The Challenge of Hedge Fund Performance Measurement: a Toolbox Rather Than a Pandora's Box

December 2006

Walter Géhin
Research Associate, EDHEC Risk and Asset Management Research Centre
Business Analyst, Atos Euronext Market Solutions
This paper, which is being written to provide an overview of the multitude of publications we have seen on hedge fund performance, is the result of a reading and analysis of about 200 studies on this subject. About 50 publications in the most famous journals and working papers written by recognized authors have been selected to provide a dynamic and comprehensive view of the improvements in hedge fund performance measurement.

The issue of performance measurement in the hedge fund industry has led to literature that is both abundant and controversial. The explanation of this complexity lies in the particular features of alternative funds. Hedge funds invest in a heterogeneous range of financial assets and cover a wide range of strategies that have different risk and return profiles. Even though the current studies on hedge fund performance appear to be confusing, due to conflicting conclusions and criticism of the methods employed in previous papers, they contribute to an improvement in our understanding of alternative funds and help to confirm the validity of leading approaches. The aim of this paper is to highlight some specific characteristics of hedge funds and their implications in terms of performance measurement.

Abstract

This paper, which is being written to provide an overview of the multitude of publications we have seen on hedge fund performance, is the result of a reading and analysis of about 200 studies on this subject. About 50 publications in the most famous journals and working papers written by recognized authors have been selected to provide a dynamic and comprehensive view of the improvements in hedge fund performance measurement.

The issue of performance measurement in the hedge fund industry has led to literature that is both abundant and controversial. The explanation of this complexity lies in the particular features of alternative funds. Hedge funds invest in a heterogeneous range of financial assets and cover a wide range of strategies that have different risk and return profiles. Even though the current studies on hedge fund performance appear to be confusing, due to conflicting conclusions and criticism of the methods employed in previous papers, they contribute to an improvement in our understanding of alternative funds and help to confirm the validity of leading approaches. The aim of this paper is to highlight some specific characteristics of hedge funds and their implications in terms of performance measurement.
Walter Géhin is a Research Associate with the EDHEC Risk and Asset Management Research Centre. He is currently working at Atos Euronext Market Solutions as a Business Analyst. He holds a master’s degree in banking and finance and an advanced graduate diploma in financial engineering. Walter has published articles on performance persistence, hedge fund performance measurement, hedge fund market capacity, and hedge fund indexes. He is in charge of the hedge fund performance research section of the www.edhec-risk.com website.
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td>5</td>
</tr>
<tr>
<td>1. PREREQUISITE: QUALITY OF THE DATA</td>
<td>6</td>
</tr>
<tr>
<td>1.1. Factors of accuracy</td>
<td>6</td>
</tr>
<tr>
<td>1.2. Biases</td>
<td>6</td>
</tr>
<tr>
<td>1.2.1. Survivorship bias</td>
<td>6</td>
</tr>
<tr>
<td>1.2.2. Instant history bias</td>
<td>7</td>
</tr>
<tr>
<td>1.2.3. Selection bias</td>
<td>7</td>
</tr>
<tr>
<td>1.2.4. State price bias</td>
<td>8</td>
</tr>
<tr>
<td>1.3. Testing the return adequacy</td>
<td>8</td>
</tr>
<tr>
<td>2. ABSOLUTE PERFORMANCE MEASUREMENT</td>
<td>9</td>
</tr>
<tr>
<td>2.1. Traditional absolute performance measures</td>
<td>9</td>
</tr>
<tr>
<td>2.1.1. Sharpe and Treynor ratios</td>
<td>9</td>
</tr>
<tr>
<td>2.1.2. Theoretical problems</td>
<td>9</td>
</tr>
<tr>
<td>2.1.2.1. Non normality of hedge fund return distributions</td>
<td>10</td>
</tr>
<tr>
<td>2.1.2.2. Presence of autocorrelation in hedge fund return series</td>
<td>14</td>
</tr>
<tr>
<td>2.2. Innovative absolute performance measures based on the Sharpe ratio</td>
<td>10</td>
</tr>
<tr>
<td>2.2.1. Autocorrelation-adjusted Sharpe ratio</td>
<td>10</td>
</tr>
<tr>
<td>2.2.2. Modified Sharpe ratio</td>
<td>11</td>
</tr>
<tr>
<td>2.3. Innovative absolute performance measures not based on the Sharpe ratio</td>
<td>12</td>
</tr>
<tr>
<td>2.3.1. Stutzer index</td>
<td>12</td>
</tr>
<tr>
<td>2.3.2. Omega</td>
<td>12</td>
</tr>
<tr>
<td>2.3.3. Sharpe-Omega</td>
<td>13</td>
</tr>
<tr>
<td>2.3.4. AIRAP</td>
<td>14</td>
</tr>
<tr>
<td>2.3.5. Kappa</td>
<td>14</td>
</tr>
<tr>
<td>3. MODELLING: IDENTIFICATION OF HEDGE FUND RETURN SOURCES</td>
<td>16</td>
</tr>
<tr>
<td>3.1. Linear multi-factor models including linear factors</td>
<td>16</td>
</tr>
<tr>
<td>3.1.1. Explicit micro-factor model</td>
<td>16</td>
</tr>
<tr>
<td>3.1.1.1. Size of the fund</td>
<td>16</td>
</tr>
<tr>
<td>3.1.1.2. Age of the fund</td>
<td>17</td>
</tr>
<tr>
<td>3.1.1.3. Manager tenure</td>
<td>18</td>
</tr>
<tr>
<td>3.1.1.4. Performance fees</td>
<td>18</td>
</tr>
<tr>
<td>3.1.1.5. Other fund factors</td>
<td>19</td>
</tr>
<tr>
<td>3.1.1.6. Comments</td>
<td>19</td>
</tr>
<tr>
<td>3.1.2. Explicit macro-factor model</td>
<td>19</td>
</tr>
<tr>
<td>3.1.3. Implicit factor model</td>
<td>20</td>
</tr>
<tr>
<td>3.2. Adaptation of the models to the non-linearity of hedge fund returns</td>
<td>21</td>
</tr>
<tr>
<td>3.2.1. Linear models including non-linear regressors</td>
<td>21</td>
</tr>
<tr>
<td>3.2.1.1. Option portfolios as non-linear regressors</td>
<td>21</td>
</tr>
<tr>
<td>3.2.1.2. Hedge fund indices as non-linear variables</td>
<td>21</td>
</tr>
<tr>
<td>3.2.2. Conditional approaches</td>
<td>22</td>
</tr>
<tr>
<td>3.2.2.1. Methods</td>
<td>22</td>
</tr>
<tr>
<td>3.2.2.2. Results</td>
<td>23</td>
</tr>
<tr>
<td>3.3. Higher-moment-adjusted CAPM</td>
<td>23</td>
</tr>
<tr>
<td>4. PERFORMANCE REPLICATION</td>
<td>25</td>
</tr>
<tr>
<td>4.1. Passive Replication</td>
<td>25</td>
</tr>
<tr>
<td>4.2. Dynamic Replication</td>
<td>26</td>
</tr>
<tr>
<td>4.2.1. Optimization-based replication</td>
<td>26</td>
</tr>
<tr>
<td>4.2.2. Copula-based replication</td>
<td>27</td>
</tr>
<tr>
<td>5. PERFORMANCE PERSISTENCE</td>
<td>29</td>
</tr>
<tr>
<td>5.1. First approach: persistence of relative returns</td>
<td>29</td>
</tr>
<tr>
<td>5.1.1. Test methods</td>
<td>29</td>
</tr>
<tr>
<td>5.1.1.1. Two-period framework</td>
<td>29</td>
</tr>
<tr>
<td>5.1.1.2. Multi-period framework, Kolmogorov-Smirnov test</td>
<td>30</td>
</tr>
<tr>
<td>5.1.2. Overview of performance persistence studies</td>
<td>31</td>
</tr>
<tr>
<td>5.1.3. Results at short-term horizons</td>
<td>32</td>
</tr>
<tr>
<td>5.1.4. Results at long-term horizons</td>
<td>32</td>
</tr>
<tr>
<td>5.1.5. Comments on relative persistence results</td>
<td>32</td>
</tr>
<tr>
<td>5.1.6. Disadvantages of a relative persistence approach</td>
<td>32</td>
</tr>
<tr>
<td>5.2. A manoeuvrable approach: pure persistence</td>
<td>33</td>
</tr>
<tr>
<td>5.2.1. Hurst exponent combined with a D-Statistic</td>
<td>33</td>
</tr>
<tr>
<td>5.2.2. Pure persistence results</td>
<td>33</td>
</tr>
<tr>
<td>5.2.3. Comments on pure persistence results</td>
<td>35</td>
</tr>
<tr>
<td>5.3. Persistence is not predictability</td>
<td>35</td>
</tr>
<tr>
<td>CONCLUSION</td>
<td>37</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>38</td>
</tr>
</tbody>
</table>
This paper, which is being written to provide an overview of the multitude of publications we have seen on hedge fund performance, is the result of a reading and analysis of about 200 studies on this subject. About 50 publications in the most famous journals and working papers written by recognized authors have been selected to provide a dynamic and comprehensive view of the improvements in hedge fund performance measurement.

The literature on hedge fund performance is both controversial and abundant. Its controversy springs from the numerous qualities specific to hedge funds. As a result, certain studies are devoted to the inadequacy of traditional approaches when applied to this universe. In parallel, new performance indicators and models have been introduced. The abundance of this literature can be explained by the wide acceptance of the fact that performance measurement is a key point of the quantitative analysis required in a rigorous fund selection process.

Due to the number of hedge funds that skyrocket, short track records, the heterogeneity of hedge fund strategies, the fact that hedge fund managers are not equal in talent, or the opacity of the hedge fund universe, fund picking is more than a challenging task. Quantitative analysis applied to hedge funds must be sophisticated and requires real expertise.

The basis of a quantitative fund selection process consists of adequate absolute performance indicators. In a hedge fund context, risk adjustment plays a primordial role because of the specificities of return distributions. However, absolute performance indicators only give a static picture of performance over a given period. It is not sufficient to rank funds. This must be completed by indicators of performance persistence that make it possible to identify funds that display a stable positive performance.

Another criterion for ranking hedge funds is the measure of the manager's talent. The distinction between this talent, known as the alpha, and the return generated by exposures to different market factors, namely the betas, is based on modelling. Moreover, the identification of the different return sources allows the style analysis to be reinforced, giving a clear understanding of what strategy is being pursued.

At first sight, the conclusions of studies that scrutinize hedge fund performance seem to be confusing. It is true that conflicting conclusions and criticism of the methods employed in previous papers are frequent. However, by highlighting leading approaches, the debate contributes to our understanding of hedge funds.

The first section of this paper discusses the importance of database quality and the impact of the different biases on returns. The second section investigates absolute performance measurement, the third section looks at modelling and the fourth focuses on performance replication. The fifth and final section is dedicated to performance persistence.
Before we look at performance measurement, the choice of an accurate database is of great interest in the context of the hedge fund industry, where a lack of transparency is often observed. Performance measurement based on an inaccurate database is unreliable in all cases. Moreover, many biases can affect performance.

1. Prerequisite: quality of the data

1.1. Factors of accuracy
Liang (2003b) enumerates some factors that have a positive effect on the quality of the database. First, funds that are audited effectively have lower absolute return discrepancies than those for which audit dates are missing. Second, the respect of a transparency principle is a reliable indicator of the quality of the data. It appears that funds of hedge funds report returns more accurately than single hedge funds. When comparing onshore funds with their offshore twins (the only difference is fund location), audited pairs reveal significantly less return discrepancy than non-audited pairs. Finally, significant positive correlation between hedge fund size and the auditing variable appears: large funds are more frequently audited than small funds.

When comparing the returns given by TASS and the US Offshore Fund Directory with the percentage changes in the net asset values, the US Offshore Fund Directory exhibits an average discrepancy of 0.29 points per year, while TASS exhibits an average discrepancy of 0 points per year. It illustrates the fact that at a given date the quality of the databases is not homogeneous. The lack of constancy of each database is illustrated by the fact that for the same database vendor, the quality differs between versions. Two different versions of TASS returns are compared, one from July 31, 1999, and the other from March 31, 2001. 3,638 observations of 461 hedge funds are different across the two dates.

1.2. Biases
Hedge fund databases can potentially suffer from several biases which have a significant impact on performance measures. The most common biases are survivorship bias, instant history bias, selection bias and stale price bias.

