Conditionally fitted Sharpe performance with an application to hedge fund rating

Serge Darolles\textsuperscript{a,}\textsuperscript{*}, Christian Gourieroux\textsuperscript{b}

\textsuperscript{a} Lyxor Asset Management and CREST, Paris, France
\textsuperscript{b} University of Toronto, CEPREMAP and CREST, Paris, France

\textbf{ARTICLE INFO}

\textbf{Article history:}
Received 13 March 2009
Accepted 26 August 2009
Available online 2 September 2009

\textbf{JEL classification:}
G11

\textbf{Keywords:}
Hedge fund
Sharpe ratio
Fitted performance
Fund rating
Segmentation

\textbf{ABSTRACT}

We define a battery of Sharpe performance measures, which differ by the information taken into account in their computation, but also by the potential use of the fund by the investor. Four advantages of Sharpe performance based rating are especially important for the investor. First, the performance measures correspond to the standard measures used for mutual funds and known by retail investors. Second, we can compare the numerical results, even if they are obtained with different assumptions. Third, the rankings are based on regression analysis and easy to compute. Fourth, we can easily use these performance measures in the design of an optimal basket of hedge funds. Finally, we can use the performance measures to partition the set of funds into homogenous segments.

\textcopyright 2009 Elsevier B.V. All rights reserved.

1. Introduction

The hedge fund industry has grown quickly during the last 10 years and represents now about $1600 billions of assets and more than 9000 funds. However, this market is still in its infancy, which explains the relative lack of information on the definition of strategies, portfolio allocations, risk and performance measures. As in the mutual fund industry,\textsuperscript{1} the current trend is to diffuse information on hedge funds by means of ratings regularly elaborated by specialized agencies and published in general or specialized newspapers. Two kinds of ratings have been historically proposed. The first ones, introduced by Standard and Poor’s in the mid-1970’s, are qualitative and based on advanced funds due diligence and manager interviews. Due to the cost of such analysis, the results were not free of charge, and only a few funds were analysed. This explains the success of quantitative ratings based on historical return data, first proposed by Morningstar in the mid-1980’s, then followed by Standard and Poor’s, Lipper, Axiomix and more recently by Edhec-Européperformance (see, e.g. Amenc and Le Sourd (2005) for a detailed study of these ratings). These ratings can easily be proposed free of charge; they already cover all mutual funds and can be successfully adapted to hedge funds, at least in a preliminary analysis. Indeed, most hedge funds use highly dynamic investment strategies, can have short sell and, in this respect, are different from mutual funds. But, even if we have full transparency on the holdings at some specific dates, we do not get the whole information on the dynamic strategy followed by the fund manager. Historical return data\textsuperscript{2} are generally the only source of information for performance estimation and risk analysis. Indeed, retail investors do not have access to managed accounts facilities (i.e. full transparency on the fund holdings) and then must only use historical returns in their funds selection process. This explains why standard quantitative ratings are also used for hedge funds, at least in a first step. These quantitative ratings are most of the time poorly explained, and often misunderstood not only by retail investors, but also by more specialized ones including pension funds, corporates and institutional investors. The aim of this paper is to review in detail the ratings based on the Sharpe performance measures and to standardize their use. We insist on the fact that all performance measures must be investor driven, and not model driven.

In Section 2, we recall the derivation of a conditional Sharpe performance measure in the mean–variance management framework. This derivation shows that different Sharpe performances can be constructed for a given fund. They depend on the information of the investor, but also on the type of other investments, that

\textsuperscript{*} Corresponding author. Tel.: +33 1 41 17 35 90; fax: +33 1 41 17 76 66.
E-mail addresses: serge.darolles@ensae.fr (S. Darolles), christian.gourieroux@ensae.fr (C. Gourieroux).

\textsuperscript{1} The gap between mutual funds and hedge funds is diminishing rapidly with the so-called “hedged mutual funds”, that are mutual funds mimicking hedge fund strategies, but still regulated by the Securities and Exchange Commission (SEC) (see Agarwal et al. (2009)).

\textsuperscript{2} And the design of the allocation between investor’s account, provision account and management account (see Darolles and Gourieroux (2009)).
he/she want to complete by buying/selling the fund and of the investment horizon. As a consequence, there exist as many rankings as ways of defining these performances, and these rankings do not necessarily provide identical results. We also reinterpret in terms of Sharpe performance the use of the so-called alphas in the ranking process. The persistence question is included by considering the term structure of performances. All these rankings are easy to obtain from simple regression models, and the numerical results can be compared since they come from the same theoretical framework. An application to a set of hedge funds is given in Section 3. We determine and compare the rankings associated with different performance measures such as the historical performance, or a performance when the fund is introduced to complete either the market portfolio, or a portfolio of equities containing the medium or large size firms. We also discuss the pattern of the term structure of performances. Section 4 explains how the performance measures can be used to construct segmentations of hedge funds. When the segmentation is based on a single performance measure, each segment includes the funds with performance in a given medium or large size firms. We also discuss the pattern of the term structure of performances. Section 4 explains how the performance measures can be used to construct segmentations of hedge funds. When the segmentation is based on a single performance measure, each segment includes the funds with performance in a given interval. We discuss the selection of thresholds providing the best contrast between segments while ensuring enough homogeneity within segments. The method is extended to get homogenous clusters based on several Sharpe performance measures. Section 5 discusses some limitations of the Sharpe performance based rankings and proposes directions for future improvements. Some proofs are gathered in appendices.

2. Sharpe performance

Let us first recall how the Sharpe performance is related to the mean–variance portfolio management, and explain how it can be used to define individual performances. Then, we study the increase in Sharpe performance, when the set of assets is extended.

2.1. Mean–variance efficient portfolio

Let us denote by \(a_0, a_1, \ldots, a_n\) the amounts invested at date \(t\) in the different assets \(i = 0, 1, \ldots, n\), where \(0\) denotes the riskfree asset. The value of the portfolio at date \(t\) is \(w_t = a_0 + \cdots + a_n\), whereas its future value is uncertain, equal to \(w_{t+1} = a_0(1 + r_t) + \sum_{i=1}^{n} a_i p_{t+1,i}/p_t\), where \(r_t\) denotes the riskfree rate at horizon 1 and \(p_t\) the unitary price at \(t\) of asset \(i\). In the mean–variance framework, the investor selects the portfolio allocation by solving the following optimization problem:

\[
\max_{a_0, \ldots, a_n} \mathbb{E}(w_{t+1}) - \frac{A}{2} \mathbb{V}(w_{t+1}), \quad \text{s.t.} \quad \sum_{i=0}^{n} a_{i} = w_{t},
\]

where \(A\) is the individual (absolute) risk aversion, and \(\mathbb{E}, \mathbb{V}\) denote the expectation and variance, respectively, given the information used by the retail investor at date \(t\). We can solve the budget constraint to get the quantity invested in the riskfree asset: \(a_0 = w_t - \sum_{i=1}^{n} a_i\). By substitution, we deduce the unconstrained quadratic optimization problem:

\[
\max_{a_1, \ldots, a_n} w_t(1 + r_t) + \mathbb{E}\left[\sum_{i=1}^{n} a_i y_{t+1,i}\right] - \frac{A}{2} \mathbb{V}\left[\sum_{i=1}^{n} a_i y_{t+1,i}\right].
\]

where \(y_{t+1,i} = (p_{t+1,i} - p_t)/p_t - r_t\). The optimal allocation in the risky assets, \(a_i^* = \left(a_{i,1}^*, \ldots, a_{i,T}^*\right)^T\), is given by (see, e.g. Markowitz (1952)):

\[
a_i^* = \frac{1}{A} \Sigma_i^{-1} m_i,
\]

where \(m_i\) [resp. \(\Sigma_i\)] is the expectation [resp. the variance–covariance matrix] of the vector of excess returns \(y_{t,i} = (y_{1t,i}, \ldots, y_{Tt,i})^T\). Moreover, the optimal value of the objective function is equal to:

\[
\Pi_t = w_t(1 + r_t) + \frac{1}{2A} \Sigma_i^{-1} m_i.
\]

It depends on the initial budget, riskfree rate, risk aversion coefficient and quantity:

\[
S_{t,1 \ldots n} = m_i^* \Sigma_i^{-1} m_i.
\]

which summarizes the stochastic properties of the risky returns (see Treynor (1965), Sharpe (1966)). This quantity is called the Sharpe performance of the set of assets \(1 \ldots n\) at date \(t\). The Sharpe performance depends on the information used by the investor to compute the means, variances and covariances. It is called historical Sharpe performance when the computation corresponds to the unconditional mean and variance, and conditional (or potential) Sharpe performance (see Jobson and Korkie (1982)), otherwise. The historical and conditional Sharpe performances have different interpretations due to the different information sets, even if their computations are similar. Moreover, the historical performance cannot be deduced from the knowledge of the conditional performance in all environments (see Appendix A1). As noted early by Sharpe (1994), the use of unadjusted historical Sharpe ratios as surrogates for unbiased estimations of conditional ratios “remained subject to serious question”.

