations from the benchmark. Differently, the other GR ratios (Growth to Defensive) are almost equivalent (the smallest rank correlation is equal to 0.955). In addition, the effect of the quantile level is minor. Building on these results, we chose to include the Moderate GR ratio at the 10% level when the sample dimension is large (10 years). In contrast, when the sample is small or medium, the GR for Aggressive investors will also be considered (again at the 10% level).

Following the performance measure groups previously introduced, we move then to measures based on partial moments that include the indices of Sortino, the Kappa 3 index and the FT ratios. Similarly to the Generalized Rachev ratios, we group the FT performance measures into two sets, separately considering the Aggressive parameter combination. The results are reported in Table 4, where the first group includes the parameter combinations Growth, Moderate, Conservative, Defensive as well as the Upside Potential Ratio (which is a special case of the FT index as we previously argued). Our analysis shows that these parameter combinations do not provide additional information or relevant differences in the ranking of the underlying assets (first to third rows). The result is marginally influenced by the sample length and the kind of returns used in the evaluation of the indices. On the other hand, the threshold used in the index construction matters, making the indices sensibly different in terms of assets ranking (fourth to sixth rows of Table 4). In fact, the rank correlations across indices computed over different thresholds are generally small and always lower than 0.822. When considering the Aggressive parameter combination, the rank correlations are always small, and sometimes negative. In addition, they are affected by both the sample dimension and the return type. Summarizing, we suggest considering the FT Moderate index (or Omega index) together with the Aggressive parameter combination, under all three of the thresholds considered. For the Sortino and Kappa 3 indices, the rank correlation with respect to the Omega index is higher than 0.98 and therefore the two indices are not considered.

Moving to the performance measures based on utility functions, we first note that the MRAR indices with risk aversion set to 10 and 50 are almost equivalent. Therefore, we decide to focus on the measure with risk aversion set to 2 and 10. By contrast, in the LAP measures, the Hwang-Satchell, Moderate and Growth parameter combinations are almost equivalent while the Conservative case is very close to them. In order to provide a selection of measures which is limited, internally consistent, and that maximizes the difference across parameter combinations, we suggest focusing on the cases Defensive, Moderate and Aggressive. Within each group, we suggest considering all performance measures even if the Moderate case reports a high within-group average rank correlation.

Finally, we consider a further group composed by most of the traditional performance measures. Table 6 includes the rank correlation of these indices with the ranking induced by the Sharpe index. As we may observe, all indices are almost identical to the Sharpe ratio in terms of ranking of the assets. Some minor exceptions are the Appraisal ratio and the M2 index for the 3 year sample. Overall, we may infer that the Treynor index, the Appraisal ratio and the indices replacing the standard deviation in the Sharpe with a proxy are all equivalent. We thus suggest introducing in the following analysis only the Sharpe index. Notably, this result is in line with the findings of Eling and Schuhmacher (2007). In our case, the rank correlations are not as high as shown by these authors. Furthermore, our results point out that the equivalence across the selected performance measures is not influenced by the return used for the evaluation and only scarcely affected by the sample dimension.

After this within-group analysis, we select the following performance measures: the Sharpe ratio; the Calmar ratio; the Sterling Ratio and the Burke ratio computed over the 5% of the sample dimension; the VR index and the Conditional Sharpe at the 5% quantile; the VaR ratio at both the 5% and 10% quantiles; the Generalized Rachev ratio with Moderate parameter combination at the 10% quantile level (one single index - the Aggressive index is included only if the evaluation window is small); the FT Moderate and Aggressive indices under all three threshold levels (6 indices); the MRAR index with risk aversion set to 2 and 10; and the LAP measures for Defensive, Moderate and Aggressive parameter combinations (9 indices). On the whole, the total number of selected measures is 26.

[TABLES 2 TO 6]

4.2 Descriptive and rolling analysis of selected measures

We run additional correlation analysis on the reduced set of performance measures identified in the previous section. As a first outcome, we highlight that some of the measures are still highly correlated. In particular, we report in Table 7 the correlation between the Sharpe ratio and some selected measures. As shown in the table, we may infer that the Calmar ratio, the Sterling ratio (5%), the VR Index (5%), and the Conditional Sharpe (5%) are all equivalent to the Sharpe ratio. These findings confirm the results of Eling and Schuhmacher (2007) and are in line with the findings of Ortobelli et al. (2005) showing that traditional risk measures induce indifference across performance measures where the reward index is the average return. However, we obtain rather different rank correlations for Omega, with values going down to 0.536 and high rank correlation for long samples (120 months) only. Note that these differences are pronounced if we compute the Omega over Excess Returns or Deviations from the benchmark, while in the case of asset returns the Omega (with a zero threshold) is very close to the Sharpe, as in Eling and Schuhmacher (2007).

[TABLE 7]

Such a result points out that ranking of performance measures and their equivalence may be influenced by the kind of assets considered, the return type (nominal or excess return), the estimation window, and the sample period. To shed some light on the last motivation, and considering also the purposes of the actual paper, we perform a rolling analysis on the rank correlation across the reduced set of selected performance measures. Considering all the 1500 stocks in the S&P Index at the end of October 2008, and available over the range January 1990 to October 2008 (226 observations), we compute the rank correlation over 23 performance measures (we drop from the set the Sterling ratio (5%), the VR Index (5%), and the Conditional Sharpe (5%)) on a rolling window of 60 months, obtaining 166 instances of the rank correlation matrix.