1.2.1. Survivorship bias
Definition
Survivorship bias occurs if the database only contains information on 'surviving funds'. Those funds are in operation and report information to the database vendor at the end of the data sample. The opposite of these are defunct funds. They stop reporting because of bankruptcy or liquidation, for example. Good funds that close generate a downward bias on returns, while bad funds that fail generate an upward bias.

As mentioned by Amin and Kat (2003), survivorship bias also generates a downward bias in the standard deviation, an upward bias in the skewness and a downward bias in the kurtosis.

Evaluation
Following Malkiel (1995), the bias is evaluated via the difference in the performance of the "observable" portfolio (investment in each fund in the database from the beginning of the data sample) and the portfolio of surviving funds. Fung and Hsieh (2000) exhibit this to be about 3% per year. A similar result is found in Brown, Goetzmann and Ibbotson (1999).

In order to estimate survivorship bias, Chen and Ibbotson (2005) form six sub-samples: live funds only with backfill data, live funds only without backfill data, live and dead funds with backfill data, live and dead funds without backfill data, dead funds only with backfill data and dead funds only without backfill data. The survivorship bias is estimated from January 1995 to March 2004 on the TASS database. When the backfill data are included, the survivorship bias is 2.74% per year. When the backfill data
are excluded, it is 5.68%, demonstrating that backfill data lead to an underestimation of the survivorship bias.

Ammann and Moerth (2005) calculate the survivorship bias from January 1994 to June 2003, on the TASS database. They use successively asset-weighted and equally weighted returns. The survivorship bias is 2.44% with equally weighted returns, while it is 0.85% with asset-weighted returns. The difference is significant (5%). It implies that the survivorship bias comes to a large extent from the smallest funds.

Correction
To correct this bias, TASS has kept the returns of defunct funds since 1994 in its database. The same method has been applied by MAR Hedge since 1995. In the same vein, Caglayan and Edwards (2001) include 496 defunct hedge funds in their sample.

1.2.2. Instant history bias
Definition
Instant history bias (or backfill bias) is the consequence of adding a hedge fund whose earlier good returns are backfilled between the inception date of the fund and the date on which it enters the database, while bad track records are not backfilled.

Evaluation
This bias is evaluated by the difference between the return of an adjusted observable portfolio (the returns corresponding to the incubation period are dropped) and the return of a non-adjusted observable portfolio. An instant history bias of 1.4% per year is calculated by Fung and Hsieh (2000) for the TASS database over the period of 1994–1998.

Using an alternative method, Posthuma and van der Sluis (2003) eliminate an individual incubation period fund by fund for the TASS database over the period of 1996–2002. In the first scenario, which is based on the hypothesis that lockup periods and fund liquidation have no impact on returns, a backfill bias of 4.35% per year is found (all strategies are considered). In the second scenario, which is based on the hypothesis that lockup periods and fund liquidation result in additional negative impact of 50% on returns, the backfill bias is 7.24% per year. In the third scenario, which is based on the hypothesis that lockup periods and fund liquidation result in additional negative impact of 100% on returns, the backfill bias is 10.13% per year.

Chen and Ibbotson (2005) form six sub-samples: live funds only with backfill data, live funds only without backfill data, live and dead funds with backfill data, live and dead funds without backfill data, dead funds only with backfill data, dead funds only without backfill data. The database, provided by TASS, is from January 1994 to March 2004. They calculate the backfill bias successively on the basis of an equally weighted portfolio, a value-weighted portfolio and an equally weighted portfolio with only funds that have reported an amount for assets under management, from January 1995 to March 2004. The equally weighted portfolio exhibits a backfill bias of 4.84%. The equally weighted portfolio with only funds that have reported an amount for assets under management gives a backfill bias of 1.29%. The value-weighted portfolio reveals a backfill bias of 4.58%.

Correction
To correct this bias, Caglayan and Edwards (2001) exclude the first 12 months of returns for all funds in their sample.

1.2.3. Selection bias
Selection bias is generated when only funds with good performance want to be included in a database. However this upward bias is limited, because some high-performance managers do
not publish their performance. This might be the case when they have reached their goal in terms of assets under management or their target size. Fung and Hsieh (2000) therefore consider this bias to be negligible.

1.2.4. Stale price bias

Some of the instruments in which hedge funds invest have a low liquidity level. For these instruments, the market price is not always available. In order to report returns from all dates, the last price of the security is often used. This generates a stale price bias.

1.3. Testing the return adequacy

Surz (2005) tests the hypothesis that “Performance is good”. The actual performance is compared to all the possible outcomes previously evaluated ("what could have happened"). More precisely, the possible outcomes correspond to the possible portfolios that a hedge fund manager could have constituted. These portfolios are created on the basis of the investment parameters followed by the hedge fund manager, for example investment style, long and short positions, fees and leverage.

This method is presented as "a credibility check on manager performance". A reported performance that is not included in the simulated range of possible outcomes should be considered with scepticism.

A Monte Carlo Simulation (henceforth MCS) is applied to the Market Neutral strategy. Three sub-strategies are distinguished: "long value/short growth", "long growth/short value" and "style-neutral". For each sub-strategy, all the possible outcomes are evaluated over a five-year period ending June 30, 2004.

MCS provides a range of possible performance levels. For the "long growth/short value" sub-strategy it ranges from -16.4% to +3%. For the "short growth/long value" sub-strategy it ranges from +2.8% to +22.1%. For the "style-neutral" sub-strategy it ranges from -5.9% to +11.7%. The best possible "long growth/short value" manager discloses a performance of +3%, while the worst possible "long value/short growth" manager discloses +8.3%. This confirms that the performance of a manager has to be compared to that of a manager following the same strategy.

Surz underlines the implications of these results in the context of fund picking conducted by a hedge fund manager. By considering the market neutral strategy as a unique group, an above-median manager (actually included in the long value/short growth sub-strategy) is only in the bottom quartile of his sub-strategy group.
2. Absolute performance measurement

An initial step involves calculating a “raw” return, where contributions, withdrawals, interest, dividends accrued, gains/losses, accrued management fees and transactional fees are taken into account. For example, the Hedgeworks method is as follows:

\[
\text{return} = \frac{(i-e)^*(1-ifa))}{b}
\]

where b is the basis (prior period ending capital plus capital contributed or withdrawn at beginning of period), i is the income earned during the period (interest, dividends accrued, realized and unrealized gains/losses, other income), e is expenses accrued during the period (interest, dividends (short), accrued management fees, transactional fees, other fees) and ifa is the incentive fee adjustment (deduction if over high watermark; gross up or giveback of prior accrued if under high watermark).

However such a performance indicator is not sufficient, because it does not provide for risk-adjustment.

2.1. Traditional absolute performance measures

2.1.1. Sharpe and Treynor ratios

These measures are considered “absolute” because no benchmark is used to calculate them. The most common indicators are the Sharpe ratio (1966) and the Treynor ratio (1965).

The Sharpe ratio is formulated as follows:

\[
Sp = \frac{E(R_p) - R_f}{\sigma(R_f)}
\]

where E(Rp) is the expected return of the portfolio, Rf is the risk-free rate and \(\sigma(R_f)\) is the standard deviation of the portfolio returns.

The Treynor ratio is formulated as follows:

\[
Tp = \frac{E(R_p) - R_f}{\beta_p}
\]

where E(Rp) is the expected return of the portfolio, Rf is the risk-free rate and \(\beta_p\) is the beta of the portfolio.

2.1.2. Theoretical problems

Traditional indicators work when returns follow a symmetrical distribution. In that case, risk is represented by the standard deviation. Unfortunately, hedge fund returns are not normally distributed, and hedge fund return series are autocorrelated. Consequently, traditional performance measures suffer from theoretical problems when they are applied to hedge funds.

2.1.2.1. Non normality of hedge fund return distributions

**Skewness**

The skewness coefficient measures the asymmetry coefficient of the return distribution. For a series of N returns, skewness is equal to:

\[
S = \frac{1}{N\sigma^3} \sum (r_i - \chi)^3
\]

where \(r_i\) is the th return of the series, \(\chi\) is the mean of the returns and \(\sigma\) is the standard deviation.

**Kurtosis**

The kurtosis coefficient measures the tail depth of the return distribution. A high coefficient indicates the presence of extreme returns. Kurtosis is equal to:

\[
K = \frac{1}{N\sigma^4} \sum (r_i - \chi)^4
\]

where \(r_i\) is the th return of the series, \(\chi\) is the mean of the returns and \(\sigma\) is the standard deviation.

When the returns are normally distributed, kurtosis is equal to 3.

**Tests of normality**

In order to test the normality of a distribution, a Jarque-Bera test can be conducted. A normal

---

1 - The Jarque Bera test is not the only method to test normality. For example, the Shapiro-Wilk test for normality is appropriate in the case of a small sample.
2. Absolute performance measurement

distribution has skewness = 0 and kurtosis = 3. The Jarque-Bera statistic is given by:

$$JB = n \left[ \frac{\text{skewness}^2}{6} + \frac{(\text{kurtosis} - 3)^2}{24} \right]$$

where \( n \) is the number of observations in the sample period. This statistic has a chi-squared distribution (with two degrees of freedom) under the null hypothesis of normality.

2.1.2.2. Presence of autocorrelation in hedge fund return series

Autocorrelation impact

A position on illiquid assets generates autocorrelation in return series. Persistent price lags in the valuation of hedge funds have an impact on the accuracy of some performance measures. Lo (2002) documents that there is an overstatement of the Sharpe ratio in the case of positive autocorrelation of the hedge fund returns. According to his study, the presence of a serial correlation in monthly returns generates an overestimation of as much as 65% of the annual Sharpe ratio. Consequently, hedge fund rankings based on the Sharpe ratio can be dramatically wrong.

Tests of autocorrelation

The Ljung-Box test is one of the most well-known autocorrelation tests. It is formulated as follows:

$$Q = T (T + 2) \sum_{k=1}^{m} \frac{\tau_k^2}{T - k}$$

where \( T \) is the number of return observations.

This statistic has a chi-squared distribution under the null hypothesis of an absence of autocorrelation.

2.2. Innovative absolute performance measures based on the Sharpe ratio

The Sortino ratio (1994) provides a solution to the asymmetry of the return distribution by replacing the standard deviation with a downside deviation. Consequently, the Sortino ratio is more appropriate when returns are left-skewed.

This is the excess return over the risk-free rate over the downside semi-variance, so it measures the return to “bad” volatility. It is formulated as follows:

$$\text{Sortino ratio} = \frac{E(R_p) - \text{MAR}}{\sqrt{\frac{1}{T} \sum_{t=1}^{T} (R_{pt} - \text{MAR})^2}}$$

where \( R_{pt} \) is the return of the portfolio in the sub-period \( t \), \( R_{p} \) is the average of the returns of the portfolio over the whole period, \( \text{MAR} \) is the minimum acceptable return and \( T \) is the number of sub-periods.

While the Sortino ratio takes asymmetry into account, it does not solve the problem of kurtosis and the problem of autocorrelation. Adjustments of the Sharpe ratio have been proposed. Lo (2002) proposes an adjustment to the autocorrelation in return series, while Gregoriou and Gueyie (2003) propose an adjustment to the skewness and the kurtosis.

2.2.1. Autocorrelation-adjusted Sharpe ratio

Presentation

This indicator is recommended by Lo (2002) to avoid the overestimation of the Sharpe ratio due to the autocorrelation of the hedge fund returns. Liang (2003a) uses the autocorrelation-adjusted Sharpe ratio with the following terms:

$$\eta(q) \text{SR}$$

with

$$\eta(q) = \frac{q}{\sqrt{q + 2 \sum_{k=1}^{q-1} (q - k) \rho_k}}$$

Many other autocorrelation tests are available. Among them we can also cite the Herfindahl index.
where SR is the regular Sharpe ratio on a monthly basis, $\rho_k$ is the kth autocorrelation for hedge fund returns, and $\eta (q)\text{SR}$ is the annualised autocorrelation-adjusted Sharpe ratio with $q=12$.

**Empirical results**

On the basis of a database provided by Zurich Capital Markets, Liang (2003a) observes an annualised Sharpe ratio of 1.2505 and an annualised autocorrelation-adjusted Sharpe ratio of 1.0743 from 1998 to 1999. This period corresponds to a bull market. From 2000 to 2001 (corresponding to a bear market), the annualised Sharpe ratio is 0.0918 and the annualised autocorrelation-adjusted Sharpe ratio 0.1417. These results do not indicate that in bull markets (respectively in bear markets) the standard Sharpe ratio is always greater (less) than the autocorrelation-adjusted Sharpe ratio, but according to the period where the performance is measured, the autocorrelation of the hedge fund returns can have various impacts on the Sharpe ratio.

