AESTIMATIO, THE IEB INTERNATIONAL JOURNAL OF FINANCE, 2011. 3: 2-23

Applications of State Contingent Stochastic Ordering Methods to the Clustering and Performance Measurement of Trading Strategies

Glaffig, Clemens H.

▸ RECEIVED: 22 DECEMBER 2010
▸ ACCEPTED: 27 MARCH 2011

Abstract

The rise in popularity of benchmark free and complex trading strategies throughout the last decade has made available a large variety of risk and performance profiles. As a consequence, to account for their complex performance characteristics, a lot of effort has been devoted to classify and value the performance of these strategies by the alterations of previous – or innovative measures. However, as most measures are often still simple path – and context independent statistics, most often the information provided proves inadequate to separate performance characteristics – as evidenced by the latest crisis.

This paper provides a methodology that integrates the clustering and performance measurement of trading strategies in a context and preference based environment. It decomposes preferred performance characteristics into fragments of context dependent behaviour for clustering purposes. It subsequently aggregates these fragments of performance characteristics into a performance measure. The methodology allows for consideration of path dependencies. Two applications, in the clustering of hedge fund styles and the ordering of alternative equity strategies are given. A further application in the statistical replication of trading strategies is highlighted.

Keywords:

Clustering, Performance measure, Stochastic dominance, State contingent preference.

JEL classification:

C10, C14, C18, C51, C65, G10.

Glaffig, Clemens H. ✉ *Panathea Capital Partners GmbH & Co. KG, Grünwälderstr. 1- 7. Freiburg, Germany.*
✆ *+49 761 3894 8961 Fax: +40 761 3894 8969. E-mail: Clemens.glaffig@panathea.de*

1. Introduction

The selection of trading strategies or funds typically is a two stage process: The first stage defines an admissible subset of the universe of available strategies, often by characterising the preferred strategy styles – like global macro, long-short equity, emerging market equity etc.– possibly further constrained by risk preferences, market perception and regulatory matters. The second stage evaluates and orders the strategies of the chosen subset, for which one or several measures of performance are used.

In this paper, we provide a framework for an integrated approach to clustering and ordering strategies, which is based on state contingent investor preferences.

Typically, grouping or clustering strategies according to their style characteristics is highly subjective and prone to abuse – depending on the flavour of the moment, a global macro strategy very quickly becomes a Futures Fund, commonly referred to as commodity trading advisor (CTA) or a "special situations" strategy. Consequently, each style cluster exhibits vastly different risk and return characteristics amongst its constituents. A more useful way of clustering would be according to preferred performance characteristics, as measured by an appropriate technique. Common performance measures are inadequate to express detailed characteristics, the perception of which varies with investors. Moreover, investor perception of perform ance quality is generally context, i.e. market dependent.

Early performance measures like the Sharpe ratio were based on the mean/variance framework, sufficient for the benchmark driven, long only world, depending on market instruments reflecting risk factors in a linear way. This framework proved insufficient for the world of alternative strategies, with their active trading and use of complex derivative instruments, producing nonlinear risk factor dependencies. As has been shown in several papers (see e.g. Spurgin, 2001, and Goetzmann *et al.*, 2002), the Sharpe ratio can easily be manipulated by employing or supplementing asymmetric return strategies with positive carry trades, like option writing strategies.

In light of this evidence, a number of improvements and innovative performance measures were developed: Amongst those, Kazemi *et al.* (2003) present a revised Sharpe ratio, addressing the issue of non normality of strategy returns. Further variations of the Sharpe ratio were developed, like the Generalised Sharpe Ratio, Sortino ratio, gain-loss ratio or performance indices based on distortions of the return distribution, see Cherny *et al.*(2009), Cherny and Madan (2009) and Eberlein and Madan (2009), all addressing and capturing various aspects of return characteristics of alternative strategies, not captured by the original measures. Many more were developed, for a comprehensive summary and discussion see Cogneau and Hubner (2009).

Still, most often these measures are simple statistics of the return distribution, often in the form of reward to risk ratios with varying concepts of risk. Common to most performance measures is the aversion to risk, with little or no allowance for risk tolerance (see e.g. Zakamouline and Koekebakker, 2009) for a utility based approach which includes preferences for higher moments of the return distribution, allowing for a symmetric treatment of risk aversion and tolerance). However, as reality has taught, assessing risk via a single number is dangerous. In the same way, judging on a strategy's performance based on a single isolated statistic is meaningless. In addition, most performance measures are path independent, contrary to investor perception of performance quality: A manager producing 50 times a -1% return followed by 50 times a +1% return produces the same Sharpe ratio and VaR as a manager showing a sequence of 50 times a -1% loss immediately followed by a 1% gain – in reality investor's perception of - and reactions to these performances would be drastically different.

In our integrated framework, we cluster trading strategies based on preferred return characteristics and order them by attaching scores as performance measures to cluster points. To make the approach context and investor dependent, we revisit the concept of state contingent utility theory, approaching the subject from a different angle than historically done. The main issues we would like to be reflected in this framework are:

- Performance characteristics have to be assessed in context, within a market environment or relative to a market benchmark. Integrating contextual dependencies at the outset will produce more robust results than comparing context independent measures ex post.
- Path dependencies of return characteristics should be reflected.
- Investor preferences, specifically context dependent risk aversion and tolerance should be included. This is more naturally achieved by working with the full return distribution rather than a statistic of it.

The first point is addressed by relating return characteristics to a benchmark, which could be a performance benchmark or a market-state-indicator. The second point is reflected by working with multiple time scales. The third point will be addressed by reflecting investor preferences via a contingent target distribution, contingent on the benchmark state. To make use of the information provided by the full return distribution, we will use a version of weak integral stochastic dominance.

Classical n-th order stochastic dominance has been well studied, as has its equivalent formulation in terms of utility theory. Also, contextual dependencies have been reflected to some degree by considering state dependent utility functions. However, n-th order stochastic dominance ("*sd*") or utility preference over large classes of utility functions is too restrictive to separate different strategies for clustering purposes and to imply a full ordering of the strategies. Specifically, first and second order *sd* are very sensitive to sampling errors and outliers, complicating practical applications. Moreover, the very definition of *sd* or the restriction to concave utility functions incipiently overemphasizes risk aversion in all circumstances.