2.2. Individual sharpe performance

As an illustration, let us consider a portfolio including the riskfree asset and a single risky asset \(j\). The associated performance is:

\[
S_{t,j} = m_j^2 / \sigma_j^2.
\]

The performance can take any positive value. Indeed, even if \(m_j\) is strictly negative, the risky asset is profitable by introducing a negative amount \(a_j\), whenever shortsell is allowed. If the investor has to choose between a portfolio including the riskfree asset and risky asset \(j\), and a portfolio including the riskfree asset and risky asset \(k\), he/she will prefer the first one if \(S_{t,j} > S_{t,k}\), whenever he/she allocates the assets in a mean–variance efficient way. This explains why competing funds (i.e. asset portfolios) are usually compared and ranked by means of these Sharpe performances \(S_t\). The practitioners generally define the performance as \(S_{t,j}^2 = m_j^2 / \sigma_j^2\). It corresponds to the so-called Sharpe ratio when expectation and volatility are annualized. Of course, \(S_t\) and \(S_{t,j}^2\) provide the same ranking. The Sharpe ratio can be interpreted as a “risk premium”. Loosely speaking, it represents the change (in %) of expected return following a 1% change of the risk \(\sigma\) and was originally called reward–to–variability ratio. It can also be computed taking into account the sign of the expected excess return as \(SR_{t,j} = m_j / \sigma_j\) (see the discussion in McLeod and Van Vuuren (2004)).

In practice, the individual performances are not known and have to be estimated from a sequence of asset excess returns \(y_{jt}, t = 1, \ldots, T\), say. The estimates depend on the type of individual performance which is considered.

2.2.1. Historical individual performance

When the mean and variance are unconditional, we approximate the Sharpe performance by its sample counterpart:

\[
\hat{S}_j = \frac{\sqrt{\frac{1}{T} \sum_{t=1}^{T} y_{jt}^2}}{\sqrt{\frac{1}{T} \sum_{t=1}^{T} (y_{jt} - \frac{1}{T} \sum_{t=1}^{T} y_{jt})^2}}.
\]

This estimator is called ex-post performance to highlight its computation from past observations.
2.2.2. Conditional individual performance

In the conditional framework, we need the list of observable variables Z_{t-1}, say, included in the information set, and we have to explain how they are used by the investor to derive the conditional mean and variance. The explanatory variables are indexed by \( t-1 \) instead of \( t \) to highlight the fact that they have to be known in advance for prediction purpose. For instance, they can include lagged returns or macroeconomic variables. Let us assume that the conditioning is performed by means of a linear regression. The computation of the estimated performance is done in two steps. We first regress the returns on the explanatory variables:

\[
y_{j,t} = \delta_j + Z_{t-1} \gamma_j + u_{j,t},
\]

which provides estimated coefficients \( \hat{\delta}_j(Z) \), \( \gamma_j(Z) \) and also an estimate of the residual variance:

\[
\hat{\sigma}_j^2(Z) = \frac{1}{T} \sum_{t=1}^{T} \hat{u}_{j,t}^2(Z) = \frac{1}{T} \sum_{t=1}^{T} (y_{j,t} - \hat{\delta}_j(Z) - Z_{t-1} \hat{\gamma}_j(Z))^2.
\]

In a second step, the estimated conditional individual performance at date \( T \) is:

\[
\hat{S}_{1j}(Z) = \frac{\hat{\delta}_j(Z) + Z_{T-1} \hat{\gamma}_j(Z)}{\hat{\sigma}_j^2(Z)},
\]

which depends on the selected regressors (asset returns or economic variables) and on date \( T \).

2.3. Fitted sharpe performance

Let us now consider the riskfree asset and two risky assets 1 and \( j \), say.\(^4\) The investor can construct an efficient portfolio with the riskfree asset and risky asset 1; he/she can also construct an efficient portfolio including the three assets. In the first case, the Sharpe performance is:

\[
\hat{S}_{1} = \frac{m_{1} - r_f}{\hat{\sigma}_{1}},
\]

and, in the second case, the Sharpe performance becomes:

\[
\hat{S}_{1j} = (m_{1}, m_{j}) \left( \frac{\sigma_{1j}^{2}}{\sigma_{j}^{2}} \right)^{-1} \left( \frac{m_{1}}{m_{j}} \right).
\]

The increase in Sharpe performance due to the inclusion of the new asset \( j \) is:

\[
\hat{S}_{1j} = \hat{S}_{1} - \hat{S}_{1}.\]

The additive decomposition formula\(^5\) for the joint performance \( \hat{S}_{1j} \) accounts for the potential correlation between the basic asset 1 and the fund \( j \). This increase is a natural criterion to rank different funds, this is the second justification of the terminology “fitted”.

Another important application of this concept can be found in the design of an optimal basket of funds. In this case, the risky asset 1 can be a candidate basket of funds. We can quantify the interest of completing this basket with fund \( j \) by a direct computation of the fitted performance measure \( \hat{S}_{1j} \). Since the conditional performance varies in time, the selection of the completing fund has to be regularly updated.

The fitted performance has another expression which is useful for interpretation and estimation purpose. Indeed, we can check that [see Appendix A.2]:

\[
\hat{S}_{1j} = \frac{ \left( m_{1} - \frac{\sigma_{1j}^{2}}{\sigma_{j}^{2}} m_{j} \right)^{2}}{\sigma_{j}^{2} - \frac{\sigma_{1j}^{2}}{\sigma_{j}^{2}}},
\]

To understand the importance of this expression and its interpretation in terms of regression residuals, we first consider the case of historical fitted performance.

2.3.1. Historical fitted performance

When the means, variances and covariances are computed unconditionally, we have:

\[
S_{1j} = \frac{ \left( m_{1} - \frac{\sigma_{1j}^{2}}{\sigma_{j}^{2}} m_{j} \right)^{2}}{\sigma_{j}^{2} - \frac{\sigma_{1j}^{2}}{\sigma_{j}^{2}}},
\]

which is estimated by its sample counterpart. Let us now interpret the numerator and denominator of the expression above. For this purpose, let us introduce the linear regression of \( y_{j,t} \) on \( y_{1,t} \):

\[
y_{j,t} = \beta_{1j} y_{1,t} + \beta_{j} y_{j,t} (1), \quad \text{say.}
\]

The OLS estimators of the coefficients are \( \hat{\beta}_{1j} \), \( \hat{\beta}_{j} \) and the estimator of the variance of the error term is \( \hat{\sigma}_{j}^{2}(1) \). The comparison with (15) shows that:

\[
\hat{S}_{1j} = \frac{ \hat{\sigma}_{j}^{2}(1) }{ \hat{\sigma}_{j}^{2}(1)}.\]

This is the so-called Treynor appraisal ratio, when asset 1 is the market portfolio. The regression (16) has the following financial interpretation: if the returns of assets 1 and \( j \) were (conditionally) uncorrelated, the joint Sharpe performance \( \hat{S}_{1j} \) would be equal to the sum of individual performances \( \hat{S}_{1} \) and \( \hat{S}_{j} \). When the asset returns are correlated, the aim of the regression is to construct the portfolio in assets 1 and \( j \), which is (conditionally) uncorrelated with the benchmark portfolio and allows for additional diversification. The allocation of this portfolio is 1 in asset \( j \), \(-\beta_{j}(1) \) in asset 1, and its return is \( \hat{\beta}_{1j}(1) + \beta_{j}(1) \). The need for a preliminary regression is the second justification for the terminology “fitted”.

2.3.2. Conditional fitted performance

If we still consider a linear regression framework and some variables \( Z_{t-1} \) in the information set, the regression becomes:

\[
y_{j,t} = \delta_j + Z_{t-1} \gamma_j + y_{1,t} \beta_{1j} + w_{j,t},
\]

with estimators \( \hat{\delta}_j(1,Z), \hat{\gamma}_j(1,Z), \hat{\beta}_{1j}(1,Z), \hat{\gamma}_j(1,Z), \) say. The estimated conditionally fitted individual performance is:

\[
\hat{S}_{1j}(Z) = \frac{\hat{\delta}_j(1,Z) + Z_{T-1} \hat{\gamma}_j(1,Z)}{\hat{\sigma}_j^2(Z)},
\]

We get as many rankings as choices of explanatory variables and benchmark portfolios.

To summarize, the unconditional/conditional, fitted/unfitted performance measures are simply Sharpe performances computed on regression residuals. They differ by the set of lagged variables and current returns which are introduced. The lagged regressors are introduced to represent the information, whereas the current
returns are included for taking into account the inclusion of the fund in the investor’s portfolio.

Finally, the use of Sharpe ratios has been criticized in the literature, since the hedge fund returns follow a non-symmetric (unconditional) distribution. This criticism applies to the historical (i.e. unconditional) performance, but not to either the fitted, or conditional measures. Indeed, the pattern of the distribution of the error term matters, not the pattern of the historical distribution of returns.