2.2.2. Modified Sharpe ratio

**Presentation**

Gregoriou and Gueyie (2003) propose an improvement to the original Sharpe ratio through the use of the Modified Value-at-Risk (MVaR). The new performance measure is known as the Modified Sharpe ratio.

The modified VaR replaces the standard deviation in the equation of the modified Sharpe ratio. It is defined as follows:

$$\text{Modified Sharpe Ratio} = \frac{(R_p - Rf)}{\text{MVaR}}$$

where $R_p$ is the return of the portfolio (i.e. a hedge fund or a fund of hedge funds), $Rf$ is the risk-free rate and MVaR is the modified VaR.

The replacement of the standard definition by the MVaR is justified by the fact that the latter takes skewness and kurtosis into account in addition to mean and standard deviation. It is of particular interest in the case of hedge funds in order to avoid underestimating risk. It should be noted that from this angle the VaR exhibits the same shortcomings as the standard deviation.

**Empirical results**

Gregoriou and Gueyie (2003) conduct an empirical application of the modified Sharpe ratio. The data, provided by Zurich Capital Markets, covers the period of January 1997 to December 2001. The whole sample contains monthly returns of 90 live funds of hedge funds, but only 30 funds are studied: the 10 funds with the highest assets under management, the middle 10 and the bottom 10 funds. The risk-free rate $Rf$ is assumed to be nil to simplify the ranking. The MVaR is calculated at a 95% confidence level.

Comparing the average of mean returns in each of the three groups, the top group (respectively bottom funds) exhibits the highest (lowest) mean return average. On the other hand, the most negative skewness is in the bottom group, where the standard deviation is also the highest. Considering the MVaR, the bottom funds display the highest in absolute value. In short, bottom funds are more frequently affected by extreme negative returns. Mostly, empirical results for the 30 selected funds confirm that a normal Sharpe ratio overestimates the performance in comparison with the modified Sharpe ratio, except when the normal Sharpe ratio is negative.

Differences in rankings obtained through the Sharpe ratio and the modified Sharpe ratio are examined by Gregoriou (2004) from January 1998 to December 2002 for 9 Canadian hedge funds. The MVaR is set at 95%. A risk-free rate of 0 is used to facilitate the rankings.

The 9 funds are divided into 3 equal groups according to assets under management (henceforth AuM), in other words the size of each fund. The top group contains the largest funds. It displays AuM of USD 236.4 million on average. The middle group displays AuM of USD 104.4 million on average and the bottom group USD 3.4 million on average.
Focusing on VaR and MVaR, the middle group exhibits the lowest levels. While the top group has the highest VaR, the bottom group has the highest MVaR. This confirms that the use of VaR or MVaR is not interchangeable, because it gives different rankings in terms of exposure to extreme market losses. Rankings based on the standard Sharpe ratio and the modified Sharpe ratio are similar. However, it is stated that the modified Sharpe ratio in the 3 groups is lower than the standard Sharpe ratio. These results show that the use of the standard Sharpe ratio provokes an underestimation of the extreme risk. It leads to an overestimation of risk-adjusted performance.

2.3. Innovative absolute performance measures not based on the Sharpe ratio

Several new performance measures are not based on the Sharpe ratio. They are innovative in that they attempt to take skewness and kurtosis into account.

2.3.1. Stutzer index

Presentation
The Stutzer index was introduced by Stutzer (2000). It is based on the behavioural hypothesis that investors aim to minimize the probability that the excess returns over a given threshold will be negative over a long time horizon. When the portfolio has a positive expected excess return, this probability will decay to zero at an exponential decay rate as the time horizon increases. It is equal to the maximum decay rate to zero of the expected excess return: the higher the Stutzer index, the longer the time horizon and the better the hedge fund.

The Stutzer index downgrades the ranking of funds whose skewness is strongly negative and whose kurtosis is strongly positive, while it upgrades the ranking of funds whose skewness is near zero and whose kurtosis is not strongly positive.

Empirical results
Bacmann and Scholz (2003) compare the rankings of 44 hedge fund indices with the Stutzer index and the Sharpe ratio. The database used, provided by CSFB/Tremont, HFR and Stark, covers the period of January 1994 to February 2003. 4 indices are drawn from the traditional universe (MSCI World Index, Russell 2000, S&P 500 and the Salomon World Government Bond Index). 15 indices are normally distributed according to the Jarque-Bera statistic at the 5% significance level.

In comparison with the Sharpe ratio, 37 funds have the same ranking according to the Stutzer index. However, if we consider the higher moments for the indices whose rank improves, the negative skewness turns positive in the case of the Stutzer index. The positive kurtosis decreases from 7.22 to 3.69. For the indices whose rank deteriorates, the negative skewness significantly increases from −0.82 to −2.95. The positive kurtosis increases strongly from 7.22 to 19.17.

In contrast to the previous results, ranks are similar when the authors only consider the traditional indices, whatever the performance measure. This confirms that higher moments are the source of the mismatch between the Sharpe ratio and the Stutzer index.

2.3.2. Omega

Presentation
The Omega measure was introduced by Keating and Shadwick (2002). It incorporates all the moments of the return distribution, including skewness and kurtosis. Moreover, in contrast to the Sharpe ratio, ranking is always possible, whatever the threshold. It requires no assumptions on the return distribution or on the utility function of the investor.
Omega is expressed as the ratio of the gain with respect to the threshold and the loss with respect to the same threshold:

\[
\Omega(L) = \frac{\int_a^b (1 - F(x))dx}{\int_a^b F(x)dx}
\]

where \(L\) is the required return threshold, \(a\) and \(b\) are the return intervals and \(F(x)\) is the cumulative distribution of returns below threshold \(L\). At a defined level of threshold, the higher the Omega the better.

Gupta, Kazemi, and Schneeweis (2003) give an intuitive expression of Omega:

\[
\Omega(L) = \frac{C(L)}{P(L)}
\]

where \(C(L)\) is essentially the price of a European call option written on the investment and \(P(L)\) is essentially the price of a European put option written on the investment.

De Souza and Gokcan (2004) provide the Omega formula in a discrete case:

\[
\Omega(L) = \frac{\sum_{i=1}^n \text{Max}(0, R^+)}{\sum_{i=1}^n \text{Max}(0, |R^-|)}
\]

where \(R^+\) (\(R^-\)) is the return above (below) a threshold \(L\).

**Empirical results**
Bacmann and Scholz (2003) compare the rankings of 44 hedge fund indices with the Omega and the Sharpe ratio.

In comparison with the Sharpe ratio, 36 funds have the same ranking according to the Omega, but if we consider the higher moments for the indices whose rank improves, the negative skewness decreases from -0.75 to -0.45. The positive kurtosis decreases from 7.18 to 4.09. For the indices whose rank deteriorates, the negative skewness significantly increases from -0.75 to -2.60. The positive kurtosis increases strongly from 7.18 to 16.85 in the case of the Omega.

Ranks are similar when only the traditional indices are considered. It confirms that the Sharpe ratio tends to underestimate or overestimate the performance results in the context of hedge funds.

### 2.3.3. Sharpe-Omega³

**Presentation**
Presented by Gupta, Kazemi and Schneeweis (2003), the Sharpe-Omega has identical features to the Omega, whilst keeping the same risk approach as the Sharpe ratio. It is introduced in the following way:

\[
\text{Sharpe - Omega} = \frac{\text{(expected return} - \text{threshold)}}{P(L)} = \frac{\text{put option price}}{\text{expected return} - \text{threshold}}
\]

This indicator has the particular quality of being proportional to \((1-\Omega)\). Consequently it provides strictly the same rankings as the Omega. Through numerical examples in the case of changes in the distribution of an investment’s return, the authors show that the Sharpe-Omega is most sensitive to the mean and the variance, and is less impacted by skewness and kurtosis.

**Empirical results**
Using monthly data from January 1994 to May 2003, Gupta et al. estimate the Omega and Sharpe-Omega for the S&P 500 index, the CSFB convertible arbitrage index and the CSFB equity market neutral index. For different levels of threshold, the two indicators give the same rankings for the three indices.

Sharpe-Omega is successively calculated by successively modifying only the mean and the threshold (while standard deviation = 5%, skewness = 0, kurtosis = 3), only the standard deviation and the threshold (while mean = 1%, skewness = 0, kurtosis = 3), only the skewness
and the threshold (while mean = 1%, standard deviation = 5%, kurtosis = 3), and only the kurtosis and the threshold (while mean = 1%, standard deviation = 5%, skewness = 0). It appears that changes in mean and standard deviation have the most pronounced impact on the Sharpe-Omega, confirming Keating and Shadwick’s (2002) conclusions on Omega.

2.3.4. AIRAP

Presentation


AIRAP is constructed on the basis of the Expected Utility theory. The selected form of utility is a Constant Relative Risk Aversion (CRRA). AIRAP is formulated as follows:

- when \( c \) (Arrow-Pratt coefficient) is different to 1 and greater than or equal to 0:

\[
\text{AIRAP} = \left[ \prod_i p_i \star (1 + \text{TR}_i)^{(1-c)} \right]^{\frac{1}{1-c}} - 1
\]

where \( \text{TR}_i = \frac{d\text{NAV}_i}{\text{NAV}_{t-1}} \) and \( p_i \) is the frequency of % returns.

- when \( c \) is equal to 1:

\[
\text{AIRAP} = \left[ \prod_i (1 + \text{TR}_i) \right]^{\frac{1}{N}} - 1
\]

Sharma recommends an Arrow-Pratt coefficient (represented by \( c \)) from 1 to 10. Because a geometric mean is used to measure the average performance, \( c = 1 \) corresponds to risk neutrality (in this case the risk premium is nil). Cases with \( c \) comprised between 0 and 1 assume that rational investors accept the risk of insolvency, but according to the author this is implausible. Adopting a cautious view, the author assumes \( c = 4 \). This corresponds to a case where investors accept a risk of a maximum loss of 20.7% of their wealth.

An approach that only involves using the ratio of gross and net assets is inadequate for taking into account the impact of leverage on the performance of hedge funds, because of the presence of derivatives. This justifies a risk-based approach. AIRAP captures the impact of leverage through a credit for the higher mean and a penalty for the higher volatility as a function of the CRRA parameter. The optimal leverage, which maximises AIRAP for a range of CRRA, can be defined by standard optimization techniques.

According to Sharma, AIRAP presents several advantages. It takes leverage and investor preferences into account. Unlike traditional risk-adjusted performance measures, AIRAP penalizes negative skewness and positive kurtosis. Moreover, it is scale invariant and can be used for non-directional strategies, unlike the Treynor ratio.

Empirical results

Data covers the period of January 1997 to December 2001. At the index level, the data is provided by EACM. At the individual fund level, the data is provided by HFR. Rank reversals between Sharpe and AIRAP and between Jensen’s alpha and AIRAP are presented with 19 different levels of Constant Relative Risk Aversion for the HFR universe.

The percentage of Sharpe ratio rank reversals is between 99% and 100%, while the percentage of Jensen’s alpha rank reversals is between 98% and 100%. The Spearman rank correlation confirms the lack of correlation between standard measures and the AIRAP. At the intra-strategy level, even if the rank reversal is somewhat lower, it also indicates discrepancies between the Sharpe ratio and AIRAP.

2.3.5. Kappa

Presentation

Kappa, introduced by Kaplan and Knowles (2004), is presented as a generalized downside risk-adjusted performance measure. “Generalized”
means that this indicator can become any risk-adjusted return measure, through a single parameter.

\[ K_n(\tau) = \frac{\mu - \tau}{\sqrt[n]{LPM_n(\tau)}} \]

where \( \mu \) is the expected periodic return, \( \tau \) is the investor’s minimum acceptable or threshold periodic return and LPM is the lower partial moment.

It becomes apparent that the Sortino ratio is equal to \( K_2 \), and Omega to \( K_1 + 1 \). \( n \) is strictly greater than 0.

Kappa can be calculated in two ways: it can use discrete return data or a parameter-based calculation. A discrete calculation gives robust results, but it is a strict requirement. A parameter-based calculation involves deriving a continuous return distribution from the values of the first four moments, i.e. mean, standard deviation, skewness and kurtosis.