Across the performance measures with the highest average rank correlations, some pairs evidence a clear instability. This is the case for Omega with zero threshold and the Sharpe index when computed on deviations from the benchmark index (see Figure 1). Even though the global level of correlation is around 0.90, there are periods where the rank correlation is below 0.70 and periods where it is much higher than 0.90. Furthermore, this behavior does not seem random but shows a clear persistence. Notably, similar behaviors are not observed when considering the same pair of indices but using standard returns or returns in excess to the risk-free investment. A second example is given by considering the rank correlations between the pairs Sharpe-MRAR(2), and Sharpe-MRAR(10). Figure 2 shows that both MRAR(2) and MRAR(10) have low rank correlation with respect to the Sharpe index, but with relatively large changes over time, with a range going from about -0.15 to about 0.15. Similar results have been observed also for other pairs of performance measures. These results show evidence of dynamics in the rank correlations. They also suggest that the use of one single index should be avoided given that, over time, alternative performance measures provide different informative contents which could be relevant for selecting the optimal assets in a more appropriate way.

In addition, we explore the relation between the sample length, the return type and the rank correlation levels. For this purpose, we run simple linear regressions across the rank correlations computed over different combinations of return types and sample periods. Let $R_S\left(X_t, T\right)$ denote the set of rank correlations computed over the returns $X_{i,t}$ $i = 1, 2, ...N$ using a sample of dimension $T = 36, 60, 120$. We consider the cross-sectional linear regressions across all different pairs of $R_S\left(X_t, T\right)$ by varying the return type and the sample dimension. We obtain nine possible sets $R_S\left(X_t, T\right)$ (three return type and three sample size) and 45 regressions of the form $R_S\left(X_t, T\right) = \beta_0 + \beta_1 R'_S\left(X'_t, T'\right) + \varepsilon$ where $R'_S\left(X'_t, T'\right)$ differs from $R_S\left(X_t, T\right)$ either for the return type ($X'_t \neq X_t$), the sample size ($T \neq T'$), or for both. We then compare the R^2 of the regressions and

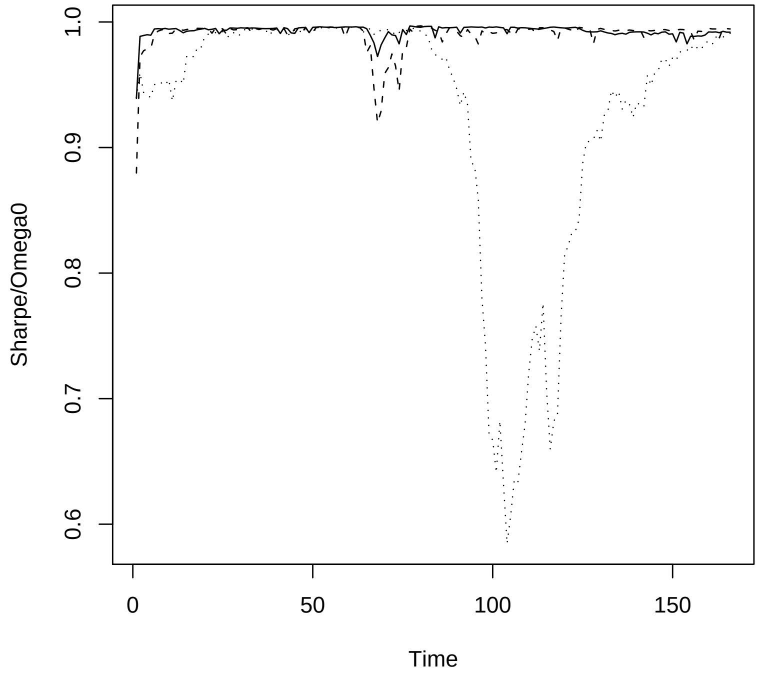

Figure 1: Ratio Sharpe/Omega for returns (full line), excess return (dashed line) and excess returns from benchmark (dotted line).

find that the sample size induces some change over rank correlations computed using the same return type. In fact, when $X'_t = X_t$, the R^2 for the regressions with $T = 120$ and $T' = 60$ are the lowest, reaching a minimum of 0.57, which is still considerable. This is a somewhat expected result given that over shorter intervals the performance measures may be more sensitive to extreme returns. However, interesting observations emerge when comparing the rank correlations computed over the same sample dimensions ($T = T'$) using different return types. In this case, we note that the return type plays an extremely limited role in the evaluation of rank correlations. The R^2 of these regressions range from 0.93 to 0.99, without any clear difference across returns. As a result, we conclude that the choice of returns is not relevant within a selection process of assets (the simple return without any benchmark or risk-free can be used), while the use of at least 60 months could be suggested in order to reduce the impact of extreme returns.