2.4. Errors to be avoided

The mean–variance approach provides a coherent framework for comparing the interests in different investments. Indeed, the various Sharpe performance measures have similar interpretations, and we can compare their values to see the effect of the information, if the lagged variables are different, or to detect the best potential use of the fund, if the current returns are different.

As noted above, these approaches do not provide a unique ranking, but different ones according to the type of portfolio that the investor wants to consider and to the information used to compute the mean and the variance. Other ranking criteria have been introduced in the mean–variance framework by academics or practitioners (see Section 5), but do not feature the coherency of the Sharpe performance approach.

2.4.1. Currency risk

The rankings are derived from returns, generally computed in US dollars and appropriate for investor with US dollars as the money unit. They are not appropriate for investors with Euro as the money unit. If returns are not hedged for currency risk, different rankings depending on the money unit, i.e. US dollars, Euro, Yen, should be published to avoid to the retail investor the correction for currency risk by him/herself, which is not standard. On the contrary, if returns are perfectly hedged\(^8\) for currency risk, we do not have to consider this point. In this case, the monetary unit used for the computation has to been mentioned explicitly.

2.4.2. Ranking based on the deviation to Security Market Line (SML)

This methodology has been initially suggested by Jensen (1968) as a measure of abnormal performance, and applied after various corrections in a lot of papers on hedge fund performance (see, e.g. Ferson and Schadt (1996), Kat and Miffre (2002), Amenc and Le Sourd (2005)). This is the core of the European performance fund rating methodology. By choosing a benchmark portfolio depending on the announced portfolio management category, the authors introduce a selectivity bias in the ranking comparison of hedge funds.

\[ y_{jt} = \beta_0 + \beta_1 y_{mt} + \beta_2 y_{at} + \beta_3 y_{zt} + \epsilon_{jt} \]

Thus, this student comes likely from a misleading interpretation of the notion of abnormal return (see Jha et al. (2009)). Indeed the intercept measures a discrepancy in the expected return, and this expected return is abnormal if it is too large compared to the corresponding volatility. This is exactly what is done when computing the conditional Sharpe performance.

2.4.3. Other current regressors

Very often the list of current regressors include different types of variables such as current values of macroeconomic variables or nonlinear transformations of current returns. As seen in the discussion above, the lagged values of any variables can be introduced when they belong to the information set of the retail investor and current value of only portfolio returns to fix the benchmark.

For instance, a nonlinear effect of market return (i.e. tracker return) \(y_{mt}\) cannot be considered by regressing \(y_{jt}\) on \(y_{mt}\) and computing the associated fitted performance. This is the so-called market timing ability introduced by Treynor and Mazuy (1966) and proposed to practitioners by Douday (2007). This approach is misleading. However, we can capture the information effect by regressing \(y_{jt}\) on \(y_{mt}\), \(y_{mt-1}\), \(y_{at}\) taking into account \(y_{mt}\), \(y_{at}\) in the information set and selecting the market portfolio as a benchmark portfolio. A second effect can be captured by regressing \(y_{jt}\) on \(y_{mt}\), \(y_{at}\), where \(y_{at}\) is the tracker return and \(y_{jt}\) is the return of the at-the-money option written on the market index. Then the benchmark assets are the tracker and the at-the-money option. It can also be captured by regressing \(y_{jt}\) on \(y_{mt}\), \(y_{mt-1}\), \(y_{mt-2}\), \(y_{at}\), \(y_{at-1}\), and so on. A similar remark applies when a call payoff \((\text{call}_\text{call} - k)^{+}\) is included among the regressors (see Agarwal and Naik (2004), Diez de los Rios and Garcia (2005)).

2.4.4. Differences regressors for different funds

It has also been proposed to introduce different regressors for the funds considered in the comparison to compute the performances and perform the ranking. More precisely, given two funds, 1 and 2, say, the conditional individual performances are computed as \(S_{t1}(Z_1)\) and \(S_{t2}(Z_2)\), respectively, and compared. As mentioned in sections above, the same information set has to be used for the comparison. This means that we have to compute \(S_{t1}(Z_1)\) and \(S_{t2}(Z_2)\) to get a first ranking, \(S_{t1}(Z_1)\) and \(S_{t2}(Z_2)\) to get a second one, \(S_{t1}(Z_2)\) and \(S_{t2}(Z_2)\) to get a third one. For instance, fund specific information sets are often introduced in the conditional performances, when univariate ARCH models are estimated separately for the different funds. In this approach the regressors are lagged (squared) returns of fund \(j\) only, when the conditional performance of \(j\) is considered. This misleading methodology is in particular followed in the ARCH based risk model developed by J.P. Morgan (see Daul (2007)).

A similar remark applies to fund specific benchmark portfolios (see, e.g. Amenc and Le Sourd (2005)). This is the core of the Edhec-Europerformance fund rating methodology. By choosing a benchmark portfolio depending on the announced portfolio management category, the authors introduce a selectivity bias in the ranking comparison of hedge funds.

2.4.5. Utility based performance

It has also been proposed to compare the funds by means of a subjective price, which can be computed by either a utility based method (see, e.g. Morningstar (2006) and Koekekbakker and Zaka-
First, there is an effect of the portfolio size. Second, the risk aversion associated with the Sharpe performance approach in two respects. Clearly, the values and ranking differ from the values and ranking by the (conditionally) Gaussian assumption. When the portfolio in mean–variance efficient, we deduce from (4) that the utility based price of the future portfolio value $w_{t+1}$ is:

$$P_t = \frac{1}{A} \log E_t [\exp (-A W_{t+1})]$$

$$= \frac{1}{A} \log \left[ \exp \left( -AE_t w_{t+1} + \frac{1}{2} V_t w_{t+1} \right) \right]$$

$$= E_t (w_{t+1}) - \frac{A}{2} V_t [w_{t+1}].$$

By the (conditionally) Gaussian assumption, when the portfolio in mean–variance efficient, we deduce from (4) that the utility based price corrected for the initial budget level is:

$$\frac{P_t}{W_t} = (1 + r_t) + \frac{1}{2A} m_t \Sigma^{-1} m_t.$$

Clearly, the values and ranking differ from the values and ranking associated with the Sharpe performance approach in two respects. First, there is an effect of the portfolio size. Second, the risk aversion coefficient $A$ (i.e. the selected utility function) matters. It is preferable to propose performance measures based on market data only, i.e. $m_t, \Sigma_t$, without introducing an ad hoc risk aversion parameter, which does not necessarily represent the retail investor’s preferences. To summarize, the Sharpe performance is an objective indicator, whereas a price computed from the utility function or an ad hoc change of probability is a subjective indicator.

### 2.4.6. Performance deduced from average ranks

Finally, another frequent practice is the following. A good fund is basically a fund with high mean returns, low volatility, low correlation with market returns (diversification effect) and positive autocorrelation (persistence effect). A naive way to get a unique ranking including all these features consists in first determining the rankings based on the mean, on the standard deviation, on the correlation with market return, etc., respectively, then averaging these rankings (see the Edhec-Europerformance methodology in Amenc and Le Sourd (2005)) to define a “performance” measure. Of course, the definition of this final ranking remains ad hoc, and clearly depends on the weights used when averaging. The numerical results are just averages of rankings, without any financial interpretation.

Moreover, whereas the Sharpe performance can be computed separately per fund, the approach based on average ranks requires the determination of the basic ranks, and then an analysis of all funds together, which is much more time consuming. Last, but not the least, this approach is highly sensitive to the population of funds which is considered.

## 3. Application to hedge funds

In this section we derive and compare the rankings of hedge funds according to a set of conditionally fitted Sharpe performance measures, computed with US $ as the money unit. In the first subsection, we describe the hedge funds database and provide summary statistics on hedge funds returns. In the two following sections we:

### Table 1

<table>
<thead>
<tr>
<th>Ticker</th>
<th>Fund name</th>
<th>Strategy</th>
<th>Manager</th>
<th>Fund assets</th>
</tr>
</thead>
<tbody>
<tr>
<td>GU EH</td>
<td>Exante Investors Gulliver Fund</td>
<td>Equity hedge</td>
<td>Exante Structured Asset Management</td>
<td>918 048 000</td>
</tr>
<tr>
<td>IC EH</td>
<td>Ibis Capital, LP</td>
<td>Equity Hedge</td>
<td>Ibis Management, LLC</td>
<td>35 000 000</td>
</tr>
<tr>
<td>OAM EH</td>
<td>Odey European Inc.</td>
<td>Equity Hedge</td>
<td>Odey Asset Management Limited</td>
<td>637 070 000</td>
</tr>
<tr>
<td>PF EH</td>
<td>Platinum Fund Ltd.</td>
<td>Equity Hedge</td>
<td>Optima Fund Management</td>
<td>1 009 357 000</td>
</tr>
<tr>
<td>PM EH</td>
<td>PM Capital Australian Opportunities Fund Ltd.</td>
<td>Equity Hedge</td>
<td>PM Capital Limited</td>
<td>19 000 000</td>
</tr>
<tr>
<td>RA EH</td>
<td>RAB Europe Fund</td>
<td>Equity Hedge</td>
<td>RAB Capital plc</td>
<td>185 000 000</td>
</tr>
<tr>
<td>RB EH</td>
<td>Robeco Boston Partners Long/Short Equity, L.P.</td>
<td>Equity Hedge</td>
<td>Robeco Investment Management</td>
<td>45 990 000</td>
</tr>
<tr>
<td>SP EH</td>
<td>Sprott Opportunities Hedge Fund, L.P.</td>
<td>Equity Hedge</td>
<td>Sprott Asset Management Inc.</td>
<td>313 000 000</td>
</tr>
<tr>
<td>RC EH</td>
<td>Robbins Capital Partners, L.P.</td>
<td>Equity Hedge</td>
<td>T. Robbins Capital Management, LLC</td>
<td>25 000 000</td>
</tr>
<tr>
<td>FS EM</td>
<td>Fairfield Sentry Ltd.</td>
<td>Equity Market Neutral</td>
<td>Fairfield Greenwich Group</td>
<td>6 800 000 000</td>
</tr>
<tr>
<td>IEM</td>
<td>Invesco QLS Equity</td>
<td>Equity Market Neutral</td>
<td>Invesco Structured Products Group</td>
<td>806 461 000</td>
</tr>
<tr>
<td>LC EN</td>
<td>Large Core Equity</td>
<td>Equity Non-Hedge</td>
<td>First Quadrant L.P.</td>
<td>3 063 300 000</td>
</tr>
<tr>
<td>TR EN</td>
<td>Thames River European Fund</td>
<td>Equity Non-Hedge</td>
<td>Thames River Capital LLP</td>
<td>42 888 960</td>
</tr>
<tr>
<td>LES SS</td>
<td>Leveraged Short Equity Index Hedge L.P.</td>
<td>Short Selling</td>
<td>Derivative Consulting Group L.L.C.</td>
<td>3 000 000</td>
</tr>
<tr>
<td>APM MA</td>
<td>APM Global Fixed Income Composite Fund</td>
<td>Macro</td>
<td>Absolute Plus Management, LLC</td>
<td>1 770 000 000</td>
</tr>
<tr>
<td>AS MA</td>
<td>Aspect Diversified Fund Limited</td>
<td>Macro</td>
<td>Aspect Capital Limited</td>
<td>1 420 000 000</td>
</tr>
<tr>
<td>FO MA</td>
<td>FORT Global Contrarian L.P.</td>
<td>Macro</td>
<td>Fort, Inc.</td>
<td>694 721 000</td>
</tr>
<tr>
<td>FX MA</td>
<td>FX Concepts Global Currency Program</td>
<td>Macro</td>
<td>FX Concepts, Inc.</td>
<td>3 600 000 000</td>
</tr>
<tr>
<td>HC MA</td>
<td>Haider Jupiter International Ltd.</td>
<td>Macro</td>
<td>Haider Capital Management, LLC</td>
<td>80 387 000</td>
</tr>
<tr>
<td>ML MA</td>
<td>Maple Leaf Macro Volatility Fund</td>
<td>Macro</td>
<td>Maple Leaf Capital</td>
<td>653 000 000</td>
</tr>
<tr>
<td>PE MA</td>
<td>Permal Global Opportunities Ltd. A</td>
<td>Macro</td>
<td>Permal Asset Management, Inc.</td>
<td>77 000 000</td>
</tr>
<tr>
<td>WI MA</td>
<td>Winton Futures Fund</td>
<td>Macro</td>
<td>Winton Capital Management</td>
<td>4 440 000 000</td>
</tr>
<tr>
<td>DF MF</td>
<td>Discus Fund Limited</td>
<td>Managed Futures</td>
<td>Capital Fund Management</td>
<td>47 989 382</td>
</tr>
<tr>
<td>CG RV</td>
<td>Clinton MultiStrategy Fund</td>
<td>Relative Value Arbitrage</td>
<td>Clinton Group, Inc.</td>
<td>685 000 000</td>
</tr>
<tr>
<td>BF RV</td>
<td>Endeavour Fund L.P.</td>
<td>Relative Value Arbitrage</td>
<td>Endeavour Capital LLP</td>
<td>2 767 000 000</td>
</tr>
<tr>
<td>WE RV</td>
<td>Western Investment Institutional Partners, LLC</td>
<td>Relative Value Arbitrage</td>
<td>Western Investment, LLC</td>
<td>62 000 000</td>
</tr>
<tr>
<td>AC CA</td>
<td>Advent Convertible Arbitrage Fund, L.P.</td>
<td>Convertible Arbitrage</td>
<td>Advent Capital Management, LLC</td>
<td>4 000 000 000</td>
</tr>
<tr>
<td>AI CA</td>
<td>Aristeia International, Ltd.</td>
<td>Convertible Arbitrage</td>
<td>Aristeia Capital LLC</td>
<td>2 746 479 000</td>
</tr>
<tr>
<td>PI MA</td>
<td>Paulson International Ltd.</td>
<td>Merger Arbitrage</td>
<td>Paulson &amp; Co., Inc.</td>
<td>3 373 000 000</td>
</tr>
<tr>
<td>CSS ED</td>
<td>Courage Special Situations Offshore Fund, Ltd.</td>
<td>Event-Driven</td>
<td>Courage Capital Management, LLC</td>
<td>350 567 000</td>
</tr>
<tr>
<td>RO ED</td>
<td>Roseau Limited Partnership</td>
<td>Event-Driven</td>
<td>Roseau Asset Management Ltd.</td>
<td>157 263 000</td>
</tr>
<tr>
<td>TT ED</td>
<td>TT Event-Driven Fund</td>
<td>Event-Driven</td>
<td>TT International</td>
<td>990 006 144</td>
</tr>
<tr>
<td>BF FI</td>
<td>BlackRock Fixed Income Global Opportunities Fund LLC</td>
<td>Fixed Income Arbitrage</td>
<td>BlackRock Financial Management, Inc.</td>
<td>1 333 000 000</td>
</tr>
<tr>
<td>BR FI</td>
<td>Blue River Advantaged Muni Fund</td>
<td>Fixed Income Arbitrage</td>
<td>Blue River Asset Management</td>
<td>1 600 000 000</td>
</tr>
<tr>
<td>DM FI</td>
<td>Drake Global Opportunities Fund, Ltd.</td>
<td>Fixed Income Arbitrage</td>
<td>Drake Management LLC</td>
<td>3 896 000 000</td>
</tr>
<tr>
<td>RCG DS</td>
<td>RCG Carpathia Overseas Fund, Ltd.</td>
<td>Distressed Securities</td>
<td>Ramius Capital Group, LLC</td>
<td>458 000 000</td>
</tr>
</tbody>
</table>
We select randomly 36 single hedge funds in different categories and report information on the management company, the self-declared strategy\(^{11}\) and the assets under management (see Table 1). We only select single hedge funds and do not consider funds of funds. Indeed, it might be more profitable for a skilled retail investor to manage a portfolio of single funds by himself, in order to avoid the additional fees corresponding to the payment of the fund of funds manager. We introduce the notion of ticker to shortly name each fund in the sample. It is not surprising to find 14 hedge funds in the Equity Long/Short categories (i.e. Equity Hedge, Equity Market Neutral, Equity Non-Hedge and Short Selling). Indeed, this is the most common investment strategy used by hedge fund managers. It consists in buying and selling short equities to generate returns or specific markets. Relative Value Arbitrage, Convertible Arbitrage, Merger Arbitrage, Event-Driven, Fixed Income Arbitrage and Managed Futures correspond to directional strategies on global or specific markets. Relative Value Arbitrage, Convertible Arbitrage, Merger Arbitrage, Event-Driven, Fixed Income Arbitrage and Distressed Securities are relative value strategies, consisting essentially on bets made on abnormal valorisation spreads.

Table 2 provides different summary statistics for the selected hedge funds returns. We compute annualized mean excess return, annualized volatility, current betas and lagged gammas with equi-
ty and bond markets. More precisely, the regressions include a single explanatory variable that is either the current value of the benchmark return to capture the fitted aspect, or its lagged value to capture the information effect. The corresponding regression coefficients are called current beta and lagged gamma, respectively. The leverage effect is the first explanation. The portfolio return and volatility. The corresponding results are provided in Table 3 and 4 give a first illustration of the impact of information sets. To be in line with industry practices, we report annualized Sharpe ratios rather than individual Sharpe performances, not to Sharpe ratios.

### 3.2. Comparison of unfitted performances

In this section and the following one, the theoretical framework of Section 2 is used to construct several rankings from the statistics computed on each fund. Let us first consider an investor with no competing holdings. For each fund, we compute the estimated individual Sharpe performance $S_j$, the estimated conditional Sharpe performances $S_j(Z)$ using different information sets $Z$. To be in line with industry practices, we report annualized Sharpe ratios instead of individual Sharpe performances. Even if the ranking associated with Sharpe ratio $S_j$ and Sharpe performance $S_j$ are identical, the additive property of fitted performances applies to Sharpe performances, not to Sharpe ratios.