More general integral stochastic orders have been applied in several financial applications; see e.g. Müller (1997) and Müller and Stoyan (2002) for a general introduction and e.g. Rüschendorf (2005) and Mainik and Rüschendorf (2010) and Levy (2006) for more specific financial applications. Classical stochastic dominance has also been studied to evaluate performance of, e.g., portfolio insurance strategies see Annaert *et al.* (2009) and hedge fund performance, see e.g. Li and Linton (2007).

We differ from standard stochastic dominance in the following way: While we use a class of state contingent target distributions, we do not demand the validity of an ordering condition valid for all members of the class to achieve dominance, but view the set of dominances over individual target class members for clustering purposes and impose an ordering by applying a functional to that set. The ordering thus obtained is weaker than the standard integral stochastic ordering over the target class.

The paper is organised as follows. Section2 will provide the main definitions with some examples for the selection of benchmarks and targets. Section 3 will provide some properties and decomposes the dominance here employed into its main sources. Section 4 will exhibit applications to both clustering and ordering. Section 5 concludes.

2. Definitions

We will investigate the time series of strategies for which we use samples of the total daily returns.

By a sample, we mean any n-fold aggregate of daily returns, to reflect n-day returns. Using returns over number of different horizonts throughout the analysis allows incorporating path dependencies of the performance characteristics of a strategy and judge on their evolution.

To express investor preferences and separate out individual performance characteristics, we will compare the distribution of a strategy sample X to a set Ω of customised, predefined target distributions. A specific target distribution reflects a specific preference of performance characteristic. Preferences are thus described by a full distribution against which the full distribution of a given strategy is compared.

Let $F_X(x)$ and $F_T(x)$ denote the cumulative distribution functions of a strategy X and a target T respectively.

To start out, we first express state independent preferences via state independent, unconditional target distributions:

Definition 1

Let Ω be a family of one dimensional target distributions, reflecting the specific investor performance preferences with $|\Omega|=N$. We call the N-dimensional vector

$$\mathcal{D}_\Omega(X)=\left\{D_{T_n}(X):=\int_{-\infty}^{\infty}(F_{T_n}(x)-F_X(x))dF_{T_n}(x),\ T_n\in\Omega,\ n=1,2...N\right\}\in R^N$$

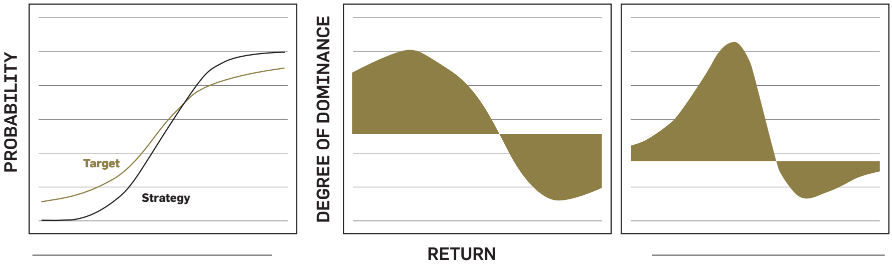
the unconditional dominance of X over the target class Ω. $D_{T_n}(X)$ is called the dominance of X over the target T_n.

We call the union of all the ranges of $\mathcal{D}_\Omega(X)$ over all available strategies the cluster space for Ω.

Notes: While we will concentrate on the targets T being random variables, we could generalize the definition by allowing generalized distributions *s.th.* T acts on the distribution of X to include cases in which T has the same moments of specified order as the strategy X it is compared to.

Figure 1 illustrates Definition 1: The left graph shows the cumulative distribution functions of a target and a strategy. The middle graph shows the un-weighted difference $F_T(x)$-$F_X(x)$. The right graph shows the weighted difference $(F_T(x)$-$F_X(x))dF_T(x)$. The (signed) area of the shaded region corresponds to $D_T(X)$, which in this example is greater than 0. The right graph conveys more information on the specific regions of greatest contribution to dominance than the simple number $D_T(X)$. Specifically, in this example, while T dominates X for larger returns, the X dominance for lower returns outweighs. This is a typical picture of a strategy with less leverage than - and positive alpha relative to the target.

Figure 1. **Left: Example of specific cdf's of a target and a strategy; Middle: Un-weighted difference of cdf's; Right: Weighted difference of cdf's.**



Examples of Specific Target Distributions:

- Let the distribution of T be defined by $dF(x)=\delta_{v_0}(x)$, the delta function centred at v_0, defined by $\delta_{v_0}(x):=\lim_{\varepsilon\to 0}\frac{1}{\sqrt{2\pi\varepsilon^2}}\exp(-\frac{(x-v_0)^2}{2\varepsilon^2})$. Then $D_T(X)=\frac{1}{2}-F_X(v_0)$, which is positive iff $median(X)>v_0$. Also, if $F_X(X)$ admits a density, $VaR_{\frac{1}{2}-D_T(X)}(X)=v_0$, i.e. the VaR of X at level $\frac{1}{2}-D_T(X)$ is v_0.

- Let $dF(x)=\frac{1}{2}\delta_{v_0}(x)+\frac{1}{2}\delta_{-v_0}(x)$, then $2D_T(X)=1-F_X(v_0)-F_X(-v_0)=P(X\geq v_0)-P(X<-v_0)$, which can serve as a measure of skew.

Given investor's preferences are typically context dependent, we add one more feature to our consideration: State contingency. It allows, e.g., to integrate different attitudes towards risk and outperformance expectations in different market environments.

We denote by B a benchmark or market-state-indicator, representing the "market" which will separate scenarios. Let $\Omega(B)$ be a family of conditional targets, represented by their conditional distribution function, conditioned on B, i.e. $T|B\in\Omega(B)$ will be identified with $dF_{T|B}(t|b)$. The distribution of B, $dF_B(b)$can represent a real benchmark distribution or a distribution weighing the importance of specific benchmark states to an investor. Again, we assume $|\Omega(B)|=N$

Definition 2

We call $\mathcal{D}_{\Omega(B)}(X)=\left\{E_B(D_{T_n|B}(X|B)),\ T_n|B\in\Omega(B), n=1,2,...N\right\}$

$$=\left\{\int_{-\infty}^{\infty}D_{T_n|B}(X|B)dF_B(b), T_n|B\in\Omega(B), n=1,2,...N\right\}$$

$$=\left\{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(F_{T_n|B}(x|b)-F_{X|B}(x|b))dF_{T_n|B}(x|b))dF_B(b), n=1,2,...,N\right\}\in\boldsymbol{R}^N$$

the state contingent dominance vector of X over the performance target class $\Omega(B)$ contingent on the benchmark B.