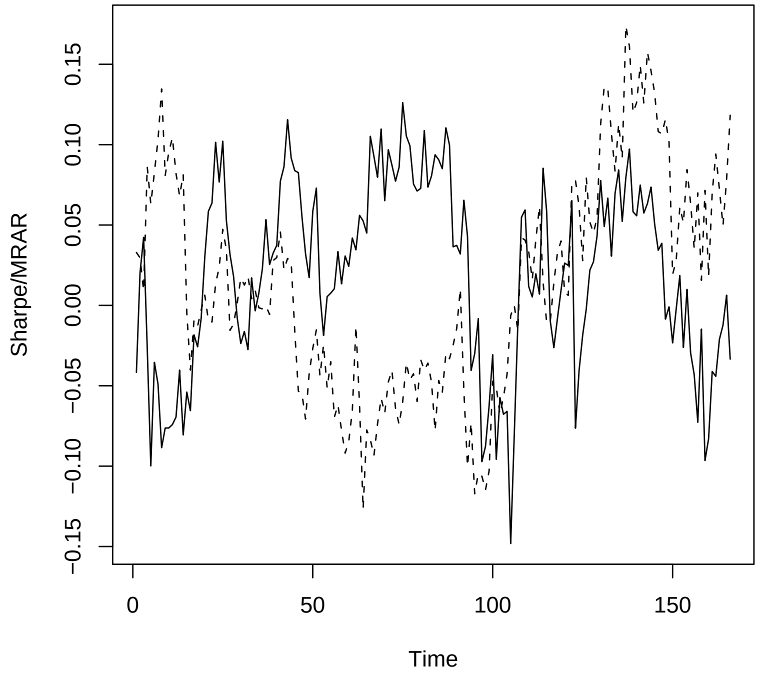

Figure 2: Ratio Sharpe/MRAR(2) (full line) and Sharpe/MRAR(10) (dashed line).

5 A simulated portfolio allocation strategy based on performance ranking

The previous section highlights that, differently from Eling and Schuhmacher (2007), not all performance measures are equivalent either on the basis of a static analysis or, and even more evidently, if we consider a rolling approach. However, such a result leaves some open problems associated with the use of performance measures with different, and maybe complementary, information. What is called for is a method allowing the combining of the selected reward-to-risk ratios. Given that this topic raises a number of additional questions related to the method of combining information and to the optimality criteria of combined measures, we limit our analysis to a very simple example. We refer to the combination of different ranks created by alternative raters (our performance measures). The combination of ranks is an optimal solution under the assumption of independence across raters, (Kendall, 1970). Unfortunately, this is not the case in our setup, because the raters (the performance measures) use the same information set to derive ranks. Being aware of this limitation, our choice of focusing on performance measures with low rank correlations can be considered as a naive approach for controlling and reducing the dependence across raters.

As a result, we propose the use of a composite performance index that, at each point in time, will be obtained as follows. Assume we monitor Q performance measures and, at time t, we want to invest in n assets chosen from a basket of N assets. Denote by P_t the $N \times Q$ matrix containing the values of the performance measures at time t. Starting from P_t we can obtain the $N \times Q$ matrix R_t of the asset ranks. In particular, let us denote by $R_{ij,t}$ the rank of the i-th asset with respect to the j-th measure $(i = 1, ..., N; j = 1, ..., Q)$ and assume that the asset with highest performance measure q has rank 1.
The composite index, $CI_{i,t}$ for the asset i-th at time t is simply the sum of the ranks of asset i

$$CI_{i,t} = \sum_{j=1}^{Q} R_{ij,t}. \tag{29}$$

The CI index can be considered a joint performance measure where the best assets are associated with the lowest values of CI_i. At each month t, CI_t is computed using the last 60 monthly returns for all the available assets. Using the composite index we determine the portfolio allocation for month $t + 1$ choosing the M assets with the highest composite index and using an equally weighted approach. This allocation strategy avoids estimation errors related to the weights (see De Miguel et al., 2009). Finally, we rotate the portfolio monthly.

We consider four possible values of M: 10, 25, 50, and 100 and the three possible return measures in (1): standard returns (R), excess returns from the risk-free (ER), and deviations from the benchmark investment (DB). In our exercise we consider a restricted set of performance measures formed by: the Sharpe ratio; the Burke ratio computed over the 5% of the sample dimension; the VaR ratio at both the 5% and 10% quantiles; the Generalized Rachev ratio with Moderate parameter combination at the 10% quantile level (one single index); the FT Moderate indices under all three threshold levels (or the Omega indices - these are 3 indices); the MRAR index with risk aversion set to 2 and 10; the LAP measures for Defensive and Moderate parameter combinations (another 6 indices). Differently from the restricted set of measures identified in Section 4, the FT and LAP performance measures with Aggressive parameter combinations have been excluded in order to consider the case of a standard investor characterized by moderate risk aversion (these measures are appropriate for investors with low risk aversion).

We consider three benchmarks for our strategy: the risk-free investment, represented by the returns over the 1-month bond index by Citigroup, the S&P 1500 index, and a similar strategy where assets are selected according to the Sharpe ratio. Furthermore, we compare the strategy results over 165 months using the cumulated returns of the strategies, their risk, measured by the standard deviation and Value-at-Risk at the 5% level, and the average monthly turnover of the portfolios.

[TABLE 8]

Table 8 lists the results of our simulations. Using the CI index, cumulated returns are higher than for the strategy based on the Sharpe index in 7 cases out of 12, but without a clear pattern, while the standard deviation is always lower for CI strategies based on R and ER and higher for DB cases. The turnover is lower in all cases but one, where it is close to the Sharpe strategy. We highlight that the turnover reduction is considerable, in particular for the cases with M low, while increasing M, the advantage over the turnover decreases suggesting that the ranks of the two methods are different only for the top performer assets. Differently, if we consider top performers and very good assets, the two methods could provide similar groups. The relevant reduction in turnover will induce a small impact of transaction and operational costs on the strategies based on the CI index. As a result, if transaction costs will be taken into account, strategies average returns, risk measures and performance measures will show a relatively higher preference for CI based allocations.