**Empirical results**

Kaplan and Knowles test Kappa on a database provided by HFR that covers from January 1990 to February 2003. They focus on 11 hedge fund indices. Firstly, for each hedge fund strategy, Kappa is calculated with \( n \) being equal to 1 or 2, with a successive threshold of 0% or 1%. It is stated that the difference between the results obtained through the two methods (discrete or parameter-based) increases when the threshold decreases. In such cases, Kappa has to be handled cautiously.

Secondly, the rankings obtained through the two methods are compared, successively with \( n = 1, 2 \) and 3, and with a threshold of –1%, –0.5% and 0%. In terms of ranking, the parameter-based method provides similar results to the discrete method. The parameter \( n \) has the greatest impact on the ranking: only two strategies (Emerging Markets and Event-Driven) have the same ranking regardless of what \( n \) is, for a threshold of 0%.

With \( n \) being equal to 1, 2 or 3, an inverse relationship between the threshold and the value of Kappa appears. The steepness of the Kappa curve decreases when the parameter \( n \) increases. Considering the sensitivity of Kappa to skewness, when the threshold is above (below) the mean return, it is insensitive (sensitive). When Kappa is sensitive, it is a negative function of \( n \).
3. Modelling: identification of hedge fund return sources

The primary model is the Capital Asset Pricing Model (CAPM), initiated by Sharpe in 1964. It is a single factor model in which security prices are governed by their market risks and not their firm-specific risks. Based on a simple statistical regression framework using $T$ historical returns:

$$ R_{it} = \alpha_i + \beta_i R_{mt} + \epsilon_{it} $$

where $R_{it}$ is the return on a given portfolio (or fund) $i$, $\alpha_i$ is the abnormal performance of the portfolio (or fund) $i$, $\beta_i$ is the sensitivity of the portfolio (or fund) $i$ and $R_{mt}$ is the market return for the period.

In a close form, Jensen’s alpha (1968) is obtained via a regression on:

$$ R_{Pt} - R_{Ft} = \alpha_p + \beta_p (R_{Mt} - R_{Ft}) + \epsilon_{Pt} $$

where $R_{Pt}$ is the expected return of the portfolio, $R_{Ft}$ is the risk-free rate, $\beta_p$ is the beta of the portfolio, and $R_{Mt}$ is the expected market return.

This measure is considered “relative” because a benchmark is used to calculate it.

Due to the wide range of instruments and techniques used by hedge funds, a single-factor model presents the risk of overestimating the alpha. Consequently, multi-factor models are more appropriate. Multi-factor models can be specified in a linear or non-linear way. A higher-moment CAPM is an alternative.

### 3.1. Linear multi-factor models including linear factors

The general presentation of a multi-factor model is as follows:

$$ R_{it} = \alpha_i + \sum_{k=1}^{K} b_{ik} F_k + \epsilon_{it} $$

where $R_{it}$ is the return on a given portfolio (or fund) $i$, $\alpha_i$ is the abnormal performance of the portfolio (or fund) $i$, $b_{ik}$ the sensitivity of the portfolio (or fund) $i$ and $F_k$ the return on factor $k$ for the period.

Three approaches can be followed to specify the model:

- Explicit micro-factor model
- Explicit macro-factor model
- Implicit factor model

While explicit macro-factor models refer to asset-based style factors (henceforth ABS factors), implicit factor models refer to return-based style factors (henceforth RBS factors).

#### 3.1.1. Explicit micro-factor model

In an explicit micro-factor model, the selected factors refer to fund-specific features, such as size, age, manager tenure or performance fees.

##### 3.1.1.1. Size of the fund

Gregoriou and Rouah (2002) focus on the relationship between the size of hedge funds and their performance. The size of a fund is defined as the total asset amount at the start of the calculation period. The relationship between size and performance is tested by Pearson’s correlation coefficient and Spearman’s rank correlation from January 1994 to December 1999 on the basis of databases obtained from ZCM and LaPorte. Using the geometric mean, the Sharpe ratio and the Treynor ratio, the correlations are not statistically significant. The authors conclude that the size of a hedge fund (and of a fund of hedge funds) has no impact on its performance. However, they suggest testing this relationship again over a longer period, because some size factors are liable to harm performance, for example slower operations due to administrative duties.

Koh, Koh and Teo (2003) study this relationship for Asian hedge funds. Their results corroborate the previous results, with a non-significant relationship.

Brorsen and Harri (2004) find that returns decrease when the market capitalization...
increases. They provide the hypothesis that the funds are created to exploit market inefficiencies, and that the inefficiencies are finite. To maintain the performance, the managers have to close the funds to new investors.

De Souza and Gokcan (2003) exhibit through a regression on the TASS database that assets under management have a positive relationship with performance. According to them, this could imply that poor performing funds have difficulty attracting new contributions, or that large size allows lower average costs to be obtained.

Amenc and Martellini (2003) study the impact of various fund characteristics on performance on the basis of several models, such as the standard CAPM, an adjusted CAPM for the presence of stale prices and an implicit factor model extracted from a Principal Component Analysis. All models indicate that the mean alpha for large funds exceeds the mean alpha for small funds, with a large share of statistically significant differences.

Getmansky (2004) uses a regression on the TASS database that includes the size squared as a factor. A positive and concave relationship between current performance and past asset size is found. This suggests that an investor should select hedge funds that are near their optimal size.

Chen and Ibbotson (2005) rank funds on the basis of their assets under management. The database, provided by TASS, is from January 1994 to March 2004. It appears that the largest 1% of the funds outperforms all the other categories. The largest 10% of the funds outperform "the average by roughly 2 percentage points". They explain these results in two ways. First, managers of larger funds have better skills and, second, managers of larger funds need not consider constraints like "paying the bills".

Ammann and Moerth (2005) select 2,317 funds from the TASS database. They study the relationship between size and successively average returns, Sharpe ratio and alpha from January 1994 to June 2003. They state that the bottom percentiles, from the 1\textsuperscript{st} to the 20\textsuperscript{th} percentile (i.e. the smallest funds), display the lowest returns, while the funds from the 21\textsuperscript{st} to the 50\textsuperscript{th} percentile display the highest returns. A linear regression reveals a significant positive relationship between size and average returns, at the 1% level. A quadratic regression exhibits a significant concave relationship. A linear regression displays a significant positive relationship between size and the Sharpe ratio, at the 1% level. A quadratic regression reveals a significant concave relationship. To evaluate the alpha, a three-factor model is used. The factors are the Goldman Sachs Commodity Index, the Lehman Aggregate Bond Index and the Wilshire Micro Cap Index. A linear regression shows a significant positive relationship between size and alpha, at the 1% level. A quadratic regression exhibits a significant concave relationship.

3.1.1.2. Age of the fund
Howell (2001) investigates the relationship between the age of hedge funds and their performance, from 1994 to 2000. Young hedge funds are usually defined as those with a track record of less than three years. The first step was to adjust the returns by applying the probability of a failure to report to the surviving funds. This provides ex-post returns, which correspond to the true costs and benefits of investing in funds with different maturities. The second step was to adjust the returns by applying the probability of future survival to the survivors' returns by age decile. This gives ex-ante returns, which correspond to the expected returns from investing in hedge funds with different maturities. Ex-ante returns infer that young funds' returns are superior to those of seasoned funds: the youngest decile exhibits a return of 21.5%, while the whole sample median exhibits a return of 13.9% (a spread of 760 basis points in favour of young funds). Moreover, the spread between the decile of youngest funds and the decile of oldest funds is 970 points, and the spread between the second youngest fund decile and the whole sample median is 290 points. The conclusion of this study is that
hedge fund performance deteriorates over time, even when the risk of failure is taken into account. Consequently, the youngest funds seem particularly attractive.

In Amenc and Martellini (2003), it appears that for all the models used, newer funds (one or two years old) exhibit an alpha exceeding the alpha of the older funds. Nevertheless, the significance of the difference between the alphas varies across the models.

In contrast to these results, Koh, Koh and Teo (2003) find that fund age is not an explanatory factor for Asian hedge fund returns, using a cross-sectional Fama and MacBeth (1973) framework.

According to De Souza and Gokcan (2003), on the basis of a regression on the TASS database, older funds outperform younger funds on average.

3.1.1.3. Manager tenure

Boyson (2003) analyses the relationship between hedge fund manager tenure and fund returns. Regressions show that each additional year of experience is associated with a statistically significant decrease in annual returns of approximately -0.8%.

To explain the relationship between experience and performance in the light of risk-taking behaviour, Boyson successively examines the relationship between manager tenure and risk-taking behaviour and the relationship between risk-taking behaviour and returns.

Focusing on the relationship between manager tenure and risk-taking behaviour, three risk measures are used: the standard deviation of a portfolio's return, a tracking error deviation\(^5\) and a beta deviation\(^6\). It appears that an increase in manager tenure, fund size or tenure/size interaction engenders less risky behaviour.

Concerning the relationship between the risk-taking behaviour and the returns, each of the three risk measures is positively related to the annual returns. In other words, when manager tenure increases, risk-taking decreases, and when risk-taking decreases, returns decrease.

These results highlight the impact on hedge fund returns of increasing career concerns over time, with risk-taking behaviour characterised by increasing risk aversion. Career concerns in the hedge fund industry are unique in that they change over time. This is due to the sources of the manager's compensation, i.e. the assets under management and the returns. Young managers generally have a lower level of assets under management than older managers. Consequently, they take more risk to obtain good returns, while the large size of the fund provides older managers with their compensation. As a result, the risk level diminishes in accordance with the hedge fund manager's rising age. Moreover, statistics show that failed hedge fund managers rarely start a new hedge fund, and if they move into the mutual fund industry, for example, this is associated with a pay cut. The amount of the pay cut is more significant for older hedge fund managers, and it is thus an incentive for them to mitigate their risk-taking behaviour. A final explanation for the lower level of risk taken by an older hedge fund manager is the large amount of personal assets invested in the fund.

3.1.1.4. Performance fees

Kazemi, Martin and Schneeweis (2002) study the impact of performance fees for Value, Growth and Small styles. Their data show that fees have a poor effect on performance.

Koh, Koh and Teo (2003) find that funds with higher performance fees have smaller post-fee returns than funds with lower performance fees.

De Souza and Gokcan (2003) find that incentive fees and performance are positively correlated. Higher incentive fees generating higher performance can be explained by the fact that incentive fees are increased when a manager improves his performance, or by the fact that the best managers in terms of performance demand higher incentive fees.

In Amenc and Martellini (2003), it appears that for all the models used, funds exhibiting high incentive fees (greater than or equal to 20%) obtain a better alpha than the funds with low

---

5 - Measure of how much a manager's tracking error (i.e. the volatility in returns not explained by market volatility) differs from that of the average manager in the same style category.
6 - Difference between the fund's beta on the fund of funds index (i.e. each individual fund's time-series coefficient obtained from a regression of the fund's returns on the fund of funds index) and the average beta on the fund of funds index for all other funds in the same style category.
incentive fees. However, the implicit factor model indicates a non-significant difference.

3.1.1.5. Other fund factors
Koh, Koh and Teo (2003) examine other fund factors. They find that Asian hedge funds returns have a positive and significant relationship with the redemption period and the size of the holding company. The size of the minimum investment is not a significant factor.

Similarly, Kazemi, Martin and Schneeweis (2002) find that the redemption period affects returns. For a given strategy, funds with a quarterly lockup have higher returns than funds with a monthly lockup.

De Souza and Gokcan (2003) show that the investment by a manager of his own capital has a positive impact on performance, as with the lockup and redemption periods.

3.1.1.6. Comments
The results obtained by De Souza and Gokcan (2003) show that the specification of an explicit micro-factor model is a difficult and hazardous task. Setting a requirement level of 5% significance, six variables are kept in the model: managed assets, the age of the fund, partner-capital participation, lockup, the required redemption notice period and incentive fees. This model displays poor explanatory power with an adjusted-R² of 0.14. This poor explanatory power is corroborated by the heterogeneity of the conclusions on fund factors. It tends to indicate that explicit micro-factor models suffer from mis-specification.

3.1.2. Explicit macro-factor model

Presentation
Observable market risk factors are included in a model through a discretionary choice. Again, the mis-specification risks are non-negligible.

Empirical results
To avoid correlation between the returns of the different asset class factors, Agarwal and Naik (2000a) select variables through a stepwise regression. As explained by the authors, “the stepwise method involves entering the independent variables into the discriminant function one at a time, based on their discriminating power. The single best variable is chosen first; the initial variable is then paired with each of the other independent variables, one at a time, and a second variable is chosen, and so on.”