We call $E_B(D_{T|B}(X|B))$ the state contingent dominance of X over the target T, contingent on the benchmark B.

For ease of notation, we will write $D_{T|B}(X|B)$ instead of $E_B(D_{T|B}(X|B))$, so that $D_{T|B}(X|B)=\mathcal{D}_{\Omega(B)}(X)$ for $\Omega(B)=\{T|B\}$.

Again, as before, if $dF_{T|B}(t|b)$ and $dF_{X|B}(x|b)$ admit densities, $D_{T_n|B}(X|B)$ can be expressed as

$$\left\{\frac{1}{2}-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}F_{X|B}(x|b)dF_{T_n|B}(x|b)dF_B(b), n=1,2,...N\right\}$$

$$=\left\{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(F_{T_n|B}(x|b)dF_{X|B}(x|b)dF_B(b)-\frac{1}{2}, n=1,2,...,N\right\}$$

$$=\frac{1}{2}-E_{T,B}(F_{X|B})=E_{X,B}(F_{T|B})-\frac{1}{2} \qquad (*)$$

where $E_{T,B}$ and $E_{X,B}$ denote expectation under the joint distribution of T and B and X and B respectively.

Examples:

- Assume $T|B = B+\alpha+\varepsilon$, for constant α and ε normally distributed, independent of B with zero mean. Then $D_{T|B}(X|B)$ expresses to what extent X dominates randomness above the benchmark outperformance of α.

- Assume the case of "active equity" with target $T|B=\beta(B)B$, with $\beta(B)=\beta_+ 1_{(B>0)}+\beta_- 1_{(B<0)}, \beta_+>\beta_-$. Then $D_{T|B}(X)$ measures the dominance relative to an active equity benchmark.

Note: If any of the target distributions T_n is independent of B, then $D_{T_n|B}(X|B)=D_{T_n}(X)$, i.e., the state contingent dominance becomes unconditional dominance.

In some applications, as seen in section 4, it is opportune to estimate the conditional distribution $F_{X|B}(x|b)$ using the empirical marginals for X and B and fitting a copula. In this context the following equivalent version of Definition 2 will be beneficial (for brevity, we assume the conditional target distribution admits a density:

$$\mathcal{D}_{\Omega(B)}(X)=\left\{ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(F_{T_n|B}(x|b)c_{X,B}(F_X(x),F_B(b))dF_X(x)dF_B(b)-\frac{1}{2}, T_n\in\Omega(B)\right\}$$

where $c_{X,B}(F_X(x),F_B(b))$ is the copula density of X and B.

A note on the connections to utility theory:
As mentioned in the introduction, classical n-th order stochastic dominance has an equivalent formulation in terms of utility theory, whereby statements on n-th order stochastic dominance correspond to statements on whole classes of utility functions, see, e.g., Levy, (2006). While utility theory considers abstract classes of utility functions and focuses on general and absolute preferences like aspects of risk aversion, we emphasize bespoke preferences. Hence there is no such correspondence between our approach to ordering and general classes of utility functions as for classical stochastic dominance. Nevertheless, a connection to utility theory can be constructed: If we assume the target T to be independent of $B, T|B=T$, and set $u_T(x):=\left(F_T(x)-\frac{1}{2}\right)$, it follows from (*), that dominance becomes expected utility with utility function $u_T(x)$, i.e. $D_T(X)=E_X(u_T)$. However, as one difference to commonly used utility functions, $u_T(x)$ would be S-shaped for a large class of target distributions and would hence, e.g., not belong to the *CARA* or *HARA* class, reflecting aspects of prospect theory instead. If T has a normal distribution, then the Arrow Pratt risk aversion is positive (i.e. risk averse) to the right - and negative (i.e. risk tolerant) to the left of the median of T (see also Meucci, 2007, p. 270 ff).

Overall, our approach can be viewed as a way to construct specific state contingent utility functions with a more detailed interpretation of functional properties. As an example, the zero-intercept of $u_T(x)$, would not only serve as the separator between negative and positive utility, but, being the median of T, as a benchmark for relative skew (see section 3.2). Also, the local convexity or concavity of a utility function would not only be a measure of risk aversion or tolerance but, in specific cases, an expression of stop loss preferences.

We feel though, that expressing state contingent preferences via distributions is more intuitive: Definitions 1 & 2 express preferences not by the individual utility of returns, but by specifying in which way targeted returns scatter around a benchmark or in what way a strategy ought to distribute and behave in a specific state of the market.

To cluster strategies that are similar in a statistical sense relative to our preferences as expressed by the unconditional and state contingent target sets, we define a semi-metric which allows weighing different preferences:

Definition 3

The semi-distance $D_{\boldsymbol{\Sigma}}$ between two strategies X and Y is defined via the Mahalanobis distance between $\mathcal{D}_{\Omega(B)}(X)$ and $\mathcal{D}_{\Omega(B)}(Y)$:

$$D_{\boldsymbol{\Sigma}}(X,Y) := \sqrt{(\mathcal{D}_{\Omega(B)}(X) - \mathcal{D}_{\Omega(B)}(Y))'\Sigma^{-1}(\mathcal{D}_{\Omega(B)}(X) - \mathcal{D}_{\Omega(B)}(Y))}$$

for a positive, symmetric NxN matrix Σ, expressing the weights attached to different preferences.

Note: The semi distance of Definition 3 is identical to considering standard Euclidean distance for the cluster-preference-weighted vectors $\mathcal{D}^*_{\Omega(B)}(X) = \Sigma^{-1/2}\mathcal{D}_{\Omega(B)}(X)$.

$D_{\boldsymbol{\Sigma}}(X,Y)=0$ does in general not imply $X=Y$, only that X and Y are indistinguishable relative to our stated preferences. Hence $D_{\boldsymbol{\Sigma}}$ does not define a metric on R^N. In the applications in section 4, we will use the semi distance $D_{\boldsymbol{\Sigma}}$ to cluster strategies. The origin in cluster space corresponds to the set of strategies with complete indifference w.r. to the set of preferred performance characteristics.

To evaluate the performance characteristics of a given strategy X, we map the cluster point $\mathcal{D}_{\Omega(B)}(X)$ to the real numbers and call that map a performance measure. In this way, we build a performance measure from fragments of preferred performance characteristics, each fragment a $D_{T|B)}(X|B)$ value for $T|B \in \Omega(B)$.