In addition, the Sharpe index is generally higher for CI-based allocations, with some preference for strategies based on ER and DB rather than the simple R case. The β with respect to the market index is always smaller in the CI-allocation case, suggesting that these strategies induce a reduction of the portfolio exposure to market shocks in all cases. On the contrary, the extra performance with respect to the market index is generally larger with a pattern similar to that of the Sharpe (a preference for ER and DB strategies). Finally, the absolute VaR at the 5% level is in most cases higher for Sharpe based strategies, but without a clear interpretation with respect to N and the return choices. Overall, for low levels of M, the CI index seems to be better than the simpler Sharpe based strategies, while the preference decreases for larger M though still remaining relevant.

6 Conclusions

A typical problem of portfolio management is to select some assets within a large group to build an optimal portfolio. However, sometimes there are not enough degrees of freedom to implement standard mean-variance optimizers, and alternative selection approaches can be used. Within this framework, performance measures could do the task. Nevertheless, a different problem emerges: which measures to use? To answer this question, we followed the approach of Eling and Schuhmacher (2007) and compared performance measures using rank correlations. Within this paper we generalize the study of Eling and Schuhmacher (2007) by enlarging the selection of performance measures compared and exploring the dynamic properties of rank correlations. We show that performance measures based on partial moments and loss aversion are generally different from the traditional ones (including the Sharpe ratio). As an additional finding, we show evidence of changing behavior in rank correlations, even across pairs considered equivalent by Eling and Schuhmacher (2007). Finally, we provide a simple example of the use of such a comparison across performances within a simplified asset allocation exercise. The approach we proposed could be generalized to different datasets, further enriching the set of performance measures. Such extensions are left to future researches as well as the identification of the optimal combination of performance measures. The results of such future contributions could provide relevant tools for quantitative asset allocation.

References

Agarwal, V., Naik, N.Y., 2004, Risk and portfolio decisions involving hedge funds. Review of Financial Studies, 17 (1), 63-98.

Ait-Sahalia, Y., Parker, J. and Yogo, M., 2004, Luxury goods and the equity premium, Journal of Finance, 59 (6), 2959-3004

Avramov, D., and Chordia, T., 2006, Asset pricing models and financial market anomalies, The Review of Financial Studies, 19-3, 1001-1040.

Barberis, N., Hwang, M. and Santos, T., 2001, Prospect theory and asset prices, Quarterly Journal of Economics, 116 (1), 1-53

Bernatzi, S. and Thaler, R.H., 1995, Myopic loss aversion and the equity premium puzzle, Quarterly Journal of Economics, 110 (1), 73-92

Biglova, A., Ortobelli, S., Rachev, S. and Stoyanov, S., 2004, Different approaches to risk estimation in portfolio theory, Journal of Portfolio Management, 31 (1), 103-112

Briner, B.G. and Connor, G., 2008, How much structure is best? A comparison of market model, factor model and unstructured equity covariance matrices, The Journal of Risk, 10 (4), 3-30

Burke, G., 1994, A sharper Sharpe ratio, Futures 23 (3), 56.

Caporin, M., and Lisi, F., 2009, Are portfolio managers really active? An evaluation based on conditional CAPM models, Working paper, Department of Economics, University of Padova

Caporin, M., and Paruolo, P., 2009, Structured Multivariate Volatility Models, Working paper, Department of Economics, University of Padova

Chen, Z., and Wang, Y., 2008, Two-sided coherent risk measures and their application in realistic portfolio optimization, Journal of Banking and Finance, 32, 2667-2673

De Miguel, V., Garlappi, L., and Uppal, R., 2009, Optimal versus naive diversification: how inefficient is the 1/N portfolio strategy?, to appear in Review of Financial Studies.

Dowd, K., 2000, Adjusting for risk: An improved Sharpe ratio. International Review of Economics and Finance, 9 (3), 209–222.

Eling, M., 2008, Does the measure matters in the mutual fund industry?, Financial Analyst Journal, 64 (3), 54-66

Eling, M. and F. Schuhmacher, 2007, Does the choice of performance measure influence the evaluation of hedge funds? Journal of Banking and Finance, 31, 2632-2647

Farinelli, S. and Tibiletti, L., 2003a, Sharpe thinking with asymmetrical preferences, available in SSRN

Farinelli, S. and Tibiletti, L., 2003b, Upside and downside risk with a benchmark, Atlantic Economic Journal, Anthology Section, 31 (4), 387

Farinelli, S., Ferreira, M., Rossello, D., Thoeny, M. and Tibiletti, L., 2008, Beyond Sharpe ratio: optimal asset allocation using different performance ratios, Journal of Banking and Finance, 32, 2057-2063

Farinelli, S., Ferreira, M., Rossello, D., Thoeny, M. and Tibiletti, L., 2009, Optimal asset allocation aid system: from "one-size" vs "taylor-made" performance ratio, European Journal of Operational Research, 192, 209-215

Favre, L. and Galeano, J.A., 2002, Mean-modified Value at Risk optimization with hedge funds, Journal of Alternative Investment, 5 (Fall), 21-25

Ferson, W.E., and Schadt, R., 1996, Measuring fund strategy and performance in changing economic conditions, Journal of Finance, 51, 425-461.