The asset class factors (i.e. the market factors) are selected from among the following indices: S&P 500 Composite, MSCI World excluding US, MSCI Emerging Markets, Salomon Brothers Government and Corporate Bond, Salomon Brothers World Government Bond Index, Lehman High Yield Composite, Federal Reserve Bank Trade-Weighted Dollar and the UK Market Price for Gold.

From January 1994 to September 1998, the model is applied to 4 directional strategies (Macro, Long, Hedge Long Bias and Short) and 6 non-directional strategies (Fixed Income Arbitrage, Event Driven, Equity Hedge, Restructuring, Event Arbitrage and Capital Structure Arbitrage). Confirming that non-directional strategies are less correlated with the market, they exhibit lower R² (from 0.38 to 0.73) than the directional strategies (from 0.49 to 0.83).

For 8 strategies, alpha is significant at the 5% level, and for 2 strategies it is significant at the 10% level. It is comprised between 0.53 (Macro strategy) and 1.25 (Short strategy).

Agarwal, Fung, Loon and Naik (2004) create 3 ABS factors by simulating the returns generated by three sub-strategies: positive carry (henceforth PCASi), volatility arbitrage (PCASv), and credit arbitrage (PCASc). Convertible Arbitrage returns are successively regressed on PCASi, PCASi plus PCASv, and PCASi plus PCASv plus PCASc. The whole process is repeated for each of the following indexes: TASS, HFR, MSCI, and CISDM. When PCASi is the sole factor, the in-sample adjusted R² ranges only from 1.7% to 5.9% according to the index provider. The inclusion
of the PCASv ABS factor improves the in-sample adjusted $R^2$ to a range of 6.3\% to 23.5\%, according to the index provider. The inclusion of the PCASc ABS factor generates an increase of the in-sample adjusted $R^2$ in three of the four indexes (the CISDM index shows a decrease), ranging from 8.7\% to 22.7\%.

The heterogeneity of the strategies used in the hedge fund industry raises the question of modelling the returns obtained by a diversified portfolio of hedge funds. The most relevant method seems to be modelling the returns strategy by strategy. Nevertheless, a combination of the factors issued from models specific to their respective strategy can lead to the modelling of the returns of a diversified portfolio.

Fung and Hsieh (2004) examine whether combining 7 ABS factors (from three distinct strategies) in a unique model allows a significant part of the returns of a diversified hedge fund portfolio to be explained.

The 7 factors are as follows: the change in credit spreads, the change in ten-year Treasury yields (Fixed-Income strategy), a stock market factor, the spread between large cap and small cap stocks (Equity Long/Short), a lookback straddle on currencies, a lookback straddle on commodities and a lookback straddle on bonds (Trend Following/CTA).

The HFR Fund-of-Funds Index returns are regressed onto the 7 ABS factors, from 1994 to 2002. The in-sample adjusted $R^2$ is 55\%. The coefficients of the equity long/short and fixed-income ABS factors are significant. This is not the case for the trend-following ABS factor.

Ammann and Moerth (2005) implement an explicit macro-factor model. They use a sample of 2,317 funds (TASS database), including both dead and living funds, from January 1994 to June 2003. Factors are the MSCI World, the NASDAQ Composite Index, the Russell 2000 Index, the Wilshire Micro Cap Index, the Lehman Aggregate Bond Index, the Goldman Sachs Commodity Index, the IPE Brent Crude Oil Index, the London Gold Bullion USD Index and the Chicago Board Options Exchange SPX Volatility Index. The dependent variable is the sample excess return on the 90-day T-bill rate.

Using a linear regression, only 3 factors are significant at the 5\% level: the Goldman Sachs Commodity Index, the Lehman Aggregate Bond Index, and the Wilshire Micro Cap Index. The adjusted $R^2$ is 0.4871.

A stepwise regression is also conducted. Again only 3 factors are significant: the Lehman Aggregate Bond Index and the Wilshire Micro Cap Index at the 1\% level, and the Goldman Sachs Commodity Index at the 5\% level. The alpha is 0.81\%, but it is not significant. The adjusted $R^2$ is 0.4897.

3.1.3. Implicit factor model

**Presentation**

Implicit factors are obtained through Principal Component Analysis. This is a purely statistical approach. The aim is to explain the return series of observed variables through a smaller group of non-observed implicit variables. From a mathematical point of view, each implicit factor is defined as a linear combination of the primary variables. The implicit factors are extracted from the time-series of returns.

The advantage is that it solves the problem of the choice of factors, because they are drawn from the series of returns. This avoids the risk of under-specifying the model (omitting true factors) or over-specifying the model (including spurious factors).

The drawback relates to the economic significance of the implicit variables obtained. Apart from the first variables, which are strongly correlated with the market index, the implicit factors are not easy to interpret.

**Empirical results**

Amenc and Martellini (2003) use a Principal Component Analysis (PCA) to extract a set of implicit factors. While a CAPM gives an annualised
mean alpha of 5.83% from a CISDM database, the implicit factor model gives an annualised mean alpha of -1.04%. Such results highlight the impact of the model specification on results.

3.2. Adaptation of the models to the non-linearity of hedge fund returns

Traditional multi-factor models appear to be inappropriate in the context of the wide range of dynamic trading strategies applied by hedge funds. Fung and Hsieh (2000) show that hedge fund returns exhibit non-linear option-like exposures to standard asset classes. Therefore, new models have been constructed in two different ways: the first consists of introducing non-linear regressors into a linear model, the second approach is followed in most recent studies which investigate the concept of “conditional performance”.

3.2.1. Linear models including non-linear regressors

To reproduce the dynamic trading strategies of hedge funds, two different non-linear variables are used in the literature: option portfolios and hedge fund indices.

3.2.1.1. Option portfolios as non-linear regressors

Presentation
This approach consists in including option-based strategies in the multi-factor model (in addition to buy-and-hold strategies) to capture hedge fund option-like payouts.

Empirical results
Agarwal and Naik (2000b) consider 3 option trading strategies: first, an at-the-money option trading strategy, where the present value of the exercise price equals the current index value; second, an out-of-the-money option trading strategy, where the exercise price is half a standard deviation away from that of the at-the-money option; and third, a deep-out-of-the-money option trading strategy, where the exercise price is one standard deviation away from that of the at-the-money option on the Russell 3000 index.

10 strategies are analysed. In the case of non-directional strategies, the average proportion of observed $R^2$ attributable to option-based strategies is 71%. In the case of directional strategies, the average proportion of observed $R^2$ due to option-based strategies is 51%. These results tend to prove the importance of including option-based factors in performance evaluation models for hedge funds.

3.2.1.2. Hedge fund indices as non-linear variables

Presentation
Lhabitant (2001) implements hedge fund indices as non-linear variables in a multi-factor model. Each index corresponds to a hedge fund style. The model is as follows:

$$R_t = \alpha + \sum_{i} \beta_i \cdot I_{i,t} + \epsilon_t$$

where $I_{i,t}$ is a hedge fund index, $\beta_i$ is the exposure to a hedge fund index $I$ and $\alpha$ is the intercept term interpreted as the portion of returns that is unexplained by the factor exposures.

All betas have to be positive, but their sum can be other than 1.

This model can be used to identify the strategy followed by a fund or a portfolio of funds. When the exposure to a given hedge fund index dominates the exposure to other indices, it indicates the main strategy actually pursued. This strategy can differ from the officially reported strategy. When several strong exposures are identified, it indicates a diversified style approach.
Empirical results
Lhabitant applies the model to 2,934 investment vehicles, including both hedge funds and CTAs. Data are provided by Managed Account Reports, Hedge Fund Research, TASS, and Evaluation Associates Capital Management. The specification of the model is based on 9 indices. The strategies are Convertible Arbitrage, Short Bias, Event Driven, Global Macro, Long Short Equity, Emerging Markets, Fixed Income Arbitrage, Market Neutral, and Managed Futures. The hedge fund indices, which are provided by CSFB/Tremont, have the advantage of being asset-weighted, corresponding to a momentum strategy. The average $R^2$ is 0.56. It ranges from 0.51 (for the funds identified as Market Neutral funds) to 0.75 (Emerging Markets funds).

3.2.2. Conditional approaches

3.2.2.1. Methods
In opposing static models, some authors consider the following issue to be important: static asset pricing models imply that risk and performance are constant over time. Due to investment decisions based on public information and dynamic trading strategies, in the case of hedge funds, static models present the risk of being mis-specified. If the risk profile is modified over the calculation period, it can have a strong impact on abnormal performance. This assumption goes against several studies which use multi-factor asset pricing models, where the risk exposure remains constant. For this reason, Kat and Miffre (2002), Kazemi and Schneeweis (2003) and Cerrahoglu, Daglioglu and Gupta (2003) attempt to improve the statistical significance of the performance evaluation by constructing a time-varying expected return asset pricing model.

Kat and Miffre (2002) estimate 3 conditional models. The set of risk factors is composed of a market factor, two microeconomic factors (size and book-to-market value) and five macroeconomic factors (exchange rate risk, term structure of interest rates, international risk of default on short maturity securities, inflation risk and industrial risk). The 3 conditional models are a market model, the Fama-French three-factor model comprising the two microeconomic factors, and an explicit macro-factor model that considers the market factor and the five macroeconomic factors. In Jensen's traditional approach and on the basis of Ferson and Schadt (1996), the authors replace $\alpha$ and $\beta$ with $(\alpha_0|z_{t-1})$ where $(\beta_1|z_{t-1})$ denotes a parameter that is conditional upon $(z_{t-1})$ in order to obtain:

$$p_t = \alpha_0 + \alpha_1 z_{t-1} + \beta_0 f_t + \beta_1 f_t z_{t-1} + \varepsilon_t$$

where $p_t$ is the excess portfolio return on a constant and $\alpha_0$ is the conditional counterpart of the Jensen measure of abnormal performance. The regressors pick up the variations through time in the performance and risk measures that are related to changing economic conditions.

Kazemi and Schneeweis (2003) argue that the distributions of hedge fund returns are neither normal nor identical through time. They propose a conditional model of performance, based mainly on previous work by Chen and Knez (1996) and Cochrane (2001): the Stochastic Discount Factor model. The SDF model has a principal advantage in that it takes the time-varying nature of the relationship between hedge fund returns and primary asset classes into account. The major assumption behind the SDF approach is the absence of arbitrage in financial markets. Under this condition, the SDF is a positive random variable which adjusts future payoffs for the passage of time and uncertainty.

Cerrahoglu, Daglioglu and Gupta (2003) consider that in the context of the dynamic trading strategies pursued by hedge fund managers, the introduction of time variation to Jensen's model (1968) is a potential source of a more accurate estimation of the parameters. A linear
relationship between beta and a set of mean zero information variables available at time ‘t-1’ enables a conditional performance evaluation model, as against static models, to be obtained. In their study, stochastic discount factors (SDF) are used as a linear function of the excess market return. On the basis of a multifactor model which accounts for time-varying betas and the non-normality of returns, two methods are applied: the Generalized Method of Moments (GMM) and the Ordinary Least Squares (OLS) method.

3.2.2.2. Results
In Kat and Miffre (2002), results highlight the fact that both abnormal performance and risk measures are time-dependent. The hypothesis of constant regression is rejected for about 79% of the funds for the market model, and for all the funds for the two multifactor models. When considering the conditional six-factor model, the best predictor of abnormal hedge fund performance is its own return (39% of the time). Next come the default spread (22.1%), the dividend yield (16.9%), the term structure (13%) and the Treasury bill (11.7%). The impact of these variables is different in periods of expansion and recession. The default spread, dividend yield and term structure have a positive (negative) impact in a down (up) market. The Treasury bill has a negative (positive) impact in an up (down) market. Moreover, abnormal performance is counter-cyclical (in a recessionary period, it is above average 75% of the time). This reveals that hedge funds are attractive in down markets.

The statistical significance of abnormal performance is raised by conditional models, by consulting the average t-ratio of the conditional models versus static models (0.72 on average, 0.98 for the six-factor model). This is confirmed by the following result: static models capture on average between 13.2% and 20% of the variation in hedge fund returns, versus a range of between 24.3% and 35.1% for conditional models.

In economic terms, the significance of the improvement in performance evaluation through the conditional models is tested via the annually compounded average measure of abnormal performance. On average, abnormal performance is increased by 1.25% through the conditional models.

Concerning hedge fund indices, Kazemi and Schneeweis (2003) find that the risk-adjusted returns obtained through Jensen’s model are significantly positive and the estimated alphas are significant. Measures based on the Stochastic Discount Factor approach show similar results for all strategies.

