Performance Measurement for Traditional Investment

Literature Survey

January 2007

Véronique Le Sourd
*Senior Research Engineer
at the EDHEC Risk and Asset
Management Research Centre*

EDHEC
BUSINESS SCHOOL
Lille - Nice

Table of contents

Abstract

The number of professionally managed funds in the financial markets is increasing. The mutual fund market is highly developed with a wide range of products proposed. The resulting competition between the different establishments has served to strengthen the need for clear and accurate portfolio performance analysis, for which portfolio return alone is not sufficient. This has led to the search for methods that would provide investors with information that meets their expectations and explains the increasing amount of academic and professional research devoted to performance measurement. The topic of performance analysis is still in expansion, meeting the needs of both investors and portfolio managers.

Performance measurement brings together a whole set of techniques, many of which originate in modern portfolio theory. Beside models issued from portfolio theory, research in the area of performance measurement has also concerned the consideration of real market conditions and has developed techniques to fit cases where the restrictive hypotheses of portfolio theory are not observed.

This article presents the state of the art of performance measurement in the area of traditional investment, from a simple evaluation of portfolio return to the more sophisticated techniques including risk in its various acceptations. It also describes models that take a step away from modern portfolio theory and allow a consideration of cases beyond mean-variance theory. It concludes with a review of performance persistence studies.

About the author

Véronique Le Sourd has a Master's Degree in Applied Mathematics from the Pierre and Marie Curie University in Paris. From 1992 to 1996, she worked as research assistant within the Finance and Economics department of the French business school, HEC, and then joined the research department of Misys Asset Management Systems in Sophia Antipolis. She is currently a senior research engineer at the EDHEC Risk and Asset Management Research Centre.
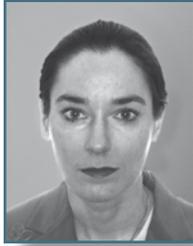
Introduction

The number of professionally managed funds in the financial markets is increasing. The mutual fund market is highly developed with a wide range of products proposed. The resulting competition between the different establishments has served to strengthen the need for clear and accurate portfolio performance analysis. Investors wish to avail of all the information necessary to carry out manager selection over comparable bases. They want to know if managers have succeeded in reaching their objectives, i.e. if their return was sufficiently high to reward the risks taken, how they compare to their peers and, finally, whether the portfolio management results were due to luck or because the manager has real skill that can be identified and repeated in the future. The portfolio return alone does not allow all these questions to be answered. This has led to the search for methods that would provide investors with information that meets their expectations and explains the increasing amount of academic and professional research devoted to performance measurement. The topic of performance analysis is still in expansion, meeting the needs of both investors and portfolio managers.

Performance measurement brings together a whole set of techniques, many of which originate in modern portfolio theory[1]. Performance evaluation is closely linked to risk. Modern portfolio theory has established the quantitative link that exists between portfolio risk and return. The Capital Asset Pricing Model (CAPM) developed by Sharpe (1964) highlighted the notion of rewarding risk and produced the first performance indicators, be they risk-adjusted ratios (Sharpe ratio, information ratio) or differential returns compared to benchmarks (alphas). Portfolio alpha measurement is at the core of portfolio managers' concerns. Sharpe's model, which explains portfolio returns with the market index as the only risk factor, has quickly become restrictive. It now appears that one factor is not enough and that other factors have to be considered. Factor models were developed as an alternative to the CAPM, allowing a better description of portfolio risks and an accurate evaluation of managers' performance, in particular a better evaluation of portfolio alpha.

Beside models issued from portfolio theory, research in the area of performance measurement has also concerned the consideration of real market conditions and has developed techniques to fit cases where the restrictive hypotheses of portfolio theory were not observed.

The choice of a performance measurement technique has to reconcile the ease of implementation and the accuracy and comprehensibility of the resulting information. In order to render this information accessible to a wide audience, rating agencies, by relying on different performance techniques, propose a ranking of funds within the various investment categories, whereby a certain number of stars is attributed to each fund. This aspect of performance measurement, which was the subject of a separate study[2], will not be presented here.

After a description of portfolio returns estimation, this article presents the state of the art of performance measurement in the area of traditional investment. As performance measurement not only serves to evaluate results previously obtained by portfolio managers, but also as a predictor for their future results, a review of studies concerning performance persistence will end this article.

1. Portfolio returns calculation

Calculating return, which is simple for an asset or an individual portfolio, becomes more complex when it involves mutual funds with variable capital, where investors can enter or leave throughout the investment period. There are several ways to proceed, depending on the area that we are seeking to evaluate. After introducing the basic formula for calculating the return on a portfolio, we then describe the different methods that allow capital movements to be taken into account, with their respective advantages and drawbacks and their improvements. In the setting of performance measurement, the frequency to which the portfolio is evaluated is also an important choice. This will be developed at the end of this section.