Definition 4

A performance measure P over the target class $\Omega(B)$ is a map from $\mathbf{R}^{|\Omega(B)|} \rightarrow \mathbf{R}$, satisfying:

(i) P is strictly non decreasing in each coordinate

(ii) $P(0) = 0$

We will say that X P-outperforms Y under preference $\Omega(B)$, if $P(\mathcal{D}_{\Omega(B)}(X)) > P(\mathcal{D}_{\Omega(B)}(Y))$.

Examples: Let $d_n = d_n(X) := D_{T_n}|B(X|B)$, $T_n|B \in \Omega(B)$,

- $P(d_n, =1,2,\ldots,N) := \inf_n (d_n(X))$
- $P(d_n, =1,2,\ldots,N) := \sum_n \omega_n d_n$ with $\sum_n \omega_n = 1$ and $\omega_n \geq 0, \forall n$.
- Assume $\{T_n|B \in \Omega(B)\}$ constitutes an ordered set of targets, ordered by some distributional property in the sense that $\{X: d_n(X)>0\} \subset \{X : d_m(X)>0\}, \forall n,m; n>m$ (e.g. identical mean and decreasing variance). Then $P(d_n, =1,2,\ldots,N) := \max\{n: d_n>0\}$ constitutes an index of performance similar to the acceptance index in Cherny *et al.* (2009), Cherny and Madan (2009) and Eberlein and Madan (2009).

P gives rise to indifference curves in n-space, along which strategies produce the same performance measure relative to $\Omega(B)$. We will explore more specific examples in Section 4.

An application which is beyond the scope of this paper is strategy replication. We will here only briefly comment on it:

A remark on preference based, state contingent strategy and hedge fund replication: Hedge fund replication has been well studied during the last decade. The two main techniques are factor based- and distributional replication. The stochastic ordering methods described in this study give rise to a new hybrid approach to hedge fund replication, which we refer to as dominance replication. It combines the intuitive construction of a replicating strategy as given by factor replication with the objective to replicate state contingent characteristics of the return distribution. By matching dominance vectors for predefined sets of target distributions, it replicates preference based distributional performance characteristics to arbitrary fine detail. Dominance replication is less ambitious in its objective than factor based replication, reducing the risk of over fitting. Its objective is concerned with distributional properties, it is however more general than classical distributional replication as introduced by Kat and Palaro (2005).

More specifically, let Y be the strategy we wish to replicate. If we choose a set a of available trading strategies, a replication, which will reflect our preferences w.r. to performance characteristics and weighting, is given by: $X^{*}=\underset{X\in a}{arg\,min}\,(D_{\Sigma}\,(X,\,Y))$ for given $\Omega(B)$. The set a could be given by: $X\in a$, $X=\Sigma\alpha_i S_i$, where $\{S_i\}$ is a set of predefined asset based risk factors and $\alpha_i=\alpha_i(B=b)$ is a state contingent (potentially rule based) allocation to S_i. More details to dominance replication will be given in a forthcoming study.

3. Properties and Sources of Dominance

3.1. Properties

In the following, we will assume that all target distributions admit densities. In the applications we will only consider target distributions that can be approximated by absolutely continuous distributions *s.th.* all of the following considerations hold. We will denote the median of X and T by m_X and m_T respectively.

The following is a list of simple properties that can be useful in the construction of target distributions. For ease of notation, we will drop the B-dependence.

(i) Sensibility: If $X>Y$ in all scenarios $\Rightarrow F_X(x)<F_Y(x),\forall x$ from which $D_T(X)>D_T(Y),\forall T$ follows directly from the definition.
(ii) $D_T(X)=-D_X(T)$.
(iii) For fixed c, $D_T(X+c)=D_{T-c}(X)$.
(iv) If T is constant, $D_T(X)>0$ if and only if $m_X>T$.
(v) Scaling with leverage: For $\lambda>0, D_T(\lambda X)=D_{T/\lambda}(X)$. For $\lambda<0, D_T(\lambda X)=-D_{T/\lambda}(X)$. Specifically, $D_T(-X)=-D_{-T}(X)$. If the density of T is scale invariant, i.e. $\lambda F'_T(\lambda x)=F'_T(x),\ \lambda>0,\forall x$, then $D_T(\lambda X)=D_T(X)$.

Properties (i) to (v) follow straightforward from the definition, using a simple change of variable in the integration. Property (v) stresses, that $D_T(X)$ is in general not scale independent as are a large number of performance measures. The preference towards risk as expressed by T will determine whether leverage is penalised, rewarded or $D_T(X)$ is indifferent w.r. to it.

3.2 Sources of Domination

To cluster strategies via different dominance values of a strategy w.r. to a given set of conditional target distributions, the individual dominance values do not matter, only their semi distance to the dominance values of other strategies, as defined in Definition 3. However, to gain a better understanding what different values for $D_T(X)$ actually mean in terms of outperforming a target, we introduce the concept of generalised alpha, which relates a dominance value to a degree of classical, CAPM-like out performance. We then decompose this alpha into its two sources: Median and skew. While different medians contribute to dominance in correspondence to classical CAPM-alpha, relative skew describes a bias of the way the strategy returns are scattered around its median relative to the way the target scatters around its median.

From (i) of the properties above, it follows for $c>0$, $D_T(T+c)>0$ and $D_T(T+c)$ is continuous and monotone increasing in c with value equal to 0 for $c=0$. Moreover, $\lim_{c\to\pm\infty} D_T(T+c)=\pm\frac{1}{2}$. Hence we define:

Definition 5

For given benchmark B and conditional target $T|B$, $\alpha_{T|B}(X)=\{\alpha|D_{T|B}(T_{|B}+\alpha)=D_{T|B}(X)\}$ is called the generalised alpha value of strategy X w.r. to target $T|B$.

$\boldsymbol{\alpha}_{\Omega(B)}(X)=\{\alpha_{T_n|B}(X), T_n\in\Omega(B)\}$ is called the generalised alpha vector of strategy X w.r. to target set $\Omega(B)$. $\boldsymbol{\alpha}_{\Omega(B)}(X)\in\mathbf{R}^{|\Omega(B)|}$.

For a set $\Omega(B)$ of targets and performance measure P,
$\alpha_{\Omega(B),P}(X)=\underset{\alpha}{\arg\min}\left\{\alpha|P(\{D_{T_n|B}(T_n|B+\alpha)\})=P(\{D_{T_n|B}(X)),n\leq|\Omega(B)|\right\}$ is called the state contingent alpha of strategy X w.r. to performance P over target set $\Omega(B)$.