The other results relate to hedge fund managers. Using a multi-factor model and an SDF approach, the estimated alphas are close to the alphas given by the single factor model, except for three strategies (large hedged equity, small hedged equity and large convertible arbitrage funds), whose alphas are more significant using the SDF model.

Kazemi and Schneeweis therefore conclude that the set of primitive assets and conditioning variables that they use are not capable of capturing the type of trading strategies followed by most hedge fund strategies.

Cerrahoglu, Daglioglu and Gupta (2003) highlight the fact that the estimations provided by both the single-factor model and the multi-factor model are the same. According to them, as well as Kazemi and Schneeweis (2003), this first indicates that the variables used in this study have weak explanatory power for the type of trading strategies used. Secondly, estimated alphas are not explained by static or dynamic strategies, but by managers’ skills. It would appear that a conditional model does not improve the estimations of the excess return, relative to a static model.

3.3. Higher-moment-adjusted CAPM

Presentation
The standard CAPM can be modified in order
to consider the impact of higher moments on excess returns. This is a higher-moment-adjusted CAPM.

Favre and Ranaldo (2003) examine the relevance of the introduction of co-skewness and co-kurtosis. They propose extensive versions of the CAPM. Firstly, they employ the Market Model to study the standard CAPM. This is expressed as follows:

\[ R_{i,t} - R_{f,t} = \alpha_1 + \alpha_2 R_{m,t} - R_{f,t} + \epsilon_t \]

where \( R_{i,t} \) is the return of the hedge fund (or the index), \( R_{f,t} \) is the return of the risk-free rate and \( R_{m,t} \) is the return of the market portfolio. \( \alpha_1 \) denotes the excess return (in other words the alpha), and \( \alpha_2 \) expresses the covariance of hedge fund returns with the market portfolio.

After that, in a Quadratic Model corresponding to a three-moment CAPM, the authors add a relation to the third moment through the following term:

\[ \alpha_3 (R_{m,t} - E(R_m))^2 \]

where \( \alpha_3 \) is a proxy of co-skewness, and \( E(R_m) \) is the expected return of the market portfolio.

Finally, in a Cubic Model corresponding to a four-moment CAPM, they add a relation to the fourth moment through the following term:

\[ \alpha_4 (R_{m,t} - E(R_m))^3 \]

where \( \alpha_4 \) is a proxy for co-kurtosis.

**Empirical results**

Favre and Ranaldo (2003) use the monthly returns of 16 hedge fund indices provided by HFR, from January 1990 to August 2002. They find that the Quadratic Model enables the adjusted \( R^2 \) (in comparison to the Market Model) in most hedge fund styles to be increased, in particular for Convertible Arbitrage (from 0.154 to 0.198), Distressed Securities (0.194 – 0.339), Event Driven (0.469 – 0.579), Emerging Markets (0.368 – 0.406), Fund of Funds (0.231 – 0.266), Market Timing (0.254 – 0.360), Merger Arbitrage (0.468 – 0.509), Relative Value Arbitrage (0.180 – 0.318) and Weighted Composite (0.585 – 0.624) indices. For these indices, \( \alpha_3 \) exhibits a high significance level.

Focusing on the Cubic Model, the adjusted \( R^2 \) is increased only for Convertible Arbitrage (from 0.198 to 0.211), Emerging Markets (0.406 – 0.429), Market Timing (0.360 – 0.374) and Merger Arbitrage (0.509 – 0.541) indices. \( \alpha_4 \) is only significant for these four strategies. However, the authors stress the fact that the high significance level of the co-kurtosis comes at the cost of a decrease in the significance level of the co-skewness coefficient. This seems to indicate co-linearities between the co-moments, thereby limiting the explanatory power of the model.

Fung, Xu and Yau (2004) implement a higher-moment-adjusted CAPM on equity-based style hedge funds on the basis of a database provided by CISDM from 1994 to 2000. They obtain similar excess returns using the standard CAPM and the extended CAPM. Such results suggest that in the case of equity-based style hedge funds, the inclusion of higher moments in the performance measure is not mandatory. It can be explained by the fact that the non-normality of the return distribution is not pronounced enough to obtain results that are significantly different.
The attempt to build portfolios mimicking the statistical properties of hedge fund strategies is a response to the criticism made of hedge funds by investors regarding their lack of liquidity, lack of transparency, lack of capacity, excessive management fees and style drift. The large amount of studies striving to replicate hedge fund performance does not therefore come as a big surprise.

Two main replication approaches can be distinguished: passive replication and dynamic replication. While a passive replication does not require the intervention of a manager after the portfolio construction (i.e. the portfolio is passively held), a dynamic replication is based on the re-estimation of the different allocations at a given frequency.

4.1. Passive Replication

Passive replication typically involves a basic constrained regression. Technically speaking, it involves looking for the portfolio — made up of long and/or short positions in a set of risk factors — that minimizes the volatility of the error term (i.e., the difference with the fund’s actual returns or “tracking error”). The mimicking portfolio brought about by the in-sample analysis is then passively held in the out-of-sample period. The wide range of risks to which hedge funds are exposed (see Géhin and Vaissié (2006)) leads to a consideration of multi-factor approaches only. This raises the question of the model parameters to be specified. Because risk factors are included through a discretionary choice, the risks of misspecification are non-negligible.

Studying 8 strategies, Agarwal and Naik (2004) use a multi-factor model where the risk factors are “buy-and-hold risk factors” and “option-based risk factors”. Buy-and-hold risk factors are equities (4 indexes), bonds (3 indexes), currencies (1 index) and commodities (1 index). The Fama-French “Size” and “Book-to-Market” factors, the Carhart “Momentum” factor and a credit risk factor are added. Option-based risk factors are ATM and OTM European call and put options on the S&P500. First, a stepwise regression is conducted to identify the significant factors, from 1990 to 2000, for 8 HFR indexes. Second, the authors examine whether “the replicating portfolios based on these factor loadings should do a good job of mimicking the out-of-sample performance of hedge funds”. A replicating portfolio is constructed, and the accuracy of the return replication is tested through a standard t-test and the Wilcoxon sign-test. Third, a figure plots the returns displayed by the replicating portfolio and its respective HFR index during the out-of-sample period for 4 strategies. In-sample adjusted $R^2$ ranges from 40.5% to 91.63%. The Standard t-test and Wilcoxon sign-test indicate that the difference in return (mean and median) between HFR indexes and their respective replicating portfolio is statistically insignificant during the out-of-sample period of 2000 to 2001. According to the authors, the figure plotting the returns displayed by replicating portfolios and their respective HFR indexes during the out-of-sample period “shows that the portfolios based on significant risk exposures estimated through our model closely track the hedge fund returns during the out-of-sample period.”

In an analysis of 8 strategies, Jaeger and Wagner (2005) implement a multi-linear asset class factor model for the in-sample period of 1994 to 2004. In a further step, the “Replicating Factor Strategy” (henceforth RFS) returns are calculated. “The RFS returns in a given month were calculated using factors obtained by a regression over data for the previous five years ending with the previous month”, in order to “avoid the problem of data-mining and in-sample over-fitting”. In-sample adjusted $R^2$ ranges from 35.3% to 88.5%. During the out-of-sample period, the difference between RFS cumulated returns and their corresponding HFR non-investable index ranges from -24.7 points to 11.8 points, while the
difference between RFS cumulated returns and their corresponding HFR investable index ranges from -5 points to 12.9 points.

Summary
Concerning passive replication, models do not generate acceptable in-sample $R^2$ to look at out-of-sample replication. Moreover, studies that provide out-of-sample $R^2$ exhibit approximate replications (see Table 1).

4.2. Dynamic Replication

As one might have expected, capturing the behavior of highly dynamic strategies with passive exposure to a set of factors is not feasible. Passive replication implies that we have to assume that return-generating processes are stationary. However, as mentioned earlier, this rarely proves to be the case in practice; hence the poor out-of-sample results obtained in most cases. Some authors therefore give up on passive replication and turn to dynamic replication.

4.2.1. Optimization-based replication

Georgiev, Schneeweis, and Spurgin (2001) construct a CTA benchmark. The first step involves constructing indexes for four futures market segments commonly traded by CTAs: currencies (3 contracts), interest rates (3 contracts), physical commodities (5 contracts) and equities (3 contracts).

<table>
<thead>
<tr>
<th>Authors</th>
<th>Strategies</th>
<th>(In-sample period) In-sample results</th>
<th>(Out-of-sample period) Out-of-sample results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Equity Long/Short</td>
<td>$R^2$ 72.50%</td>
<td>approximate replication of index returns</td>
</tr>
<tr>
<td></td>
<td>Equity Non Hedge</td>
<td>$R^2$ 91.63%</td>
<td>(standard $t$-test &amp; Wilcoxon sign-test)</td>
</tr>
<tr>
<td></td>
<td>Relative Value</td>
<td>$R^2$ 73.40%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Restructuring</td>
<td>$R^2$ 52.20%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Risk Arbitrage</td>
<td>$R^2$ 65.60%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Short Selling</td>
<td>$R^2$ 44%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$R^2$ 82%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Distressed Securities</td>
<td>$R^2$ 68.40%</td>
<td>RFS+7.6%/HFRI+2.4%</td>
</tr>
<tr>
<td></td>
<td>Equity Long/Short</td>
<td>$R^2$ 72.50%</td>
<td>RFS+20.1%/HFRI+44.8%</td>
</tr>
<tr>
<td></td>
<td>Equity Market Neutral</td>
<td>$R^2$ 35.30%</td>
<td>RFS+32.2%/HFRI+32.1%</td>
</tr>
<tr>
<td></td>
<td>Event Driven</td>
<td>$R^2$ 79.30%</td>
<td>RFS+6.2%/HFRI+10.9%</td>
</tr>
<tr>
<td></td>
<td>Global Macro</td>
<td>$R^2$ 49.70%</td>
<td>RFS+29.8%/HFRI+40%</td>
</tr>
<tr>
<td></td>
<td>Merger Arbitrage</td>
<td>$R^2$ 52.90%</td>
<td>RFS+16.7%/HFRI+24.6%</td>
</tr>
<tr>
<td></td>
<td>Short Selling</td>
<td>$R^2$ 81.20%</td>
<td>RFS+13%/HFRI+15.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RFS-28.2%/HFRI-23%</td>
</tr>
</tbody>
</table>

Table 1. In-sample and out-of-sample results of passive replication

4.2. Dynamic Replication

The position on a given contract is obtained by combining the positions given through 15-, 27- and 55-day momentum trading. The second step consists in replicating a particular CTA or CTA style via a linear combination of the four market indexes. This is the optimization stage. The model is re-estimated with indexes that are statistically significant at the 10% level. In the out-of-sample analysis, “Weights for style benchmarks are re-estimated for each year using data from the previous five years.”
Indexes are constructed for the period of January 1988 to December 2000, while the out-of-sample period is from 1993 to 2000. Out-of-sample R² figures are as follows: 54% for the Dollar Weight sub-category (replication of the MAR Dollar Weight index), 60% for Equal Weight, 62% for Currency, 4% for Discretionary, 50% for Diversified, 62% for Financial, 58% for Systematic, and 62% for Trendfollowing. These figures are lower than in-sample R², which range from 17% to 72%. The authors observe that “While the passive style benchmarks explain a considerable proportion of the variation in CTA returns, the model is less successful in tracking the level of CTA returns.” In fact, for seven of the eight sub-categories, in the out-of-sample period the replicating indexes underperform the MAR indexes.

4.2.2. Copula-based replication
Kat and Palaro (2005) consider that the failure of factor modelling in replicating hedge funds, illustrated by low out-of-sample R², justifies the need for a new replication method that does not require the identification of significant risk factors. They propose a replication which is not only a replication of the hedge fund return distribution, but also a replication of the relationship between the fund and the investor's existing portfolio.