### Tables 3 and 4

<table>
<thead>
<tr>
<th>Ticker</th>
<th>Hist. Perf.</th>
<th>Conditional performance</th>
<th>Parameters</th>
<th>SP = +1%</th>
<th>SP = −1%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\gamma$ (%)</td>
<td>$\delta$ (%)</td>
<td>$\theta$ (%)</td>
</tr>
<tr>
<td>FS EM</td>
<td>2.57</td>
<td>1</td>
<td>−6.9</td>
<td>0.3</td>
<td>1.3</td>
</tr>
<tr>
<td>PM EH</td>
<td>2.32</td>
<td>2</td>
<td>−4.8</td>
<td>1.2</td>
<td>5.9</td>
</tr>
<tr>
<td>SP EH</td>
<td>2.23</td>
<td>3</td>
<td>−5.8</td>
<td>2.1</td>
<td>11.5</td>
</tr>
<tr>
<td>WE RV</td>
<td>2.02</td>
<td>4</td>
<td>−2.0</td>
<td>0.6</td>
<td>3.6</td>
</tr>
<tr>
<td>FO MA</td>
<td>1.65</td>
<td>5</td>
<td>−13.8</td>
<td>1.0</td>
<td>6.4</td>
</tr>
<tr>
<td>I EM</td>
<td>1.68</td>
<td>6</td>
<td>−24.9</td>
<td>1.2</td>
<td>7.6</td>
</tr>
<tr>
<td>PE MA</td>
<td>1.68</td>
<td>7</td>
<td>−16.6</td>
<td>1.1</td>
<td>8.7</td>
</tr>
<tr>
<td>DF MF</td>
<td>1.67</td>
<td>8</td>
<td>−61.5</td>
<td>2.7</td>
<td>16.3</td>
</tr>
<tr>
<td>TT ED</td>
<td>1.63</td>
<td>9</td>
<td>−20.4</td>
<td>0.9</td>
<td>7.5</td>
</tr>
<tr>
<td>RO ED</td>
<td>1.57</td>
<td>10</td>
<td>−80.7</td>
<td>3.2</td>
<td>19.7</td>
</tr>
<tr>
<td>BB EH</td>
<td>1.57</td>
<td>11</td>
<td>6.0</td>
<td>0.7</td>
<td>5.8</td>
</tr>
<tr>
<td>TR EN</td>
<td>1.49</td>
<td>12</td>
<td>−12.5</td>
<td>1.3</td>
<td>9.4</td>
</tr>
<tr>
<td>RCG DS</td>
<td>1.49</td>
<td>13</td>
<td>26.0</td>
<td>0.8</td>
<td>7.1</td>
</tr>
<tr>
<td>PE MA</td>
<td>1.46</td>
<td>14</td>
<td>−12.4</td>
<td>1.3</td>
<td>10.5</td>
</tr>
<tr>
<td>AI CA</td>
<td>1.35</td>
<td>15</td>
<td>6.6</td>
<td>0.3</td>
<td>3.3</td>
</tr>
<tr>
<td>LC EN</td>
<td>1.34</td>
<td>16</td>
<td>−17.6</td>
<td>1.0</td>
<td>8.1</td>
</tr>
<tr>
<td>HC MA</td>
<td>1.24</td>
<td>17</td>
<td>−3.8</td>
<td>0.8</td>
<td>7.8</td>
</tr>
<tr>
<td>GSE ED</td>
<td>1.21</td>
<td>18</td>
<td>8.5</td>
<td>0.5</td>
<td>4.9</td>
</tr>
<tr>
<td>LES SS</td>
<td>1.15</td>
<td>19</td>
<td>12.3</td>
<td>1.1</td>
<td>11.2</td>
</tr>
<tr>
<td>DM FI</td>
<td>1.10</td>
<td>20</td>
<td>−30.0</td>
<td>1.2</td>
<td>11.5</td>
</tr>
<tr>
<td>RA EH</td>
<td>1.09</td>
<td>21</td>
<td>−12.5</td>
<td>0.9</td>
<td>9.1</td>
</tr>
<tr>
<td>PE EH</td>
<td>1.08</td>
<td>22</td>
<td>5.3</td>
<td>0.6</td>
<td>7.1</td>
</tr>
<tr>
<td>BR FI</td>
<td>1.01</td>
<td>23</td>
<td>15.2</td>
<td>0.7</td>
<td>9.6</td>
</tr>
<tr>
<td>APM MA</td>
<td>0.99</td>
<td>24</td>
<td>−1.4</td>
<td>0.6</td>
<td>6.6</td>
</tr>
<tr>
<td>WI MA</td>
<td>0.99</td>
<td>25</td>
<td>−2.4</td>
<td>1.2</td>
<td>14.6</td>
</tr>
<tr>
<td>CG RV</td>
<td>0.98</td>
<td>26</td>
<td>2.0</td>
<td>0.3</td>
<td>4.2</td>
</tr>
<tr>
<td>PF MA</td>
<td>0.96</td>
<td>27</td>
<td>1.6</td>
<td>0.9</td>
<td>10.7</td>
</tr>
<tr>
<td>GU EH</td>
<td>0.96</td>
<td>28</td>
<td>−3.0</td>
<td>0.2</td>
<td>2.5</td>
</tr>
<tr>
<td>ML MA</td>
<td>0.95</td>
<td>29</td>
<td>7.4</td>
<td>0.4</td>
<td>5.8</td>
</tr>
<tr>
<td>OAM EH</td>
<td>0.90</td>
<td>30</td>
<td>1.8</td>
<td>0.7</td>
<td>9.2</td>
</tr>
<tr>
<td>BF FI</td>
<td>0.82</td>
<td>31</td>
<td>−1.7</td>
<td>−0.1</td>
<td>1.0</td>
</tr>
<tr>
<td>AS MA</td>
<td>0.78</td>
<td>32</td>
<td>−1.8</td>
<td>0.8</td>
<td>12.6</td>
</tr>
<tr>
<td>IC EH</td>
<td>0.48</td>
<td>33</td>
<td>−33.3</td>
<td>0.7</td>
<td>11.4</td>
</tr>
<tr>
<td>RC EH</td>
<td>0.38</td>
<td>34</td>
<td>−42.5</td>
<td>−0.3</td>
<td>16.6</td>
</tr>
<tr>
<td>AC CA</td>
<td>0.17</td>
<td>35</td>
<td>11.5</td>
<td>0.0</td>
<td>3.8</td>
</tr>
<tr>
<td>EF RV</td>
<td>0.12</td>
<td>36</td>
<td>−17.6</td>
<td>0.1</td>
<td>4.1</td>
</tr>
</tbody>
</table>
columns 2 and 3 of Tables 3 and 4. The most performant fund is FS EM, with a Sharpe ratio equal to 2.57. When lagged SP500 returns are included in the definition of the conditional Sharpe performance, we use observed past return of this equity market benchmark to rank funds. We distinguish two different regimes of performance, we use observed past return of this equity market benchmark to rank funds. In Table 3 according if the SP500 is up 1% or the SP500 is down 1%. In both cases, FS EM remains the most performant hedge fund, but with a higher performance ratio in down market than in up market. We can also discuss the case of RCG DS. It is a medium fund if we consider its individual performance (Sharpe ratio of 1.49 – rank 8), but with a higher performance ratio in down market than in up market. Let us still consider the same 36 funds as before, and an investor with initial holdings. The initial holdings are either a portfolio fully invested in US Equity, denoted $P = E$ (the portfolio return is equal to the SP500 return), or a portfolio fully invested in US Bonds, denoted $P = B$ (the portfolio return is equal to the Lehmann Bonds index return). These two portfolios are proxies for the usual investment strategies followed by retail investors. For example, for life insurance products, three risk profiles are often defined: dynamic, diversified and defensive. The Equity [resp. Bond] portfolio can be associated with the dynamic [resp. defensive] profile and the diversified one corresponds to an equally weighted allocation between the Equity and Bond portfolios.\(^{13}\) We compute estimated Sharpe performance $S_{jP}$ and estimated conditionally fitted Sharpe performance $S_{jP}$ for each of these initial holdings $P = E$ or $B$.

Table 5 provides ranking of the set of hedge funds when the initial holding $P$ is the Equity portfolio. The left part of the Table (Fitted performance) can be first compared to the column providing the historical performance. If FS EM is still among the most performant funds, other funds have their rank significantly modified. For example, LES SS improves its ranking due to its negative beta (−151%) with the equity market. This characteristics is intrinsic to the strategy – short selling, and ensures a good protection for life insurance products, three risk profiles are often defined: dynamic, diversified and defensive. The Equity [resp. Bond] portfolio can be associated with the dynamic [resp. defensive] profile and the diversified one corresponds to an equally weighted allocation between the Equity and Bond portfolios.\(^{13}\) We compute estimated Sharpe performance $S_{jP}$ and estimated conditionally fitted Sharpe performance $S_{jP}$ for each of these initial holdings $P = E$ or $B$. 