Notes:

(i) From the comment above, it follows that for each given strategy X, target $T|B$ and benchmark B, the generalised alpha value exists and is uniquely defined.

(ii) The performance measure P induces a map P^* on the generalised alpha vector via

$$P(\mathcal{D}_{\Omega(B)}(X))=P(D_{T_1|B}(X),\ldots,D_{T_N|B}(X))$$

$$=P(D_{T_1|B}(T_1|B+\alpha_{T_1|B}(X)),\ldots,D_{T_N|B}(T_N|B+\alpha_{T_N|B}(X)))$$

$$=P^*(\alpha_{T_1|B}(X),\ldots,\alpha_{T_N|B}(X))=\alpha_{\Omega(B),P}(X)$$

For the following remarks, we assume that the target set $\Omega(B)$ is just the singlet $\{T|B\}$ and drop the B-contingency, as the remarks are valid for the conditional distributions as well as the marginal distributions.

Skew

For distributions with identical medians, the source of dominance is some sort of positive skew of one distribution relative to the other. The standard, statistical form of skew, defined as the third centralised and normalised moment is inadequate to measure this relative skew. We will use a different notion of skew to describe this source of dominance.

To start out we look at symmetrical distributions. A symmetric target distribution with zero mean reflects complete risk indifference. Assume that a given target distribution T admits a symmetrical density and $m_X = m_T = 0$. $D_T(X)$ is then a reflection of the extent to which X outperforms randomness and is due solely to relative skew. Specifically, we have

$$D_T(X)=\frac{1}{2}-\int_{-\infty}^{\infty} F_X(x)dF_T(x)=\frac{1}{2}-\int_{-\infty}^{0} F_X(x)dF_T(x)-\int_{-0}^{\infty} F_X(x)dF_T(x)$$
$$=\frac{1}{2}-\int_{-0}^{\infty}(F_X(x)+F_X(-x))dF_T(x)$$

as $dF_T(x)=dF_T(-x)$. If the distribution of X is also symmetrical, then $F_X(x)=1-F_X(-x)$ or $F_X(x)+F_X(-x)=1$ and hence we have:

Lemma: Assume that $m_X=m_T$. Then,

(i) If X and T have both symmetric distributions, $D_T(X)=0$ independent of their respective volatilities.
(ii) If T has a symmetric distribution, $D_T(-X)=-D_T(X)$.
(iii) If X has a symmetric distribution, $D_T(-X)=D_T(X)$.

Intuitively, whenever the distribution of X with $m_X=0$ has more weight to the right of $|x|$ than to the left of $-|x|$, i.e.

$$P(X>|x|)-P(X<-|x|)=1-F_X(x)-F_X(-x)>0, \forall x\neq 0, \qquad (**)$$

we will view X has having positive skew or positive tail skew if the this relationship holds $\forall x>x_0$, for some x_0.

Motivated by this, if $m_X=m_T=0$, we call X strongly positively skewed relative to T if

$$P(X>|x|)-P(X<-|x|)\geq P(T>|x|)-P(T<-|x|), \forall x$$
$$\Leftrightarrow 1-F_X(x)-F_X(-x)\geq 1-F_T(x)-F_T(-x), \forall x\geq 0$$
$$\Leftrightarrow F_T(x)-F_X(x)>F_X(-x)-F_T(-x), \forall x>0$$

If this relationship holds under weighting with T-probabilities, we obtain a weaker version of relative skew:

Definition 6

Let $X^*:=X\text{-}m_X$ and $T^*:=T\text{-}m_T$. The weak x_0-tail skew of X relative to T, is defined by

$$Skew_{x_0}(X|T):=\int_{x_0}^{\infty} F_{T^*}(x)-F_{X^*}(x))dF_{T^*}(x) -\int_{x_0}^{\infty} (F_{X^*}(-x)-F_{T^*}(-x))dF_{T^*}(-x)$$

If $x_0=0$, $Skew(X|T):=Skew_0\,(X|T)$ will be called the weak skew of X relative to T.

Using a change of variable, the following is immediate:

Proposition: If $m_X = m_T$, then $Skew\,(X|T)=D_T(X)$.

The skew of a target distribution in the sense of (**) is an indication of the risk attitude of the expressed preference. Skew in the sense of (**) and relative skew are different from statistical skew.

Median and Skew

As seen from the example above, relative symmetry does not contribute in either way to dominance; the only source of dominance in such case is a difference in medians. To distinguish their respective contributions to dominance, we will decompose the generalised alpha value into a pure median portion and a skew portion:

Let $X^*=X\text{-}m_X$ and $X^*=T-m_T$, we set

(a) $\alpha^{S}_{\{T\}}(X)=D_{T^*}(X^*)$
(b) $\alpha^{m}_{\{T\}}(X)=D_T(X)-D_{T^*}(X^*)$

We then have $\alpha_{\{T\}}(X)=\alpha^{m}_{\{T\}}(X)+\alpha^{S}_{\{T\}}(X)$, i.e. the generalized alpha can be decomposed into an alpha portion solely due to distributional dominance caused by the difference in medians and an alpha portion due solely due to the relative skew.

Correspondence to Classical CAPM Alpha

$\alpha^{m}_{\{T\}}(X)$ is the distributional correspondence of standard (CAPM) alpha. To illustrate this correspondence, let $T|_{B=b}(x)=\delta_b(x)$, $s.th.$ B acts as a direct benchmark. Assume $X=\beta B+\alpha$.

Then for the Sharpe Ratio of X, $SR(X)$, we have: $SR(X) = \dfrac{EX}{\sigma_X} = SR(B)+\dfrac{\alpha}{\beta\sigma_B}$.

In other words, in the classical CAPM setting, X dominates B if $\alpha>0$. In our approach, if the CAPM-$\alpha > 0$, then $\alpha^{m}_{\{T\}}(X)> 0$, so for Skew($X|T$)=0, the dominance is equivalent and hence $\alpha^{m}_{\{T\}}(X)$ corresponds to α. If, however, the skew is nonzero, $D_T(X)$ may differ from the SR-induced dominance. The same holds true if T acts as a target for outperformance with the information ratio replacing the Sharpe ratio.