Fisher, R.A., 1915, Frequency distribution of the values of the correlation coefficient in samples of an indefinitely large population, Biometrika, 10, 507-521

Gemmill, G., Hwang, S. and Salmon, M., 2006, Performance measurement with loss aversion, Journal of Asset Management, 7 (3), 190-207

Gregoriou, G.N., Gueyie, J.P., 2003, Risk-adjusted performance of funds of hedge funds using a modified Sharpe ratio. Journal of Wealth Management 6 (Winter), 77–83.

Hwang, S. and Salmon, M., 2002, An analysis of performance measures using copulae, in Knigth, J., and Satchell, S. (eds), Performance measurement in finance: firms, funds and managers, Butterworth-Heinemann Finance, Quantitative Finance Series.

Jensen, M., 1968, The performance of mutual funds in the period 1945–1968, Journal of Finance 23 (2), 389–416.

Kahnemann, D. and Tversky, A., 1979, Prospect theory: an analysis of decision under risk, Econometrica, 47, 263-291

Kaplan, P.D., Knowles, J.A., 2004, Kappa: A Generalized Downside Risk-Adjusted Performance Measure, Morningstar Associates and York Hedge Fund Strategies, January 2004.

Kazemi, H., Schneeweis, T. and Gupta, R., 2004, Omega as a performance measure, Journal of Performance Measurement, 8 (3), 16-25

Kendall, M.G., 1970, Rank correlation methods, Charles Griffin & Co. Ltd., London

Kestner, L.N., 1996, Getting a handle on true performance, Futures 25 (1), 44–46.

Knigth, J., and Satchell, S. (eds), 2002, Performance measurement in finance: firms, funds and managers, Butterworth-Heinemann Finance, Quantitative Finance Series.

Konno, H., 1990, Piecewise linear risk functions and portfolio optimization, Journal of the Operations Research Society of Japan, 33, 139-156

Konno, H. and Yamazaki, H., 1991, Mean-absolute deviation portfolio optimization model and its application to Tokyo stock market, Management Science, 37, 519-531

Ledoit, O., and Wolf, M., 2003, Improved estimation of the covariance matrix of stock returns with an application to portfolio selection, Journal of Empirical Finance, 10, 603-621

Ledoit, O., and Wolf, M., 2004, Honey I shrunk the sample covariance matrix, Journal of Portfolio Management, 30, 110-119

Lintner, J., 1965, The valuation of risky assets and the selection of risky investment in stock portfolios and capital budgets, Review of Economics and Statistics, 47, 13-37

Markowitz, H., 1959, Portfolio selection: efficient diversification of investments, John Wiley

Modigliani, F., Modigliani, L., 1997, Risk-adjusted performance – how to measure it and why. Journal of Portfolio Management 23 (2), 45–54.

Mossin, J., 1969, Security pricing and investment criteria in competitive markets, American Economic Review, 59, 749-756

Ortobelli, S., Rachev, S., Stoyanov, S., Fabozzi, F.J. and Biglova, A., 2005, The proper use of risk measures in portfolio theory, International Journal of Theoretical and Applied Finance, 8 (8), 1107-1133

Pedersen, C.S., Rudholm-Alfvin, T., 2003, Selecting a risk-adjusted shareholder performance measure. Journal of Asset Management 4 (3), 152–172.

Pedersen, C.S., and Satchell, S.E., 2002, On the foundation of performance measures under asymmetric returns, Quantitative Finance, 2 (3), 217-223

Rachev, S., Martin, D. and Siboulet, F., 2003, Phi-alpha optimal portfolios and Extreme Risk Management, Wilmott Magazine of Finance, November, 70-83

Rachev, S., Ortobelli, S., Stoyanov, S., Fabozzi, F.J., and Biglova, A., 2008, Desirable properties of an ideal risk measure in portfolio theory, International Journal of Theoretical and Applied Finance, 11 (1), 19-54

Shadwick, W.F., Keating, C., 2002, A universal performance measure. Journal of Performance Measurement, 6 (3), 59–84.

Sharma, M., 2004, A.I.R.A.P. - Alternative RAPMs for Alternative Investments, Journal of Investment Management, 3 (4)

Sharpe,W.F., 1964, Capital Asset Prices: A Theory of Market Equilibrium Under. Conditions of Risk, Journal of Finance,19, 425-442

Sharpe, W.F., 1966, Mutual fund performance, Journal of Business, 39 (1), 119-138

Sharpe, W.F., 1992, Asset allocation: management style and performance measurement, Journal of Portfolio Management, 18 (2), 7-19

Sharpe, W.F., 1994, The Sharpe Ratio, Journal of Portfolio Management, Fall, 45-58

Sortino, F.A., 2000, Upside-potential ratios vary by investment style, Pension and Investment, 28, 30-35

Sortino, F.A., 2001, Managing downside risk in financial markets, Butterworth-Heinemann Finance, Oxford

Sortino, F.A., van der Meer, R., 1991, Downside risk. Journal of Portfolio Management 17 (Spring), 27–31.

Sortino, F.A., van der Meer, R., Plantinga, A., 1999, The Dutch triangle. Journal of Portfolio Management, 26 (Fall), 50–58.

Treynor, J.L., 1965, How to rate management of investment funds, Harvard Business Review, 43 (1), 63–75.