3.3. Comparison of fitted performances

Let us still consider the same 36 funds as before, and an investor with initial holdings. The initial holdings are either a portfolio fully invested in US Equity, denoted $P = E$ (the portfolio return is equal to the SP500 return), or a portfolio fully invested in US Bonds, denoted $P = B$ (the portfolio return is equal to the Lehmann Bonds index return). These two portfolios are proxies for the usual investment strategies followed by retail investors. For example, for life insurance products, three risk profiles are often defined: dynamic, diversified and defensive. The Equity [resp. Bond] portfolio can be associated with the dynamic [resp. defensive] profile and the diversified one corresponds to an equally weighted allocation between the Equity and Bond portfolios.\(^{13}\) We compute estimated Sharpe performance $S_{jP}$ and estimated conditionally fitted Sharpe performance $S_{jP}$ for each of these initial holdings $P = E$ or $B$.

Table 5 provides ranking of the set of hedge funds when the initial holding $P$ is the Equity portfolio. The left part of the Table (Fitted performance) can be first compared to the column providing the historical performance. If FS EM is still among the most performant funds, other funds have their rank significantly modified. For example, LES SS improves its ranking due to its negative beta (−151%) with the equity market. This characteristics is intrinsic to the strategy – short selling, and ensures a good protection for life insurance products, three risk profiles are often defined: dynamic, diversified and defensive. The Equity [resp. Bond] portfolio can be associated with the dynamic [resp. defensive] profile and the diversified one corresponds to an equally weighted allocation between the Equity and Bond portfolios.\(^{13}\) We compute estimated Sharpe performance $S_{jP}$ and estimated conditionally fitted Sharpe performance $S_{jP}$ for each of these initial holdings $P = E$ or $B$.

\(^{13}\) We miss in this setup the real estate exposure met in most of retail investor portfolios (which is left for future research) and only consider liquid investments.
equity investors. More generally, funds with low absolute beta w.r.t. equity markets improve their ranking. On the contrary, I EM, announced as equity market neutral, features a surprising high equity beta, and is ranked 15 instead of 6 regarding historical performance only. Finally, due to the decomposition formula (12), the fitted performance is strongly dependent of the fund. They have to be accounted for, according to the need of the investors. To summarize, S&P500 and Lehmann US Agg are market factors with instantaneous effect on the hedge fund returns. These effects are strongly dependent of the fund. They have to be accounted for, according to the need of the investors.

We can also include lagged values of market factors in the information set. The right part of Table 5 concerns conditionally fitted performance and reports the ranking when lagged and diversification effects are both included. We only consider one regime: past SP500 return is up to 1%. FS EM is no longer the most performant fund and LES SS is now the fund ranked no. 1. This latter fund takes advantage of its negative beta with SP500 returns and positive correlation with lagged SP500 returns.

The same type of analysis can be done for investors who complete a portfolio including bonds. Table 6 includes the correlation with bond market in the ranking. As before FS EM benefits from its lack of correlation with the bond market and increases its ranking compared to the static case. The other funds do not have their rank significantly modified. On the contrary, we recover some changes of ranking if we consider conditionally fitted performances. This result contradicts that “other performance measures (than the standard Sharpe ratio) results in virtually identical rank ordering across hedge funds” (see Eling and Schuhmacher (2007)).

### 3.4. Term structure of performance

Retail investors can differ in their holding horizons often larger than of banks and it is important to analyse how the performance depends on this horizon (see, e.g. Levy (1972), Chen and Lee (1986)), Levy and Samuelson (1992) and Hodges et al. (1997)). The performances depend on the selected investment horizon by means of the definition of the returns and the lag in conditioning variables. By considering a varying horizon, we construct the term structures of (unconditional/conditional, fitted/unfitted) Sharpe ratios and the associated term structures of ranking.

In the simple case of i.i.d. (continuously compounded) returns, the conditional and unconditional performances at horizon \( h \) coincide and are equal to \( \frac{\Delta h}{\Delta t} \bar{r}_h = \bar{h}, \) since the mean and variance are both multiplied by \( h. \) Thus, in the i.i.d. framework, we get a linear term structure of performance, a square root pattern of Sharpe ratios and a flat term structure of ranks. The term structure pattern is often very different in practice, due to more complex return dynamic. The recent literature on hedge funds emphasized the importance of persistence in hedge fund returns. While some authors (see, e.g. Agarwal and Naik (2004) and Edwards and Caglayan (2001)) have found overall evidence of persistence, the results are...
Sharpe performance at horizon $h$ is:

$$S_h = \frac{[E(y_{j,t+1} + \cdots + y_{j,t+h})]^2}{V(y_{j,t+1} + \cdots + y_{j,t+h})}$$

$$= \frac{h^2E(y_{j,t+1})}{hS_h}$$

$$= \frac{1 + 2(1 - \frac{1}{h})\rho(1) + \cdots + 2(1 - \frac{1}{h})\rho(h - 1)}{1 + 2(1 - \frac{1}{h})\rho(1) + \cdots + 2(1 - \frac{1}{h})\rho(h - 1)}$$  \hspace{1cm} (20)

where $\rho(h)$ denotes the autocorrelation at order $h$ of the geometric returns. Thus, the term structure of Sharpe performances provides the complete information on the serial correlations of return. Since the compounding has been neglected, we cannot really consider: $\lim_{h \rightarrow \infty}(1/hS_h) = S_\infty = [1 + 2(1 - \frac{1}{h})\rho(1) + \cdots + 2(1 - \frac{1}{h})\rho(h - 1)]^{-1},$ as the long term unconditional Sharpe performance (see Stutzer (2000) for the effect of compounding in the long run). Note also that the interpretation of the term structure of Sharpe performance in terms of autocorrelation (i.e. Eq. (20)) is no longer valid for conditional/fitted performances.

As an illustration, we provide in Fig. 1 the term structure of historical Sharpe performances for seven funds. The term structures feature various patterns. They can be increasing, or decreasing, convex or concave, possibly with bumps. We observe crossing of the term structures of Sharpe ratios, which will imply non-flat term structures of rating and very different rankings for short and long term investors. The term structure of Sharpe performance can provide information on the strategy of the fund, especially on the part of its allocation which is illiquid and on the restricted investments.

If the fund FS EM is ranked number 1 at all horizons, some others funds can suffer when the horizon considered in the ranking procedure increases. It is in particular true for SP EH, third fund at horizon 1, but worst fund when the horizon is higher than 4. This is due to a higher volatility estimate when we use multiperiod returns in the estimation procedure. The i.i.d. assumption is in this case clearly broken, and autocorrelation between returns can explain this result.

As an illustration, we provide in Fig. 1 the term structure of historical Sharpe performances for seven funds. The term structures feature various patterns. They can be increasing, or decreasing, convex or concave, possibly with bumps. We observe crossing of the term structures of Sharpe ratios, which will imply non-flat term structures of rating and very different rankings for short and long term investors. The term structure of Sharpe performance can provide information on the strategy of the fund, especially on the part of its allocation which is illiquid and on the restricted investments.
4. Segmentation

The conditional fitted performance measures are summary statistics, which can be used to construct homogenous segments of hedge funds. These segments can be based on one or several performance measures. Using cluster analysis to get homogenous segments of hedge funds has already been proposed by Das (2003) and Gibson and Gyger (2007). A recent study by BNY Mellon (2007) emphasizes the importance of controlling the hedge funds strategies self-declared by the managers and currently used by the industry without any validation by the regulator.

4.1. The principle

4.1.1. Segmentation based on a single performance measure

Let us choose a given type of performance measure, such as the unconditional unfitted performance measure $S_{jt}$, $j=1,\ldots,N$. Then, $K$ segments, denoted $A_k$, $k=1,\ldots,K$, can be defined by introducing performance thresholds. Segment $A_k$ includes the funds such that: $a_{k-1} < S_{jt} < a_k$, where $a_0 = 0 < a_1 < \ldots < a_{K+1} = \infty$ are given thresholds. For a fixed number $K$ of classes, these thresholds can be selected to get the best contrast between segments. Let us denote by $S_k$ the average performance of the $N_k$ funds belonging to class $k$ and $\bar{S}$ the average performance computed on the whole population. The total variance can be decomposed as the sum of the between and within variances:

$$\frac{1}{N} \sum_{j=1}^{N} (S_{jt} - \bar{S})^2 = \frac{1}{N} \sum_{k=1}^{K} \sum_{j\in A_k} (S_{jt} - S_k)^2 + \frac{1}{N} \sum_{k=1}^{K} N_k (S_k - \bar{S})^2. \quad (21)$$

An optimal segmentation is derived by looking for segments which are as homogenous as possible. It is obtained by minimizing the between component (or equivalently by maximizing the between component). The clustering solution of the optimisation problem is easily derived by following a recursive approach, which explains how to merge two intermediate segments while increasing the criterion value. This procedure starts with the $N$ segments including only one hedge fund, then merges the two closest hedge funds and so on. It is known that this optimal approach provides segments corresponding to intervals of performance values, or equivalently to an optimal partition $a_0 = 0 < a_1 < \ldots < a_{K+1} = \infty$.