In general, when compared to classical performance measures of the form return/risk, $\alpha^{m}_{\{T\}}(X)$ corresponds to the "return portion", while the "risk portion" is reflected by $\alpha^{S}_{\{T\}}(X)$. However, $\alpha^{S}_{\{T\}}(X)$ can reflect much more detail w.r. to risk attitude than any typical risk measure.

Lemma: Given two strategies X, Y and a target T. Assume T is symmetric and $m_T=0$, $m_X=0$ and $Skew(X|T)>0$, Y is symmetric and $m_Y>0$. Let $\alpha=\alpha_{\{T\}}(X)=\alpha^{S}_{\{T\}}(X)$ denote the generalised alpha of X w.r. to target T. Then,

(a) $D_{T+\alpha}(X)=0$,
(b) $D_T(Y)>D_T(X)$, if $m_Y>\alpha$ with equality for $m_Y=\alpha$.

(a) follows from $D_T(X)–D_{T+\alpha}(X)=D_T(T+\alpha)$, (b) follows from the definitions.

The Lemma relates the dominance of pure outperformance in terms if standard, classical alpha to the dominance derived solely by relative skew.

4. Applications

The emphasis in the applications is on illustrating the methodology. Hence, we will concentrate on general strategies or Hedge fund indices and rather simple target distributions. The real value of the method however, lies in evaluating single strategies with more detailed and bespoke target distributions, which is beyond the scope of this paper.

4.1 Clustering with a Market State Indicator as Benchmark

In the first application, the universe of strategies we consider is a set of Hedge Fund styles, each represented by their corresponding HFR index. Specifically, we consider the indices Global Hedge Fund, Equity, Equity Market Neutral, Macro, Distressed, Merger Arbitrage and Convertible Arbitrage. We use the daily return series from January 1, 2006 to December 1, 2010.

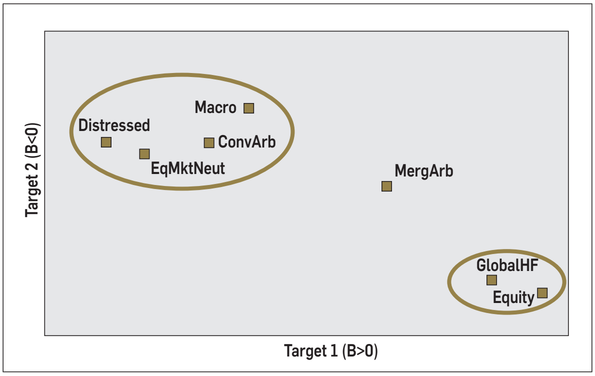
Given the strategies' benchmark-free, absolute return status, we use a market state indicator as the benchmark. The indicator we use is based on the S&P 500 index and distinguishes seven states: {very negative, negative, slightly negative, neutral, slightly positive, positive and very positive}, defined via quantiles of the 5-day return series of the S&P 500. For example, for the week starting November 16, 2008 in which the S&P 500 lost 8.39%, the state indicator would be "very negative", while for the week starting September 6, 2010 with the S&P 500 return of 0,46%, the state would be "slightly positive".

The targets will be defined via their conditional distributions, i.e. for each state by a separate distribution. The conditional distribution of the strategies is taken by a kernel approximation of their empirical data.

The samples we use will be the 3-day, 5-day and 15-day returns of the indices.

As a first quick illustration, we choose for the first two targets the following conditional distributions: $T_1|_{B=b}(x)=\delta_0(x)1_{(b>0)}$ and $T_2|_{B=b}(x)=\delta_0(x)1_{(b>0)}$, which measure the degree to which positive returns overweigh negative returns in positive respectively negative market environments. This produces the following simple cluster space for 5-day returns:

Figure 2. **Dominance Vectors and Induced Clustering by T_1 and T_2**



Besides the clustering, what can be seen is that Merger Arbitrage is closest to a parallel of the diagonal and hence most independent of the state of the market. The Figure also shows that the strategies are closer to each other for the second target, *s.th.* for negative market states, Merger Arbitrage would join the Distressed/EqMktNeutral/ConvArb/Macro cluster. This also highlights the gain in information and detail when moving to higher dimensions in dominance/cluster space by adding further targets.

For further detail, while still keeping the analysis simple, we add the following groups of target distributions:

$$T_3|_{B=b}(x)\sim N(0,\sigma)1_{(b=\{k\})},\ k=1,.2,...7$$
$$T_4|_{B=b}(x)\sim N(E(X|B=b),\sigma)1_{(b=\{k\})},\ k=1,.2,...7$$
$$T_5|_{B=b}(x)\sim 0.5\left[\{\text{“negative tail event”}\}+\{\text{“positive mean event”}\}\right]1_{(b=\{k\})},\ k=1,.2,...7$$
$$T_6|_{B=b}(x)\sim 0.5\left[\{\text{“negative mean event”}\}+\{\text{“positive tail event”}\}\right]1_{(b=\{k\})},\ k=1,.2,...7$$

where $N(0,\sigma)$ denotes the normal distribution with mean 0 and standard deviation σ, $E(X|B=b)$ denotes the conditional expectation of strategy X, given B is in state b

and $b=\{k\}$ denotes that B is in state k, for k one of the seven states described above. The tail events are the 10% and 90% quantiles respectively.

The target groups $T_3|_{B=b}(x)$ and $T_4|_{B=b}(x)$ measure degrees of bias/skew and to what degree strategies dominate randomness in a state contingent way. We will use these target groups for the 3-day, 5-day and 15-day strategy returns, incorporating not only a sense of path dependence, but also the expectation, that strategies of absolute return character "build" their performance over time and a positive, market state independent bias has to evolve – in other words, we want to see increasing dominance with increasing return periods. The target groups $T_5|_{B=b}(x)$ and $T_6|_{B=b}(x)$ measure state contingent tail skews in both directions. The dominance space is thus N=58 dimensional. According to Definition 3 and the note following it, we cluster the strategies by identifying a cluster preference matrix Σ.

Figures 3 to 5 represent the resulting clusters by projecting the weighted dominance vectors $\mathcal{D}_{\Omega(B)}(X)=\Sigma^{-1/2}\mathcal{D}_{\Omega(B)}(X)$ onto the two-dimensional subspace $(B<0; B\geq 0)$. We have considered three different cluster preferences: Pref1: Indifference as to targets, Pref2: Overweigh risk aversion, Pref3: Overweigh risk tolerance.