Their method is as follows:
• First, data are collected on the fund to be replicated, the investor's portfolio, and the reserve asset.
• Second, the joint distribution of the fund return and the investor's portfolio return is calculated, referred to as the “desired distribution”. To form the joint distributions (or bivariate distributions), two univariate distributions and a copula are combined. In their work, Kat and Palaro combine 3 distribution types with 6 copula types. It gives 54 (3 possible types of distribution for the 1st distribution * 3 types of distribution for the 2nd distribution * 6 copula types) possible joint distributions. The relevant joint distribution is selected via the Akaike Information Criterion.
• Third, the joint distribution of the investor's portfolio return and the reserve asset return is calculated, referred to as the “building block distribution”.
• Fourth, the “desired payoff function” is determined. It is the cheapest payoff function that allows the “building block distribution” to be turned into the “desired distribution”.
• Fifth, the “desired payoff function” is priced. In the empirical application presented by the authors, the desired payoff function is built on the basis of the last 24 monthly returns. It allows the number of units of the investor's portfolio and reserve assets to be determined.
• Sixth, during the month (1st out-of-sample month), the required positions are daily adjusted to the market values.
• Seventh, at the beginning of the next month, the dataset is extended to the last month, and the desired payoff function is recalculated. Steps 6 and 7 are repeated each month, until the predetermined horizon is reached.

Assuming the investor's portfolio consists of 50% US equities and 50% US Treasury bonds, and the reserve assets consist of Eurodollar futures, the replication method is applied to the returns of 3 funds: Leveraged Capital Holdings N.V. (a fund of hedge funds), Calamos Multi-Strategy Fund L.P. (a convertible arbitrage fund), and Rocker Partners L.P. (a short selling fund). Out-of-sample periods are respectively from 1987 to 2004, 1991 to 2004, and 1987 to 2001.

Over the whole period, the distribution of the fund returns and the relationship with the investor's portfolio are replicated in a very similar shape. Only extreme events are not replicated. However, this replication is not a month-to-month replication, but a replication from one date to a predefined horizon. The plot of the fund returns versus the replicated returns “shows a random scatter”, in other words “although the fund returns versus the replicated returns have statistical properties which are very similar...
4. Performance replication

[...], they come to the investor in a completely different order.” The authors recognize that the superiority of their method comes at the cost of a month-to-month checking: “giving up the requirement that returns need to be similar on a month-to-month basis [...] allows us to do so much better than the standard factor model approach.”

Moreover, each month the portfolios are adjusted on a daily basis, driven by the daily changes in the underlying index values. Consequently it induces high frequency trading on the reserve assets and on the assets contained in the investor’s portfolio. It is illustrated by the evolution of the strategy controls, in other words the units of the investor’s portfolio and reserve assets. Additionally, each month it is necessary to collect the data related to the replicated fund, in order “to determine the desired payoff function from the available 24 month returns, calculate the accompanying strategy controls and set up the required positions.” In summary, this method requires heavy data collection and complex calculations.
Top-ranking funds are very attractive. However, it’s important to ask one fundamental question: do winners repeat? This question relates to a primumordial aspect of the performance measurement process, namely performance persistence.

While a measure of performance over a given period, whatever the performance indicator, provides a static picture of performance (even if an extended period is used), its persistence reveals a more dynamic picture of a track record.

The definition of an adequate method is an obvious condition for using performance persistence as a viable quantitative feature of the selection process. Identification of potentially persistent ‘hot hands’ managers cannot be conducted through a tiresome comparison of their performance rankings in successive periods – as is done in less meticulous studies – but requires statistically significant results.

The non-relevance of selecting funds on the basis of a given period was illustrated by Géhin (2005). 182 funds of hedge funds from the Alternative Asset Center (AAC) database are ranked into quartiles, each year from January 2000 to December 2004. A Sortino ratio is used to measure the risk-adjusted performance. Table 2 exhibits that an investor has, at best, less than a one-in-two chance, and at worst, less than a one-in-four chance of selecting a fund that will be ranked in quartile one the following year by using the current quartile one ranking as a selection criterion.

It is important to stress that two concepts of performance persistence co-exist. The first approach consists of examining what can be called relative persistence, i.e. the persistence of a ranking between winners and losers. The second approach involves examining pure persistence, i.e. the persistence of the performance of a fund without considering other funds at the same time. These approaches offer two independent visions of persistence, and the results of one do not necessarily match the results of the other.

5. Performance persistence

5.1. First approach: persistence of relative returns

5.1.1. Test methods

The persistence of relative returns can be tested in two ways. The first is by using a two-period framework, and the second a multi-period framework.

5.1.1.1. Two-period framework

This is the most common method employed. Many non-parametric methods are available. Non-parametric methods are based on the construction of a two-way winner-and loser contingency table. Winners are funds whose returns are higher than the median return. Losers are funds whose alpha is weaker than the median return. Funds which are winners over two consecutive periods are denoted WW. Funds which are losers over two consecutive periods are denoted LL.

5.1.1.1.1. Cross Product Ratio test

The CPR numerator corresponds to the funds which persist, and the denominator corresponds to the funds which do not persist:

$$\text{CPR} = \frac{(WW \times LL)}{(WL \times LW)}$$

Under the null hypothesis of no persistence, the ratio is equal to 1. This implies that each of the four categories WW, LL, WL and LW represent 25% of all the funds.

Table 2. Distribution by quartile of the 45 funds of hedge funds ranked in quartile 1 in the previous year, from 2001 to 2004 (funds ranked by their Sortino ratio)
The statistical significance of CPR is tested via the calculation of the Z-statistic, corresponding to the ratio of the natural logarithm of the CPR to the standard error of the natural logarithm of CPR, expressed as follows:

\[
Z - \text{statistic} = \frac{\ln(\text{CPR})}{\alpha_{\ln(\text{CPR})}}
\]

where \(\alpha_{\ln(\text{CPR})}\) is the standard error of the natural logarithm of CPR, equal to:

\[
\alpha_{\ln(\text{CPR})} = \frac{1}{\sqrt{\frac{1}{\text{WW}} + \frac{1}{\text{WL}} + \frac{1}{\text{WL}} + \frac{1}{\text{LL}}}}
\]

For example, a Z-statistic greater than 1.96 indicates significant persistence at a 5% confidence level.


### 5.1.1.1.2. Chi-square test

The chi-square test is carried out by comparing the distribution of the observed frequencies for the four categories WW, LL, WL and LW with the expected frequencies of the distribution. The chi-square is equal to:

\[
\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}
\]

where \(O_i\) is the observed number of funds in each case of the contingency table, and \(E_i\) is the expected number of funds in each case. The degree of freedom is equal to (Line-1)*(Column-1). Carpenter and Lynch (1999) consider that tests based on the chi-square are more robust than the CPR in the presence of survivorship bias.


### 5.1.1.1.3. Spearman rank correlation test

Spearman (1904) proposes a non-parametric (distribution-free) rank as a measure of the strength of the associations between two variables.

Assume a set of funds \((1, 2, 3, \ldots, n)\), which have been ranked by two regimes \((x, y)\). In the case of performance persistence, the two regimes are the two different periods. Let \(x(i)\) and \(y(i)\) be the rank (rank one is the highest and rank \(n\) is the lowest) of fund \(i\) in the two regimes respectively and define \(d_i = x(i) - y(i)\) as the distance between these rankings.

Spearman’s rank correlation is obtained with the following formula:

\[
r_s = 1 - \frac{6 \sum d_i^2}{n^3 - n}
\]

The result will always range from 1 (a perfect positive correlation, i.e. a perfect positive persistence of performance) to minus 1 (a perfect negative correlation, i.e. a perfect negative persistence of performance). A coefficient close to 0 indicates an absence of performance persistence over two periods.

The Spearman rank correlation test is used by Park and Staum (1998).

### 5.1.1.2. Multi-period framework, Kolmogorov-Smirnov test

The Kolmogorov-Smirnov test (henceforth K-S test) also tests the persistence of relative returns, but it is performed in a multi-period framework. It checks whether the distributions of winning funds and losing funds are statistically different from the theoretical distribution. Observed frequencies of wins and losses are recorded. This frequency distribution is compared with that generated from a normal distribution and the maximum difference in cumulative densities between the observed and the normal distribution is used to construct the K-S statistic.

On the one hand, it is more complex to apply than two-period methods. On the other hand, an advantage of the K-S test is the reduction of the likelihood of observing persistent ‘hot hands’
due to the chance factor. Since it discriminates between the chance and skill factors, it is the most powerful method for testing relative persistence.

The K-S test is used by Agarwal and Naik (2000b) and Koh, Koh and Teo (2003).

5.1.2. Overview of performance persistence studies

Table 3 shows the heterogeneity of the studies on hedge fund performance persistence. The databases, periods, performance measures, test methods, and persistence horizons differ between the studies.

<table>
<thead>
<tr>
<th>Authors (year of the first version)</th>
<th>Database</th>
<th>Period</th>
<th>Performance measure</th>
<th>Test methods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Relative persistence</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Two period</td>
</tr>
<tr>
<td>Park and Staum (1998)</td>
<td>TASS</td>
<td>1986-1997</td>
<td>appraisal ratio(^7) (alpha divided by standard deviation)</td>
<td>• Chi-square</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>specific appraisal ratio(^9)</td>
<td>CPR</td>
</tr>
<tr>
<td>Agarwal and Naik (2000)</td>
<td>HFR</td>
<td>1982-1998</td>
<td>specific alpha(^8)</td>
<td>CPR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>specific appraisal ratio(^9)</td>
<td>CPR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sharpe ratio</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>post-fee returns</td>
<td>CPR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sharpe ratio</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sortino ratio</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Overview of the main studies relating to hedge fund performance persistence, in chronological order

Source: W. Géhin (2006)

With regard to the database used, Agarwal and Naik (2000c) for example cover US hedge funds, while Koh, Koh and Teo (2003) focus on Asian hedge funds.

---

7 - The appraisal ratio is equal to the alpha divided by the standard deviation.
8 - This specific alpha is equal to the return of a hedge fund using a particular strategy minus the average return for all hedge funds following the same strategy.
9 - This specific appraisal ratio is equal to the previously cited alpha divided by the residual standard deviation resulting from a regression of the hedge fund return onto the average return of all the hedge funds following that strategy.
5.1.3. Results at short-term horizons
At 3-month and 6-month horizons, Agarwal and Naik (2000c) find significant persistence, using CPR and chi-square tests. Conducting the K-S test, the level of persistence is dramatically lower, with only one case of persistence in losers and none in winners at a 6-month horizon.

From 1-month to 9-month horizons, Koh, Koh and Teo (2003), find significant persistence, using CPR and chi-square tests. Conducting the K-S test, as the horizon lengthens to 6 months and beyond, the persistence weakens.

5.1.4. Results at long-term horizons
At a 1-year horizon, Park and Staum (1998) state that persistence strength varies substantially from year to year, using Spearman rank correlation and the chi-square tests.

At a 1-year horizon, Agarwal and Naik (2000c) find that the extent of the return persistence diminishes, using the chi-square and the CPR test. According to the K-S test, there is an absence of persistence.

At 1-year and 2-year horizons, Caglayan and Edwards (2001) find both winner and loser persistence for funds of funds, global macro, market neutral funds, and at the whole sample level, using the CPR test.

At a 3-year horizon, Kat and Menexe (2003) find no evidence of persistence, using the CPR test.


At a 1-year horizon, Koh, Koh and Teo (2003), find no evidence of persistence, using the CPR, the chi-square and the K-S tests.

At a 3-year horizon, De Souza and Gokcan (2004) find no evidence of persistence, using the CPR test.

5.1.5. Comments on relative persistence results
Even though some studies present conflicting conclusions, the main characteristics of hedge fund performance persistence can be highlighted.

Firstly, short-term persistence, for horizons of up to six months, is reported by two-period tests. Using a multi-period test (the K-S test), the persistence levels are less significant. A multi-period framework allows for better discrimination between persistence due to chance and persistence due to manager skill.

Secondly, return persistence weakens as one lengthens the measurement horizons. The annual horizon is the most undecided horizon. Park and Staum (1998), Agarwal and Naik (2000c), Kat and Menexe (2003), Koh, Koh and Teo (2003) and De Souza and Gokcan (2004) find no evidence of persistence, or persistence that depends on the periods. However, Caglayan and Edwards (2001), and Kouwenberg (2003) arrive at significant persistence for some strategies.

The sensitivity of persistence to the test horizon becomes important when the lockup period is considered. If the lockup period is greater than the persistence horizon, the investor is exposed to a potential reversal in the trend of the returns reported by the hedge funds.

5.1.6. Disadvantages of a relative persistence approach
Tests on relative persistence suffer from a problem that comes from a sample approach which generalises the conclusions to apply to all the funds. If one concludes that persistence of relative performance exists, and if one extends the conclusion to all the funds, the fact that the test is conducted on the whole sample can lead to the selection of an initially winning fund that
is not persistent. This is because the persistence of other funds in the sample allows statistical significance to be reached. Conversely, if one concludes that there is an absence of persistence in the initial winning funds, there are perhaps persistent winning funds among those funds.