The standard procedures (and softwares) propose also methodologies for selecting an “optimal” number $K$ of classes, based on the analysis of the ratio between the within and total variances. Since hedge fund performances can be seen as rating of hedge fund managers, it seems preferable to follow the standard practice suggested by the regulators for risk and to impose a standardized number of classes, between 8 and 10. These classes will be the analogues of the standard AAA, AA, etc.

We can extend the approach to conditional performance measures. Since these measures depend on time by means of the environment, it is important to distinguish between environment dependent and environment independent clustering. More precisely, we can consider an environment dependent variance decomposition formula:

$$\frac{1}{N} \sum_{j=1}^{N} (S_{jt} - \bar{S})^2 = \frac{1}{N} \sum_{j=1}^{N} \sum_{k=1}^{K} (S_{jt} - S_k)^2 + \frac{1}{N} \sum_{k=1}^{K} N_k (S_k - \bar{S})^2, \quad (22)$$

and apply optimal clustering to get environment dependent partition $a_{0t} < a_{1t} < \ldots < a_{K+1t}$, for any $t$. Alternatively, we can look for an environment independent variance decomposition formula:

$$\frac{1}{NT} \sum_{j=1}^{N} \sum_{t=1}^{T} (S_{jt} - \bar{S})^2 = \frac{1}{NT} \sum_{t=1}^{T} \sum_{j=1}^{N} (S_{jt} - S_k)^2 + \frac{1}{NT} \sum_{k=1}^{K} N_k (S_k - \bar{S})^2, \quad (23)$$

where:

$$\bar{S} = \frac{1}{T} \sum_{t=1}^{T} \bar{S}_t, \quad S_k = \frac{1}{T} \sum_{t=1}^{T} S_{kt}, \quad N_k = \frac{1}{T} \sum_{t=1}^{T} N_{kt},$$

and derive an optimal clustering $a_0 < a_1 < \ldots < a_{K+1}$. The first approach is the most appropriate, since the definition of segments is

---

10 In the variance decomposition formula, the different hedge funds are equi-weighted. Alternative decompositions could account for either the size of the hedge funds, or the accuracy of the estimated performance measures.
intuitively not the same for high volatility periods and low volatility periods, or during the expansion and recession phases of the cycle. The standard regulation for risk (Basel 2) proposes in a first step to select environment independent thresholds for the regulatory analysis of risk, but allows a point-in-time (PIT) approach to account for the position within the cycle when using the internal model. A similar strategy might be followed for performance-based segmentation.

A similar discussion could be done concerning the self-declared strategies. Indeed, these strategies are currently declared at the issuing of the hedge fund and not regularly updated later on. They are environment independent, and the information on the strategies could have been improved by producing environment-dependent self-declared strategies. This is especially important when some strategies, such as Short Selling, are automatically inappropriate in some environment and leading to big losses if really followed. This question is clearly out of the scope of this paper.

To summarize, there exist as many segmentations as underlying Sharpe performance measures and environments. These segmentations have to be considered jointly and compared to the self-declared strategies.

4.1.2. Segmentation based on several performance measures

As above, we can derive optimal homogenous segments by minimizing the within component in a “variance decomposition equation”. The criterion has to involve jointly the different performance measures $S_i, i = 1, \ldots, L$, say. For this purpose, let us introduce an Euclidean norm: $||S||^2 = (S_1, \ldots, S_L)^T \Sigma^{-1} (S_1, \ldots, S_L)$, where $(\cdot)^T$ denotes transpose and $\Sigma$ is a symmetric positive definite matrix defining the selected metric. In practice $\Sigma$ is often selected as $\Sigma = ld$, where ld denotes the identity matrix, or as an estimated variance–covariance matrix of the estimated performance to account for the correlation structure between the different ranking criteria. We can decompose the total inertia into the between and within inertia as:

$$1 \sum_{j=1}^{N} \frac{1}{N} ||S_j - S||^2 - 1 \sum_{k=1}^{K} \frac{1}{N_k} \sum_{k 
eq k'} ||S_k - S_k'||^2 + \frac{1}{N} \sum_{k=1}^{K} N_k ||S_k - S||^2.$$ (24)

Whereas the optimal segments based on a single performance measure correspond to consecutive intervals and are ranked without ambiguity, the optimal segments based on several performance measures can have various patterns. The optimal segments are not obtained in general by crossing intervals based on the single performance measures, and cannot be totally ordered. A performance-based classification provides different results than a classification performed on other summary statistics such as return histories (see, e.g. Brown and Goetzmann (2003)), or binary attributes (see, e.g. Das (2003)).

4.2. Application to hedge funds

Let us illustrate the segmentation approaches described in Section 4.1. For this purpose, we consider the set of hedge funds for which the return data are complete. We get 2294 funds and compute five performance measures, that are the historical Sharpe Ratio, the SP500 Fitted Sharpe Ratio, the SP500 conditionally Fitted Sharpe Ratio, the Bond Fitted Sharpe Ratio and the conditionally Bond Fitted Sharpe Ratio. For this illustration, the conditional performances are computed for a fixed return of the conditioning variable (up to 1%) and they are not averaged over time.

4.2.1. Segmentation based on a single performance measure

Let us first consider the unconditional performance measure $S_i$. Then, we apply the automatic segmentation approach described in Section 4.1 for different numbers of clusters. Table 7 gives the

<table>
<thead>
<tr>
<th>Table 7</th>
<th>Inertia decomposition as function of the number of clusters.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>Total inertia</td>
<td>Within cluster inertia</td>
</tr>
<tr>
<td></td>
<td>Value</td>
<td>%</td>
</tr>
<tr>
<td>5</td>
<td>1.598</td>
<td>0.304</td>
</tr>
<tr>
<td>6</td>
<td>1.598</td>
<td>0.235</td>
</tr>
<tr>
<td>7</td>
<td>1.598</td>
<td>0.170</td>
</tr>
<tr>
<td>8</td>
<td>1.598</td>
<td>0.127</td>
</tr>
<tr>
<td>9</td>
<td>1.598</td>
<td>0.094</td>
</tr>
<tr>
<td>10</td>
<td>1.598</td>
<td>0.080</td>
</tr>
<tr>
<td>11</td>
<td>1.598</td>
<td>0.068</td>
</tr>
<tr>
<td>12</td>
<td>1.598</td>
<td>0.057</td>
</tr>
<tr>
<td>13</td>
<td>1.598</td>
<td>0.046</td>
</tr>
<tr>
<td>14</td>
<td>1.598</td>
<td>0.038</td>
</tr>
<tr>
<td>15</td>
<td>1.598</td>
<td>0.031</td>
</tr>
</tbody>
</table>

associated variance decomposition. The number $K = 10$, which is suggested by the regulator for credit risk, seems also appropriate in performance analysis, since the gain in inertia becomes rather small for larger value $K$. We fix $K = 10$ in further analysis. Table 8 displays the cluster characteristics by increasing order of performances. For each cluster, we give the number of funds, the lowest, the highest and the mean performances. The clusters can have rather heterogeneous sizes, but the clusters with small size correspond often to extreme performances. This explains why it is not necessarily useful to merge them with others classes. A similar analysis can be done for the conditionally (fitted) performance measures. We consider the fitted performances with SP500 benchmark, the fitted performances with Bond benchmark, the conditionally fitted performances with SP500 benchmark and the conditionally fitted performances with Bond benchmark. Tables 9 and 10 display the distribution of funds among the segments of both types, conditional and unconditional. The number of large segments, i.e. including more than 100 funds, is 6, 5, 6, 5, 5, respectively, and these segments correspond to different limiting
thresholds. In particular, the ratings based on two different performances are not comparable, that is, the ranks of the clusters cannot be compared between Tables 8 and 9 (or Table 10). For instance, rating 2 (i.e. AA in the standard terminology) corresponds to segments with size 8, 2, 8, 10, 2, respectively, and to intervals [7.8, 0.87] respectively. Even if the 1-dimensional performance direction. We observe that the projected clusters define intervals of performances which generally intersect. Table 11 provides the inertia decomposition related to this case. The clusters are defined in a 5-dimensional space of Sharpe ratios and can admit rather complicated patterns. Table 12 provides the projections of the clusters on each 1-dimensional performance direction. We observe that the projected clusters define intervals of performances which generally intersect. Table 13 provides the distribution by self-declared strategy for each cluster. It is difficult from Table 13 to see if the self-declared strategy is informative of the future results, at least when they are measured in terms of Sharpe performance. The existence of several styles with highly different performances show that “the term hedge fund encompasses investment philosophies that range far from the risk-neutral strategy of (Alfred Winslow) Jones” (see Jones (1949), Brown and Goetzmann (2003)). Moreover, “many investment firms are simply renaming their trading desk as hedge funds”. They satisfy the regulatory controls concerning leverage, Short Selling, holding shares of other investment companies, but they have a complete freedom, otherwise, especially for using the self-declared strategies as a marketing tool.