Tversky, A., and Kahnemann, D., 1992, Advances in prospect theory: cumulative representation of uncertainty, Journal of Risk and Uncertainty, 5, 297-323

Young, T.W., 1991, Calmar ratio: A smoother tool, Futures 20 (1), 40.

Young, M.R., 1998, A MiniMax portfolio selection rule with linear programming solution, Management Science, 44, 673-683

Zakamouline, V., and Koekebakker, S., 2009, Portfolio performance evaluation with generalized Sharpe ratios: beyond the mean and variance, Journal of Banking and Finance, forthcoming

Performance measures (cases)	Returns	Excess returns	Deviations from benchmark
Sharpe ratio	X	X	X
Treynor index	X	X	NA
Appraisal ratio	X	X	NA
Average R over MAD	X	X	X
Average R over MiniMax	X	X	X
Average R over Range	X	X	X
M2	X	NA	NA
Calmar ratio	X	X	X
Sterling ratio (2)	X	X	X
Burke ratio (2)	X	X	X
Sortino ratio	X	X	X
Kappa 3 measure	X	X	X
Farinelli-Tibiletti (18)	X	X	X
Average R over VaR (2)	X	X	X
Average R over ES (2)	X	X	X
VaR ratio (2)	X	X	X
Generalized Rachev Ratios (20)	X	X	X
MRAR (3)	X	X	X
LAP (19)	X	X	X
Total	80	79	77

Table 1: List of performance measures considered

Rank correlations	Returns			Excess Returns			Deviations from benchmark		
Window length	36	60	120	36	60	120	36	60	120
Sterling (5%) and (10%)	1.000	0.997	0.998	1.000	1.000	0.998	1.000	0.999	0.992
Burke (5%) and (10%)	0.997	0.974	0.880	0.996	0.969	0.866	0.995	0.979	0.903
VR index (5%) and (10%)	0.993	0.991	0.985	0.992	0.992	0.990	0.994	0.994	0.987
VaR Ratio (5%) and (10%)	**0.727**	**0.699**	**0.600**	**0.728**	**0.701**	**0.611**	**0.795**	**0.743**	**0.586**
Conditional Sharpe (5%) and (10%)	0.997	0.998	0.997	0.998	0.998	0.998	0.998	0.999	0.998

Table 2: Rank correlations across selected performances measures - drawdowns and quantile based measures.

Rank correlations	Returns			Excess Returns			Deviations from benchmark		
Window length	36	60	120	36	60	120	36	60	120
Average within GR (10%) excluding Aggressive	0.989	0.986	0.982	0.989	0.986	0.982	0.99	0.986	0.981
Average within GR (5%) excluding Aggressive	0.997	0.995	0.992	0.997	0.995	0.992	0.998	0.995	0.991
Average between GR (5%) and GR (10%) excluding Aggressive	0.957	0.955	0.955	0.956	0.955	0.956	0.962	0.956	0.955
Average GR Aggressive (10%) wrt other GR (10%)	0.941	0.898	0.848	0.955	0.912	0.871	**0.772**	**0.699**	**0.788**
Average GR Aggressive (5%) wrt other GR (5%)	**-0.474**	**0.233**	0.946	**-0.489**	**0.150**	0.937	**0.047**	**0.660**	0.954
GR Aggressive (5%) - GR Aggressive (10%)	**-0.452**	**0.192**	0.853	**-0.459**	**0.128**	0.868	**-0.072**	**0.415**	**0.810**

Table 3: Rank correlations across selected performances measures - Generalized Rachev Ratios. The first and second rows report the average rank correlation across the Generalized Rachev ratios for the parameter combinations associated to Moderate, Conservative, Growth and Defensive investors with a given quantile level. The third row reports the average rank correlation between the two groups associated to the first and second row. The other three rows report the average rank correlation between the Generalized Rachev ratios for Aggressive investors with respect to other indices
.

Rank correlations	Returns			Excess Returns			Deviations from benchmark		
Window length	36	60	120	36	60	120	36	60	120
Within and Between UPR, Defensive, Conservative, Moderate and Growth									
Within -0.02	0.941	0.944	0.914	0.938	0.942	0.907	0.956	0.960	0.927
Within 0	0.915	0.919	0.871	0.912	0.917	0.872	0.928	0.926	0.882
Within 0.02	0.915	0.926	0.922	0.918	0.929	0.931	0.902	0.907	0.915
Between -0.02 and 0	0.865	0.843	**0.716**	0.848	0.825	**0.682**	0.881	0.843	**0.725**
Between -0.02 and 0.02	**0.557**	**0.428**	**0.147**	**0.534**	**0.406**	**0.140**	**0.581**	**0.414**	**0.213**
Between 0 and 0.02	**0.790**	**0.736**	**0.649**	**0.794**	**0.744**	**0.682**	**0.798**	**0.735**	**0.701**
Within Aggressive and Between Aggressive and Defensive, Conservative, Moderate and Growth									
Within -0.02	**-0.268**	**0.045**	**0.498**	**-0.365**	**-0.100**	**0.356**	**-0.111**	**0.067**	**0.470**
Within 0	**0.383**	**0.05**	**-0.108**	**0.532**	**0.255**	**0.163**	**0.014**	**0.027**	**-0.011**
Within 0.02	0.857	0.853	**0.819**	0.874	0.864	0.840	0.852	0.847	0.823
Between -0.02 and 0	**0.058**	**0.025**	**0.103**	**0.086**	**0.057**	**0.152**	**-0.047**	**0.025**	**0.146**
Between -0.02 and 0.02	**0.125**	**0.104**	**-0.029**	**0.112**	**0.073**	**-0.041**	**0.171**	**0.09**	**0.018**
Between 0 and 0.02	**0.513**	**0.349**	**0.242**	**0.574**	**0.421**	**0.333**	**0.386**	**0.346**	**0.269**