1.1. Basic formula

The simplest method for calculating the return on a portfolio for a given period is obtained through an arithmetic calculation. We calculate the relative variation of the price of the portfolio over the period, increased, if applicable, by the dividend payment. The return R_{Pt} of the portfolio is given by:

$$R_{Pt} = \frac{V_t - V_{t-1} + D_t}{V_{t-1}}$$

where:
V_{t-1} denotes the value of the portfolio at the beginning of the period;
V_t denotes the value of the portfolio at the end of the period;
D_t denotes the cash flows generated by the portfolio during the evaluation period.

However, this formula is only valid for a portfolio that has a fixed composition throughout the evaluation period. In the area of mutual funds, portfolios are subject to contributions and withdrawals of capital on the part of investors. This leads to the purchase and sale of securities on the one hand, and to an evolution in the volume of capital managed, which is independent from variations in stock market prices, on the other. The formula must therefore be adapted to take this into account. The modifications to be made will be presented below.

1.2. Taking capital flows into account

Calculation methods have been developed to take into account the volume of capital and the time that capital is present in a portfolio. The methods that are currently listed and used are the internal rate of return, the capital-weighted rate of return and the time-weighted rate of return. Each of these methods evaluates a different aspect of the return. These methods are presented in detail below. We then look at how these various methods are perceived and used through the analysis of various and sometimes conflicting viewpoints contained in the academic literature.

1.2.1. Capital-weighted rate of return method

This rate is equal to the relationship between the variation in value of the portfolio during the period and the average of the capital invested during the period. Let's first consider the case where a single capital flow is produced during the period. The calculation formula is as follows:

$$R_{CWR} = \frac{V_T - V_0 - C_t}{V_0 + \frac{1}{2}C_t}$$

where:
V_0 denotes the value of the portfolio at the beginning of the period;
V_T denotes the value of the portfolio at the end of the period;
C_t denotes the cash flow that occurred at date t, where C_t is positive if it involves a contribution and negative if it involves a withdrawal.

This calculation is based on the assumption that the contributions and withdrawals of funds take place in the middle of the period. A more accurate method involves taking the real length of time that the capital was present in the portfolio. The calculation is then presented as follows:

$$R_{CWR} = \frac{V_T - V_0 - C_t}{V_0 + \frac{T-t}{T}C_t}$$

where T denotes the total length of the period.

1. Portfolio returns calculation

Let's now assume that there are n capital flows during the evaluation period. The formula is then generalised in the following manner:

$$R_{CWR} = \frac{V_T - V_0 - \sum_{i=1}^{n} C_{t_i}}{V_0 + \sum_{i=1}^{n} \frac{T - t_i}{T} C_{t_i}}$$

where t_i denotes the date on which the ith cash flow C_{t_i} occurs.

This calculation method is simple to use, but it actually calculates an approximate value of the true internal rate of return of the portfolio, because it does not take the capitalisation of the contributions and withdrawals of capital during the period into account. If there are a large number of capital flows, the internal rate of return, which is presented below, will be more precise. The advantage of this method, however, is that it provides an explicit formulation of the rate.

The capital-weighted rate of returns measures the total performance of the fund, so it provides the true rate of return from the fund holder's perspective. The result is strongly influenced by capital contributions and withdrawals.

1.2.2. Internal rate of return method

This method is based on an actuarial calculation. The internal rate of return is the discount rate that renders the final value of the portfolio equal to the sum of its initial value and the capital flows that occurred during the period. The cash flow for each sub-period is calculated by taking the difference between the incoming cash flow, which comes from the reinvestment of dividends and client contributions, and the outgoing cash flow, which results from payments to clients. The internal rate of return R_I is the solution to the following equation:

$$V_0 + \sum_{i=1}^{n-1} \frac{C_{t_i}}{(1 + R_I)^{t_i}} = \frac{V_T}{(1 + R_I)^T}$$

where:

T denotes the length of the period in years (this period is divided into n sub-periods);

t_i denotes the cash flow dates, expressed in years, over the period;

V_0 is the initial value of the portfolio;

V_T is the final value of the portfolio;

C_{t_i} is the cash flow on date t_i, withdrawals of capital are counted negatively and contributions positively.

As the formula is not explicit, the calculation is done iteratively. The internal rate of return only depends on the initial and final values of the portfolio. It is therefore independent from the intermediate portfolio values. However, it does depend on the size and dates of the cash flows, so the rate is, again, a capital-weighted rate of return.

The internal rate of return method allows us to obtain a more precise result than the capital-weighted rate of return when there are a significant number of capital flows of different sizes, but it takes more time to implement. The capital-weighted rate of return and the internal rate of return are the only usable methods if the value of the portfolio is not known at the time the funds are contributed and withdrawn.

1.2.3. Time-weighted rate of return method

The principle of this method is to break down the period into elementary sub-periods, during which the composition of the portfolio remains fixed. The return for the complete period is then obtained by calculating the geometric mean of the returns calculated for the sub-periods. The result gives a mean return weighted by the length of the sub-periods. This calculation assumes that the distributed cash flows, such as dividends, are reinvested in the portfolio.