Figure 3. Pref-1, Indifference to Targets

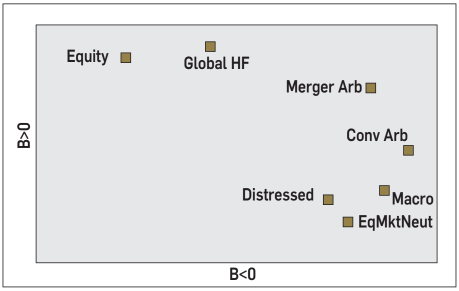


Figure 4. Pref-2, Risk Aversion

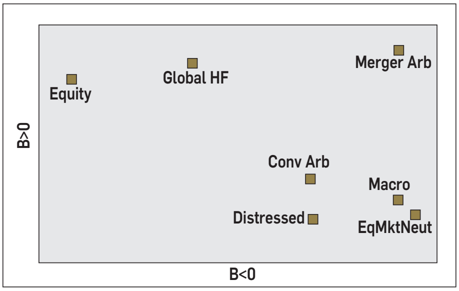
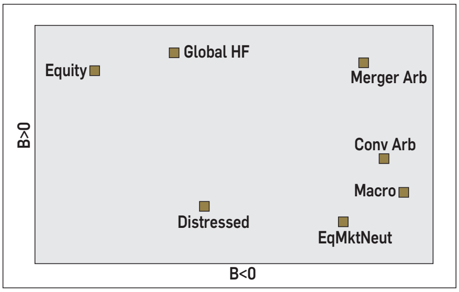


Figure 5. Pref-3, Risk Tolerance



Selected conclusions:

- The Global Hedge Fund and HFR Equity Indices confirm their relative high correlation to the S&P 500 by their position in Figures 3 to 5 (top left position, dominance and market state co-monotonic).
- The Macro-, Distressed- and Equity Market Neutral Indices cluster close together in the simplified risk neutral projection of Figure 2.
- Increasing risk tolerance benefits mostly the Macro Index, while the distressed index drops out of the previous cluster to the negative.
- Increasing risk aversion lets the Convertible Arbitrage Index drop to the negative relative to the rest.
- The evolution of dominance when increasing the return period is best for the Merger Arbitrage Index and worst for the Distressed and Convertible Arbitrage indices.
- Overall, in these simplified projections, a measure of quality is closeness to the top right corner. In this respect, the Merger Arbitrage Index comes out best for the chosen setting of preference targets and projections.

In projecting the 58-dimensional dominance space to just two dimensions as the only consideration, a lot of detail will be lost. But as the scope of this paper is to illustrate the methodology, the illustration is kept short and simple. These conclusions reflect the specific choices of targets and the specific aggregation reflected by the chosen projections as given by Figures 3 to 5. The interest in the diagrams provided is in the relative position of strategies to each other and their changes under changes in preference. Further targets and other aggregations will reveal further details of the individual performance characteristics.

4.2 Performance Ordering with a Market Index as Performance Benchmark

In this application, we consider a number of equity strategies, which we will evaluate against an equity index as the performance benchmark. Specifically, we consider a CPPI strategy on the € Stoxx 50 - floor at 90%, cap at 150%, multiplier of 5, annual reset ("CPPI"), a passive strategy consisting of the index plus a short Collar (long Put, short Call) on the index with option expirations of 3 months and rule based rolls ("Index plus Collar"), the HFR Equity index ("HFR Equity") and a rule based dynamic short-put option strategy on the index ("Put Options"). The samples used are the daily returns of the strategies, for the period from January 1, 2006, to December 1, 2010. As in the previous example, the targets are described by their conditional distributions. The conditional distributions of the strategies cannot be reliably estimated by kernel methods from empirical data, as in this case, not enough data points are available. Instead, we estimate the bi-variate copula between the strategies and the benchmark and rely on the empirical distributions for estimating the marginals.

As can be seen from their respective empirical marginal distribution functions in terms of broad distributional properties, all strategies are in essence "β–times" the index for various levels of β, i.e. they are examples of pure Alternative Index Betas. In this application we will derive various performance evaluations by mapping their dominance vectors to the real numbers under different preferences. We consider the following conditional target distributions:

$$T_1|_{B=b}(x)=\delta_b(x)$$
$$T_2|_{B=b}(x)\sim N(b,\sigma)(x)$$
$$T_3|_{B=b}(x)\sim N(b,\sigma)(x)(1_{(b\leq q_3)}(b)+1_{(b\leq q_4)}(b))$$

where σ is the standard deviation of the strategy distribution. T_1 measure the degree to which the median of the conditional strategy distribution exceeds b, T_2 measures the conditional out performance of random scattering around b. The difference of the dominance over T_2 and T_1 indicates what portion of the T_2 dominance is driven by the median and what is driven by skew. q_3 and q_4 are s.th. $F_B(q_3)=1-F_B(q_4)=0.1$ i.e. T_3 measures the dominance over random scattering around b, but only in cases of B-tail events.

Restricting the conditional targets to negative respectively positive b, yields positive respectively negative dominance w.r. to targets T_1 and T_2 for all strategies, reflecting a β of less than one for all strategies.

To derive four preference based performance measures, we distinguish four different preferences reflecting different attitudes towards risk, resulting in four different weightings in mapping the dominance vector to the real numbers:

Pref-1: Indifference w.r. to targets, equal weighting. Pref-2: Strong preference of tail dominance, i.e. heavy overweight on T_4. Pref-3: Risk averse, overweigh dominance w.r. to negative b and overweigh dominance w.r. to left tails. Pref-4: Risk tolerant, overweigh complementary to Pref-3. This results in the following ordering of the strategies:

Table 1. Orderings produced by common Performance Measures and by Preference weighted Dominance Ordering

RANK	TOTAL PERFORMANCE	SHARPE RATIO	INFORMATION RATIO	PREF-1	PREF-2	PREF-3	PREF-4
1	Index plus Collar	Index plus Collar	Index plus Collar	HFR Equity	Index plus Collar	Index plus Collar	HFR Equity
2	HFR Equity	HFR Equity	CPPI	Put Options	HFR Equity	HFR Equity	Put Options
3	CPPI	CPPI	HFR Equity	Index plus Collar	CPPI	CPPI	CPPI
4	Put Options	Put Options	Put Options	CPPI	Put Options	Put Options	Index plus Collar

Not surprisingly, different weightings for the preferences will eventually lead to different orderings specifically when strategies like in this case are, from a distributional view, unlevered versions of the benchmark. The increased weighting of risk tolerance at the cost of risk aversion profits the Hedge Fund index and hurts the Collar strategy, as the short calls cap upside participation. Given the choice of targets, Pref-1 with its equal weighting of preferences emphasizes the degree to which strategies dominate or outperform random scattering around the benchmark, i.e. the degree of systematic outperformance. With this view, the HFR Equity index fares relatively better than under performance evaluation by Sharpe – or Information ratios. Under Pref-1, with its target – and risk indifference and Pref-4, with the overweight on risk tolerance, the dynamic option strategy ranks clearly higher than in all other orderings, suggesting it to be much more a "pro risk" strategy, in contrast to what the strategy looks like in calm markets, where it behaves more like a low volatility alpha generator. Note that the period under consideration includes September/October 2008, where the asymmetric return profile of un-hedged, short put option strategies becomes apparent. In other market phases, the comment in the introduction about distorting the Sharpe ratio by this type of strategy applies.