Table 4: Rank correlations across selected performances measures - Farinelli-Tibiletti Ratios. We considered two groups composed by the parameter combinations associated to Defensive, Conservative, Moderate and Growth investors (in all possible pairs) on the one side, while on the other side the pairs of performance measures involving at least one measure associated to Aggressive investors. "Within" stands for the average of rank correlations across the pairs of measures with the same minimum acceptable return. "Between" stands for the average rank correlations across the pairs of measures with two different minimum acceptable returns. As an example, the line "Within -0.02" in the first group, contains the averages of rank correlations over the following pairs: UPR-Defensive, UPR-Conservative, UPR-Moderate, UPR-Growth, Defensive-Conservative, Defensive-Moderate, Defensive-Growth, Conservative-Moderate, Conservative-Growth, Moderate-Growth
.

Rank correlations	Returns			Excess Returns			Deviations from benchmark		
Window length	36	60	120	36	60	120	36	60	120
MRAR 2 - MRAR 10	**0.252**	**-0.076**	**-0.006**	**0.462**	**0.068**	**0.256**	**0.021**	**-0.124**	**0.024**
MRAR 2 - MRAR 50	**0.168**	**-0.132**	**0.003**	**0.404**	**0.047**	**0.260**	**-0.153**	**-0.160**	**0.052**
MRAR 10 - MRAR 50	0.894	0.858	0.955	0.940	0.922	0.974	**0.707**	**0.806**	0.97
LAP - Within HS	0.951	0.939	**0.685**	0.950	0.938	**0.685**	0.932	0.932	**0.715**
LAP - Within Defensive	**-0.067**	**-0.054**	**0.081**	**-0.052**	**-0.046**	**0.074**	**-0.101**	**-0.069**	**0.052**
LAP - Within Conservative	0.813	**0.759**	**0.534**	0.853	**0.776**	**0.510**	**0.777**	**0.767**	**0.494**
LAP - Within Moderate	0.958	0.927	**0.670**	0.957	0.929	**0.665**	0.944	0.924	**0.666**
LAP - Within Growth	0.897	0.864	**0.432**	0.898	0.861	**0.446**	0.850	0.837	**0.469**
LAP - Within Aggressive	**0.523**	**0.591**	**0.508**	**0.477**	**0.580**	**0.520**	**0.408**	**0.514**	**0.517**
LAP - Between HS-Defensive	**0.081**	**0.027**	**-0.121**	**0.078**	**0.012**	**-0.125**	**0.061**	**-0.008**	**-0.145**
LAP - Between HS-Conservative	**0.741**	**0.698**	**0.154**	**0.751**	**0.703**	**0.123**	**0.684**	**0.634**	**0.064**
LAP - Between HS-Moderate	0.948	0.928	**0.671**	0.947	0.929	**0.669**	0.935	0.925	**0.686**
LAP - Between HS-Growth	0.835	0.838	**0.562**	0.831	0.837	**0.567**	**0.787**	**0.807**	**0.590**
LAP - Between HS-Aggressive	**0.339**	**0.434**	**0.477**	**0.322**	**0.439**	**0.506**	**0.240**	**0.367**	**0.531**
LAP - Between Defensive-Conservative	**0.156**	**0.104**	**0.173**	**0.141**	**0.097**	**0.161**	**0.100**	**0.071**	**0.156**
LAP - Between Defensive-Moderate	**0.108**	**0.059**	**-0.023**	**0.103**	**0.045**	**-0.031**	**0.076**	**0.013**	**-0.061**
LAP - Between Defensive-Growth	**0.052**	**0.008**	**-0.042**	**0.044**	**-0.003**	**-0.048**	**0.036**	**-0.021**	**-0.060**
LAP - Between Defensive-Aggressive	**-0.051**	**-0.074**	**-0.145**	**-0.055**	**-0.077**	**-0.146**	**-0.037**	**-0.078**	**-0.139**
LAP - Between Conservative-Moderate	0.817	**0.783**	**0.417**	0.832	**0.791**	**0.388**	**0.769**	**0.738**	**0.326**
LAP - Between Conservative-Growth	**0.734**	**0.672**	**0.253**	**0.745**	**0.679**	**0.237**	**0.673**	**0.623**	**0.205**
LAP - Between Conservative-Aggressive	**0.128**	**0.134**	**-0.179**	**0.127**	**0.146**	**-0.177**	**0.028**	**0.043**	**-0.199**
LAP - Between Moderate-Growth	0.831	**0.819**	**0.536**	0.824	**0.817**	**0.541**	**0.787**	**0.791**	**0.568**
LAP - Between Moderate-Aggressive	**0.264**	**0.333**	**0.257**	**0.249**	**0.341**	**0.284**	**0.174**	**0.277**	**0.319**
LAP - Between Growth-Aggressive	**0.485**	**0.556**	**0.442**	**0.463**	**0.551**	**0.461**	**0.392**	**0.481**	**0.467**

Table 5: Rank correlations across selected performances measures - utility based performance measures. We separately consider the MRAR measures and the Loss Aversion Performance measures. The last are grouped depending on their parameters in Hwang-Satchell (HS), Defensive, Conservative, Moderate, Growth and Aggressive. Apart the Hwang-Satchell case, the groups do not include the LAP^S measure which is equivalent to the FT Moderate Index with threshold set at zero. For MRAR measures we report the rank correlation coefficients. For LAP groups we report the average rank correlation within each group and between each pair of groups
.