Moreover, the presence of significant persistence (for both winners and losers) only at horizons that are equal to or less than three months leads to relative performance persistence being considered as a short-term phenomenon. The persistence of relative performance is too short to be used as a viable selection criterion. It is thus preferable to seek funds whose positive performance is the most persistent over long periods. This justifies the introduction of pure persistence through the Hurst exponent.

5.2. A manoeuvrable approach: pure persistence

While relative persistence tests require a whole sample of funds to be analysed, a pure persistence test isolates a fund and focuses on it. Unlike the relative persistence tests, which indicate persistence of the outperformance or underperformance relative to other funds, a pure persistence test allows the funds exhibiting the strongest persistence of their positive returns to be identified fund by fund.

5.2.1. Hurst exponent combined with a D-Statistic

The presence of pure persistence is examined by the calculation of a Hurst exponent, combined with a D-statistic. It is the solution to the following equations:

$$H_{t} = (ct)^H$$

or

$$\ln R_S = \ln(c) + H \ln(t)$$

where $RS_i$ is the range of cumulative deviations from the mean divided by the standard deviation and $H$ is the Hurst exponent.

The Hurst exponent reveals the nature of the persistence: an exponent close to 0 suggests reverse persistence (positive returns succeed negative returns and vice-versa), an exponent close to 0.5 implies a random walk of returns, and an exponent close to 1 indicates positive persistence.

When the Hurst exponent is greater than 0.5, a D-statistic is calculated in order to determine whether positive or negative returns persist. The D-statistic is calculated as follows:

$$D - \text{statistic} = \frac{\text{sum} |\text{negative returns}|}{\text{sum} |\text{all returns}|}$$

Equal to the sum of the absolute value of the negative returns divided by the sum of the absolute value of all returns, and ranging from 0 to 1, a low D-statistic indicates a large share of positive returns.

The combination of a high Hurst exponent and a low D-statistic denotes the presence of pure positive return persistence.

5.2.2. Pure persistence results

On the basis of the Hurst exponent, De Souza and Gokcan (2004) examine the pure persistence of the performance of seven strategies, namely convertible arbitrage, distressed securities, merger arbitrage, fixed income arbitrage, equity market neutral, equity long/short and global macro. The period is from January 1997 to December 2002. The sample contains 314 hedge funds provided by HFR.

The period from January 1997 to December 2002 is divided into two sub-periods of three years. The first period is considered as the in-sample period, and the second period is considered as the out-of-sample period. The funds are divided into
three groups according to the level of the Hurst exponent exhibited by each fund. To determine the weights of each manager in the groups, two methods are used. The simpler one consists of equal weighting. The second method introduces the concept of risk budgeting. Risk budgeting is "an asset allocation technique where [...] capital is allocated to risk buckets with no consideration of associated returns."

The "low Hurst" group contains 105 managers, where exponents range from 0.32 to 0.58. The "medium Hurst" group contains 105 managers, where exponents range from 0.59 to 0.69. The "high Hurst" group contains 104 managers, where exponents range from 0.70 to 0.98. Except for the distressed securities strategy, which does not appear in the low Hurst group, all the strategies are represented in each group. Using an equal-weighting scheme, during the in-sample period returns, standard deviations and Sharpe ratios do not differ significantly among Hurst groups.

During the out-of-sample period, the high Hurst group displays the highest rate of return, the lowest volatility (thus automatically the highest Sharpe ratio), and the highest number of months with consecutive gains. Using the risk budgeting approach, the high Hurst portfolio presents the best statistics too. In other words, persistent managers outperform non-persistent managers during the out-of-sample period.

A D-Statistic is calculated for the managers included in the high Hurst group only in order to filter negative persistence. The 104 managers of the high Hurst group are ranked into three portfolios according to their D-statistic calculated during the in-sample period. The equal-weighting scheme and risk budgeting are successively used to construct the three portfolios. During both periods, using either equally-weighted portfolios or a risk budgeting approach, the low D-statistic portfolio exhibits the lowest standard deviation, the highest Sharpe ratio, and the highest number of months with consecutive gains.

In other words, funds exhibiting the highest Hurst exponent and lowest D-Statistic are more liable to have persistent performance.

De Souza and Gokcan stress the ability of such a method based on the Hurst exponent and the D-statistic "to identify a sample of 35 managers with a "future" return of 9.12% and a volatility of 2.57% (on an equally weighted basis) from a pool of 314 managers with an average return of 6.16% and a volatility of 4.52%.”

Géhin (2005) calculates the Hurst exponent for each of the 182 funds of hedge funds from the AAC database, over a 4-year period from January 2000 to December 2003. The minimum level of the Hurst exponent required to have positive persistence is increased to 0.6 and the D-statistic is considered to be low when it is less than 0.3. Table 4 indicates that over the four-year period, the Hurst exponent ranges from 0.4257 to 0.7253, and 29.67% of the 182 funds exhibit positive persistence of their positive returns. Two equally-weighted portfolios that contain, respectively, the 54 persistent funds and the 128 other funds are constituted.

<table>
<thead>
<tr>
<th></th>
<th>Min Hurst</th>
<th>Max Hurst</th>
<th>Positive Persistence of Positive Returns</th>
<th>Total Number of Funds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H &gt; 0.6 and D &lt; 0.3</td>
<td></td>
<td>Number of Funds</td>
<td>%</td>
</tr>
<tr>
<td>4 year-period</td>
<td>0.4257</td>
<td>0.7253</td>
<td>54</td>
<td>29.67%</td>
</tr>
<tr>
<td>5 year-period</td>
<td>0.4216</td>
<td>0.7328</td>
<td>46</td>
<td>25.27%</td>
</tr>
</tbody>
</table>

Table 4: Positive persistence of performance, according to a Hurst exponent combined to a D-statistic on two periods of 4 and 5 years beginning in January 2001, from a sample of 182 funds
The non risk-adjusted return and the Sortino ratio for the year 2004 are calculated (see Table 5). While the “persistent funds” and the “other funds” portfolios display similar non-risk-adjusted returns (7.63% and 7.60% respectively), the “persistent funds” portfolio exhibits a better Sortino ratio (5.55% versus 3.19%). Since the Sortino ratio penalizes funds with the most pronounced downside deviation, one possible interpretation is that selecting funds on the basis of the combination of the Hurst exponent and the D-statistic allows funds with obviously stable returns to be isolated.

<table>
<thead>
<tr>
<th></th>
<th>Annualized Average of Monthly Returns</th>
<th>Monthly Downside Deviation</th>
<th>Annualized Sortino</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persistent Funds</td>
<td>7.60%</td>
<td>0.311</td>
<td>5.55%</td>
</tr>
<tr>
<td>Other Funds</td>
<td>7.63%</td>
<td>0.543</td>
<td>3.19%</td>
</tr>
</tbody>
</table>

Table 5. Comparison of the risk-adjusted performance displayed by the persistent funds over the 4 previous years and the other funds, from January 2004 to December 2004

5.2.3. Comments on pure persistence results

Even if intuitive funds selected on the basis of pure persistence do not represent all top-ranked funds in the different sub-periods, the persistence of their positive returns gives them attractive risk-adjusted returns. The Hurst exponent appears to be a powerful indicator for analysing performance, in conjunction with appropriate risk-adjusted performance measures.

5.3. Persistence is not predictability

Even if evidence of persistence is found, it does not allow returns to be predicted. Some papers confuse performance persistence and performance predictability. Performance predictability tries to forecast the direction and/or the magnitude of future returns on the basis of past returns, while performance persistence is not a forecast, it is a retrospective observation. In other words, two different goals are pursued, and different techniques are used.

Two leading approaches can be followed to predict returns: a return-level approach and a classification-based approach. The first method involves predicting a future return level, while the second method involves predicting the direction of the returns.

Dash and Kajiji (2004) present the results of 3 different test algorithms: the SPSS radial basis function (henceforth RBF) algorithm, the Statsoft RBF algorithm, and the Kajiji RBF algorithm. Monthly returns of 13 hedge fund indices are provided by CSFB/Tremont. They cover the period from January 1994 to December 2002. A Shapiro-Wilk test shows that returns are not normally distributed. Moreover, except for the Equity Market Neutral index, all the indices have fat-tailed distributions (kurtosis is positive). Skewness is negative in 8 cases. The evidence of abnormal returns raises the question of the adequacy of the algorithmic methods for forecasting returns. First, hedge fund return predictors are identified. From the 21 predictors that are most widely used by alternative neural network models, 8 factors are extracted through a Varimax rotational technique, in order to reduce the co-linearity between the variables. The 8 resulting predictors explain more than 82% of the total variation in the predictor variables. The appropriateness of the models is decomposed.
into two parts: a return-level error measure and a classification error measure. Three return-level error measures are calculated: in-sample, out-of-sample, and over the whole period. Three classification error measures are calculated: a direction measure, a modified direction measure, and a time dependent directional profit (henceforth TDDP). A direction measure tests the number of times a prediction neural network predicts the direction of the predicted return movement correctly. The modified direction measure is presented as a more robust measure. TDDP penalizes the prediction errors more heavily. The lower it is, the more adequate the model is.

The Kajiji RBF algorithm is applied to the 13 hedge fund indices. Focusing on return-level error measures, the mean square error is the highest for the indices that display the highest skewness and kurtosis in absolute terms. This reduces the prediction power of the level of return. Nevertheless, this reduction is small when the Kajiji RBF algorithm is used. Focusing on classification error measures, the direction measure and the modified direction measure give the same results for all indices. The direction measure is equal to 0.988, and the modified direction measure is equal to 1, the best result that can be reached. The low level of the classification errors obtained through the Kajiji RBF algorithm is confirmed by the TDDP, which is comprised between 0.0001 (Equity Market Neutral index) and 0.1418 (Dedicated Short Bias index).

The SPSS RBF algorithm suffers from return-level error measures that are larger than those produced by the Kajiji RBF algorithm. Over the whole period, it is 1.975 times larger when applied to the Dedicated Short Bias index, and 3.425 times larger for the Equity Market Neutral index. Classification errors obtained by the SPSS RBF algorithm are also more pronounced. The modified measure is comprised between -0.0137 (Managed Funds index) and 0.5255 (Distressed index). The TDDP ranges from 0.1597 (Fixed Income Arbitrage index) to 7.8533 (Dedicated Short Bias index).

The Statsoft RBF algorithm gives return-level error measures and classification error measures that are higher than those produced by the Kajiji RBF algorithm. Focusing on the Dedicated Short Bias and Equity Market Neutral indices, a graphical view of predicted returns and stated returns shows that the direction of the returns is reasonably well predicted, but the Statsoft RBF algorithm fails to predict the magnitude of the changes.

According to this study, the Kajiji RBF algorithm is the most appropriate neural network method for predicting hedge fund returns on a monthly frequency. This method reconciles the return-level approach and the classification-based approach by presenting very satisfying results for both approaches.

5. Performance persistence
Conclusion

This survey illustrates the progress achieved in hedge fund performance measurement over roughly the last 10 years. First, major biases have been identified and their impact has been evaluated. This allows a second treatment of the database to obtain non-biased calculations. Second, it is now established that absolute and relative performance measurements cannot be conducted in a framework similar to that employed for mutual funds. Traditional absolute performance measures lead to erroneous fund rankings. New performance measures have been introduced, taking into account the particular qualities of hedge fund return distributions. Traditional models display a low explanatory power when they are applied to hedge funds. More sophisticated models have been elaborated, giving stronger results. Third, performance persistence appears as a useful indicator in a hedge fund quantitative selection process. Again, all methods are not equal. A pure persistence approach allows us to identify the most persistent funds with accuracy.

The issue of hedge fund performance measurement is not frozen. Every year, new absolute performance measures and new models are proposed. We think that refinements are still possible. We think also that on the basis of current knowledge, investors have viable criteria in the context of a quantitative selection process, on condition that they employ the most rigorous methods.
References


References


References


EDHEC is one of the top five business schools in France owing to the high quality of its academic staff (100 permanent lecturers from France and abroad) and its privileged relationship with professionals that the school has been developing since its establishment in 1906. EDHEC Business School has decided to draw on its extensive knowledge of the professional environment and has therefore concentrated its research on themes that satisfy the needs of professionals.

EDHEC pursues an active research policy in the field of finance. Its Risk and Asset Management Research Centre carries out numerous research programs in the areas of asset allocation and risk management in both the traditional and alternative investment universes.

Copyright © 2006 EDHEC