As before, while those illustrative conclusions and orderings can also be obtained by simpler performance measures, the formulation of preferences and wealth of detail attainable is substantially larger in this approach.

5. Comparison to common Performance Measures and Clustering Techniques

A measure of the performance quality of a trading strategy is not an absolute truth, but is relative to individual investor's perception. While common performance measures are typically preference free and strictly risk averse, the stochastic ordering methods as introduced in this study allow for a graded and state contingent attitude towards risk. Moreover, it is an approach without any restrictions as to which preference and state contingencies are included. It is this feature which distinguishes this approach from common methods and makes it a useful tool for practical applications in asset allocation, portfolio construction and analysis. The result can be seen in Table 1 of the last application, where the produced orderings under varying preferences are compared to three common performance measures (total return, Sharpe ratio and information ratio). Specifically, the preference set Pref-4, reflecting a risk tolerant attitude produces a completely different ordering over the same time period, than the common measures. Dominance weighted performance ordering yields different results than previous approaches due to its sensitivity to bespoke preferences.

With respect to clustering, the most common approach to cluster trading strategies is by allocating funds to benchmarks or style classes, with no specific regard of realized performance characteristics. To the author's knowledge, there have been no previous approaches to the state contingent, preference based clustering of trading strategies.

6. Conclusions

The paper illustrates a methodology to cluster and evaluate trading strategies, addressing a number of inefficiencies of previous and common methods. It uses information obtained by comparing the full return distribution of a trading strategy to distributions of state contingent targets, reflecting specific preferences, using a version of weak integral stochastic dominance. It thus separates trading strategies by distinguishing their state contingent and preference based performance characteristics.

Specifically, this allows including a diverse range of preferences w.r. to tail behaviour and different attitudes towards risk, particularly a state contingent preference of risk aversion and tolerance. The methodology does not rely on any distributional assumptions and can detect non normality and any sort of skew or moment related distributional characteristic. It also allows including and reflecting path dependencies

In addition, expressing preferences via target distributions by stating how returns should scatter around a benchmark or ought to behave in different states of the market environment is very intuitive and easy to formulate.

While the methodology is very flexible to include all kinds of contingencies and preferences and is thus able to distinguish and differentiate performance characteristics to very fine detail, it does potentially have a large number of degrees of freedom. As such, while it is an ideal tool to gain insight into the finer details of trading strategies, it is not a quick way to obtain a rough overview of performance quality.

Another application only briefly highlighted is the replication of trading strategies.

References

- Annaert, J., van Osselaer, S. and Verstraete, B. (2009). Performance evaluation of portfolio insurance strategies using stochastic dominance criteria, *Journal of Banking and Finance*, **33**, pp. 272-280.
- Cherny, A., Douady, R. and Molchanov, S. (2009). On Measuring Hedge Fund Risk, Preprint, available at http://papers.ssrn.com

■ Cherny, A. and Madan, D. (2009). New Measures for Performance Evaluation, *Review of Financial Studies*, **22**(7), pp. 2571-2606.

■ Cogneau, P. and Hubner, G. (2009). The 101 Ways to Measure Portfolio Performance, Preprint, available at http://papers.ssrn.com

■ Eberlein, E. and Madan, D. (2009). Hedge Fund performance: Sources and Measures, *International Journal of Theoretical and Applied Finance*, **12**(3), pp. 267-282.

■ Goetzmann, W. N., Ingersoll, J. E., Spiegel, M. I. and Welch, I. (2002). Sharpening Sharpe ratios, Yale School of Management Working Papers, WP[02-08], Yale.

■ Kat, H. and Palaro, H. (2005). Who needs hedge funds? A copula-based approach to hedge fund returnreplication, Alternative Investment Research Centre Working Paper Series, WP[027], Cass Business School, City University, London.

■ Kazemi, H., Mahdavi, M., and Schneeweis, T. (2003). Generalized Sharpe Ratio: A Defense Against Sharpened Sharpe Ratios, Working Paper of Isenberg Management School, University of Massachusetts and Smith College.

■ Levy, H. (2006). *Stochastic Dominance*. Springer Verlag, 2nd Ed, Berlin.

■ Li, S. and Linton, O. (2007). Evaluating hedge fund performance: a stochastic dominance approach, Discussion Paper [591], LSE, London.

■ Mainik, G. and Rüschendorf, L. (2010). Ordering of multivariate probability distributions with respect to portfolio losses. Preprint.

■ Meucci, A. (2007). *Risk and Asset Allocation,* Springer Verlag, Berlin.

■ Müller, A. (1997). Stochastic orders generated by integrals: A unified study, *Advances in Applied Probability*, **29**, pp. 414-428.

■ Müller, A. and Stoyan, D. (2002). *Comparison Methods for Stochastic Models and Risk*, Wiley, New York.

■ Rüschendorf, L. (2005). Stochastic Ordering of risks, influence of dependence and a.s. construction, in Balakrishnan, N., Bairanov, I.G., Gebizlioghu, O.L. (Eds.) *Advances on Models, Characterizations and Applications,* Chapman & Hall/CRC Press, New York, pp. 19-56.

■ Spurgin, R.B. (2001). How to game your Sharpe ratio, *Journal of Alternative Investments,* **4**(3), pp. 38-46.

■ Zakamouline, V. and Koekebakker, S. (2009). Portfolio Performance Evaluation with Generalized Sharpe Ratios: Beyond the Mean and Variance. *Journal of Banking and Finance,* **3**, pp. 1242-1254.