Rank correlations	Returns			Excess Returns			Deviations from benchmark		
Window length	36	60	120	36	60	120	36	60	120
Treynor	0.950	0.934	0.858	0.945	0.940	0.906	—	—	—
Appraisal ratio	**0.792**	0.940	0.906	**0.555**	0.915	0.893	—	—	—
ERMAD	0.999	0.999	0.997	0.999	0.999	0.998	0.999	0.999	0.998
ERR	0.995	0.992	0.966	0.994	0.994	0.985	0.994	0.989	0.978
ERMM	0.991	0.987	0.956	0.990	0.990	0.978	0.992	0.986	0.969
M2	**0.719**	0.928	0.964	—	—	—	—	—	—

Table 6: rank correlation of traditional and similar performance measures with the Sharpe ratio over different sample length.

Rank correlations	Returns			Excess Returns			Deviations from benchmark		
Window length	36	60	120	36	60	120	36	60	120
Calmar ratio	0.983	0.977	0.930	0.982	0.978	0.965	0.987	0.977	0.946
Sterling ratio (5%)	0.982	0.978	0.932	**0.785**	0.903	0.912	0.833	0.951	0.911
VR Index (5%)	0.993	0.991	0.983	0.993	0.993	0.992	0.994	0.994	0.990
Conditional Sharpe (5%)	0.996	0.994	0.983	0.997	0.996	0.994	0.995	0.996	0.993
Omega	**0.723**	0.933	0.991	**0.532**	0.866	0.991	0.865	0.907	0.993

Table 7: Rank correlation of selected measures with the Sharpe ratio.

CI	M=10			M=25			M=50			M=100				
allocation	R	ER	DB	R	ER	DB	R	ER	DB	R	ER	DB	S&P1500	RF
Mean	0.019	0.016	0.016	0.018	0.016	0.017	0.017	0.016	0.016	0.017	0.016	0.014	0.005	0.003
St.dev.	0.059	0.057	0.064	0.059	0.053	0.057	0.056	0.052	0.054	0.056	0.054	0.053	0.044	0.001
Sharpe	0.276	0.234	0.202	0.253	0.253	0.253	0.256	0.256	0.245	0.254	0.245	0.211	0.055	—
Alpha	0.014	0.012	0.011	0.013	0.011	0.012	0.012	0.011	0.011	0.012	0.011	0.009	—	—
Beta	0.780	0.693	0.882	0.936	0.776	0.856	0.913	0.834	0.845	0.944	0.895	0.883	—	—
R^2	0.344	0.291	0.375	0.497	0.426	0.438	0.515	0.497	0.483	0.553	0.549	0.540	—	—
VaR(5%)	-0.080	-0.072	-0.066	-0.076	-0.069	-0.072	-0.074	-0.072	-0.061	-0.065	-0.067	-0.060	-0.079	—
Ave.Turnover	0.399	0.408	0.426	0.321	0.327	0.342	0.264	0.273	0.272	0.201	0.205	0.206	—	—
Cumulated	16.779	10.153	8.815	13.045	10.644	12.408	12.349	10.737	10.234	12.06	10.09	7.157	1.078	0.641
Sharpe	M=10			M=25			M=50			M=100				
allocation	R	ER	DB	R	ER	DB	R	ER	DB	R	ER	DB	S&P1500	RF
Mean	0.016	0.017	0.015	0.019	0.015	0.015	0.019	0.017	0.014	0.018	0.016	0.014	0.005	0.003
St.dev.	0.070	0.065	0.058	0.062	0.055	0.052	0.061	0.054	0.052	0.059	0.056	0.053	0.044	0.001
Sharpe	0.182	0.221	0.211	0.259	0.222	0.229	0.257	0.250	0.209	0.251	0.233	0.210	0.055	—
Alpha	0.010	0.012	0.010	0.014	0.010	0.010	0.013	0.011	0.009	0.012	0.011	0.009	—	—
Beta	0.992	0.916	0.961	0.982	0.884	0.951	0.983	0.925	0.945	1.000	0.959	0.951	—	—
R^2	0.401	0.390	0.537	0.496	0.513	0.651	0.510	0.573	0.645	0.560	0.585	0.642	—	—
VaR(5%)	-0.112	-0.090	-0.081	-0.083	-0.066	-0.069	-0.074	-0.064	-0.066	-0.071	-0.071	-0.067	-0.079	—
Ave.Turnover	0.817	0.884	0.921	0.588	0.725	0.809	0.372	0.516	0.644	0.194	0.285	0.407	—	—
Cumulated	7.705	11.122	8.201	15.356	8.331	8.246	14.634	10.763	6.806	12.96	9.509	6.984	1.078	0.641

Table 8: Composite Index and Sharpe based simulated allocation

.