To perform a more accurate analysis, we consider a chi-square test for independence. A chi-square test can only be applied if the size of the marginal cells are large enough. For this reason, we consider 7 aggregate clusters {1,2,3,4}, 5, 6, 7, 8, 9, 10 and 5 management styles, i.e. Equity Hedge, Equity Market Neutral, Macro, Managed Futures, Fixed Income Arbitrage. Let us denote by \( Y_{ij} \) the number of funds in aggregate cluster \( i \) and aggregate category \( j \). We want to test the null hypothesis of independence:

\[
H_0 : p_{ij} = P_i j, \quad i = 1, \ldots, 7, \quad j = 1, \ldots, 5,
\]

where \( p_{ij} \) (resp. \( P_i \)) denotes the probability of being in aggregate cluster \( i \) and aggregate category \( j \) (resp. in cluster \( i \), in category \( j \)). We then compute the chi-square test statistic:

\[
Q = \sum_{i=1}^{7} \sum_{j=1}^{5} \frac{(Y_{ij} - N_i P_j)^2}{N_i P_j},
\]

where \( N \) is the total number of funds, \( p_i \) and \( p_j \) the frequency counterpart of \( P_i \) and \( P_j \), respectively. The terms in the chi-square sum are provided in Table 14.

Under the null hypothesis, the test statistic \( Q \) follows a chi-square distribution with \((7-1)(5-1) = 24\) degrees of freedom. We must reject \( H_0 \) at a 5% level if \( Q > 36 \). We get \( Q = 364 \) and then reject the independence hypothesis. Therefore, the performance of a particular fund depends on the self-declared strategy. A more detailed analysis can be done by comparing each element of Table 14 with the critical value 3.86 corresponding to a chi-square with 1 degree of freedom. The main reason for the global dependence between style and performance is the high performances of fixed arbitrage funds.

This result supports the top-down approach used in the most common funds of funds allocation process. First, the fund manager
defines a strategic allocation between strategies, and then picks funds inside each strategy. If the independence between strategies and performance were not rejected, it would be possible to directly define an optimal allocation in selecting the best performing funds, without any consideration on the strategies. This approach is called bottom-up in the asset management industry.

5. Concluding remarks

The notion of Sharpe performance is easy to interpret in the mean–variance portfolio framework. By considering different information sets, benchmark portfolios and horizons, this measure provides a battery of coherent ranking procedures. These rankings are easy to implement by performing the appropriate regressions and computing the corresponding Sharpe performance measures, as illustrated in the application to hedge funds. The analysis of a battery of various Sharpe performances is clearly the first approach to be applied for hedge fund comparison, especially for retail investors. Moreover, considering jointly several performance measures avoids largely the possibility of manipulation discussed in Goetzmann et al. (2007).\(^{17}\) Indeed, a manipulation to improve one of the measure can have a negative impact on the other ones. In particular, manipulations by means of information-free dynamic strategy are highly diminished when the whole term structure of performance is considered.

However, the Sharpe performance measures considered in this paper have at least three limitations:

(i) They assume that the riskfree asset is introduced in the investor portfolio. Without the riskfree asset the formula for the performance decomposition has to be modified and the expression of the fitted performance is not so simple.

(ii) They assume that short sell is allowed for the retail investor. Under short-sell restrictions, we have to solve the optimization problem under positivity constraints on the allocation. This complicates a lot the formulas of the performance and fitted performance. However, this can highlight the discussion on the importance on taking into account the skewness of excess return distribution for retail investor. If short sell is allowed, the retail investor can balance negative expected excess return by introducing negative allocation of the asset. Thus, with a mean–variance behavior, it is not very sensitive to the skewness features. At the contrary, the effect of skewness significantly appears if the Sharpe performance is computed under no short-sell restriction. More generally, the standard Sharpe performance is based on myopic portfolio optimization; it does not incorporate the dynamic aspects of the fund manager’s investment process, that are the allocation updating frequency, the constraints on allocations and its implicit objective function. Intuitively, some benchmark active investment management should have been defined for each management style.

(iii) They suppose that the (conditionally, fitted) mean and variance are appropriate summary statistics for the location and risk parameters of the historical, conditional or residual return distributions. This is not necessarily the case when these distributions feature high skewness and fat tails, since the sample mean is very sensitive to outliers, and the variance is not an appropriate measure of risk. High kurtosis and skewness are generally observed on the unconditional distribution\(^{18}\) of hedge funds, which can follow dynamic management strategies. Other historical performance mea-

---

\(^{17}\) Typically, they assume that the funds are ranked by a single valued measure (see condition 1, p. 1506), which eliminate partial ranking associated with a battery of measures.

\(^{18}\) However, fat tails and asymmetries are often highly reduced for conditional distributions.
sures have been proposed in the literature (see Farinelli et al. (2008) in the context of optimal allocation). They are generally obtained by considering a modified ratio, in which the variance (resp. the mean) is replaced by another measure of risk (resp. location). The measure of risk can be based on a Value-at-Risk (see, e.g. Dowd (1999), Amenc and Le Sourd (2005)), a semi-standard deviation (see, e.g. Roy (1952), Bawa (1975), Levy and Levy (2009)), the so-called Sortino ratio: (see Sortino and Van der Meer (1991) and the survey by Pedersen and Satchell (2002)), historical prices of put and call, as in the Omega measure (see, e.g. Keating and Shadwick (2002)), a Tail-Var (see, e.g. Gourieroux and Liu (2007)), or L-moments (see Darolles et al. (2009)). However, all these alternative measures can be implemented in the industry, if they allow a coherent analysis of historical/conditional, fitted/unfitted measures, as Sharpe performance measures. Moreover, it is difficult to believe that they will be understood by retail investors and use in an appropriate way in their portfolio management.

The Sharpe performance based rankings assume that the investor is able to reallocate efficiently his/her portfolio. These rankings can be misleading for an investor, who is allocating his/her portfolio in a naive way. For instance, if two funds have to be included in a portfolio, they do not have to be selected by looking for the largest sum of individual performances $S_j + S_i$, but for the largest sum $S_j + S_{ij} = S_i + S_{ji}$ to correct for the possible dependence between returns. To avoid misunderstanding of hedge fund rankings, which are already largely diffused by generalist or specialized newspapers, it is important to explain carefully how they are derived and how they can be used. In this respect, the knowledge of Sharpe performance, that is $m_i/\sigma_i$ does not allow to compute the new optimal allocation, that is $m_i/\sigma_i^2$. For deriving this optimal allocation separate information on $m_i$ and $\sigma_i^2$ is also required and generally not provided. In other words, if a high Sharpe ratio is a good thing, a portfolio strategy cannot be based on the Sharpe ratio only.

Acknowledgement
The authors gratefully acknowledge financial support of the chair Groupama/Risk Foundation: “Behavioral Finance”.

Appendix A

A.1. Unconditional versus conditional performance
The unconditional Sharpe ratio is defined by:

$$Sr = \frac{E(y_{i,t})}{V(y_{i,t})}$$

Let us denote by $m_i$ [resp. $\sigma_i^2$] the conditional expectation $E(y_{i,t})$ [resp. $V(y_{i,t})$] the conditional variance. By the iterated expectation theorem, we get: $E(y_{i,t}) = Em_i$, and by the variance decomposition equation: $V(y_{i,t}) = V(m_i) + E(\sigma_i^2)$. We deduce that:

$$Sr = \frac{E(m_i)}{\sqrt{V(m_i) + E(\sigma_i^2)}} = \frac{E(m_i)}{\sqrt{V(m_i) + E(\sigma_i^2)}} \frac{E(\sigma_i)}{E(\sigma_i)} = E_{\sigma_i}(Sr_i)\left[\frac{V[\sigma_i(Sr_i) + E(\sigma_i^2)]}{E(\sigma_i)}\right]^{1/2}$$

where $Sr_i = m_i/\sigma_i$ is the conditional Sharpe ratio, $Q_i$ a modified measure with density $\sigma_iE(\sigma_i)$ and the factor $E(\sigma_i)(V[\sigma_iSr_i] + E(\sigma_i^2))^{1/2}$ is smaller than 1, by Cauchy-Schwarz inequality. Thus, the knowledge of the conditional performance $Sr_i$ is not sufficient to derive the historical performance. In fact the volatility $\sigma_i$ has also to be observed.

A.2. Expression of the fitted performance

We have:

$$S_{r_{ij,t}} = \frac{1}{\sigma_i^2 - \sigma_{ij}}(m_i, m_j) \left(\frac{\sigma_j^2 - \sigma_{ij}}{\sigma_i^2} \right) \left(\frac{m_j}{m_i} \right) - m_i^2$$

where $\sigma_i$ and $\sigma_{ij}$ differ from $\hat{\sigma}_i$ and $\hat{\sigma}_{ij}$ by the coefficient $1 - \sum_{t=1}^T V_y_{i,t}/V_y_{i}$, This multiplicative coefficient depends on the regressor only. In particular, $\hat{S}_{r_{ij,t}}$ and $\hat{\sigma}_{ij}$ provide the same ranking.

References