We take a period of length T during which capital movements occur on dates $(t_i)_{1 \leq i \leq n}$. We denote the value of the portfolio just before a capital movement by V_{t_i} and the value of the cash flow by C_{t_i}. C_{t_i} is positive if it involves a contribution and negative if it involves a

1. Portfolio returns calculation

withdrawal. The return for a sub-period is then written as follows:

$$R_{t_i} = \frac{V_{t_i} - (V_{t_{i-1}} + C_{t_{i-1}})}{V_{t_{i-1}} + C_{t_{i-1}}}$$

This formula ensures that we compare the value of the portfolio at the end of the period with its value at the beginning of the period, i.e. its value at the end of the previous period increased by the capital paid or decreased by the capital withdrawn.

The return for the whole period is then given by the following formula:

$$R_{TWR} = \left[\prod_{i=1}^{n} (1 + R_{t_i}) \right]^{1/T} - 1$$

This calculation method provides a rate of return per dollar invested, independently of the capital flows that occur during the period. The result depends solely on the evolution of the value of the portfolio over the period. Gray and Dewar (1971) show that the time-weighted rate of return is the only well-behaved rate of return that is not influenced by contributions or withdrawals. To implement this calculation, we need to know the value and the date of the cash flows, together with the value of the portfolio at each of the dates.

There is one small reservation, however, when applying this method. To simplify matters, we often assume that the cash flows all occur at the end of the month, instead of considering the exact dates. In this case, the use of a continuous version of the rate smoothes the errors committed. It is given by the following formula:

$$r_{TWR} = \frac{1}{T} \left[ln\left(\frac{V_T}{V_0}\right) + \sum_{i=1}^{n-1} ln\left(\frac{V_{t_i}}{V_{t_i} + C_{t_i}}\right) \right]$$

The time-weighted rate of return enables a manager to be evaluated separately from the movements of capital, which he does not control. This rate only measures the impact of the manager's decisions on the performance of the fund. It is thus the best method for judging the quality of the manager. It allows the results of different managers to be compared objectively. It is considered to be the fairest method, and for that reason, is recommended by GIPS and used by the international performance measurement bodies.

1.2.4. Choice of methodology

The existence of several methods for calculating returns, which give different results, shows that a return value should always be accompanied by more information. It is appropriate to indicate the calculation method used, together with the total length of time for the historical data and the frequency with which the returns were measured.

In the setting of performance evaluation and performance attribution, the decision to take into account the movements of capital depends on what is measured. Several authors have considered the various methods of evaluating the rate of returns.

Chestopalov and Beliaev (2004/2005) describe an analytical approximation method for calculating the internal rate of return. They show that approximation of the IRR equation using linear Taylor expansion at a point with zero rate of return results in a Modified Dietz formula, both for discrete and continuous compounding. This means that separation of performance measurement methods into money-weighted and time-weighted rates of return is somewhat artificial. In fact, the time-weighted rate of return presently adopted as the CFA Institute standard is derived from the money-weighted rate of return as a particular approximation.

Spaulding (2003) also seems to share the opinion that the boundary between time-weighted and money weighted computation can sometimes be slim. He notices that when periods are relatively short and cash flows few, especially when market volatility is low, time-weighted and money-weighted tend to be relatively close. But, as we

1. Portfolio returns calculation

lengthen the time periods and increase the cash flows, especially with increased market volatility, the differences diverge and demonstrate the true differences between the two methodologies.

Campisi (2004) explains that performance attribution has evolved in parallel with performance measurement by accepting the time-weighted return as the preferred calculation method. In addition, the investment industry has accepted the assumption that increasing the frequency of calculation leads to improved accuracy in both the calculation and attribution of return. These assumptions have led to the wholesale abandonment of the money-weighted return calculation, both for performance measurement and performance attribution. He argues that while there is an irrefutable case for accepting the time-weighted return as the preferred method for measuring the return of an investment manager, there is an equally compelling case for accepting the money-weighted return as the appropriate method for evaluating the source of active return, i.e. that the money-weighted return is the correct method for performance attribution. He notices that time-weighted methodology cannot explain the active investment process as it excludes the very factors that define the active investment process, i.e. volatility, the timing of cash flows and the amount of cash flows. Time-weighting is appropriate for calculating the active return, while money-weighting is appropriate for analysing the manager's contribution to return and attribution return.

According to Campisi, an added benefit of a money-weighted methodology is the intuitive nature of the calculation. The portfolio's excess return is simply the weighted average of the issue alphas or sector alphas, and these can be "sliced and diced" to accommodate a variety of sector and industry groupings, style groupings or other risk factors that describe the active process or answer the client's questions. Furthermore, periods of less than one year can be calculated in a single step, eliminating the need to chain link attribution effects calculated over shorter periods, a process that often results in residuals that are difficult to resolve or explain to clients. Meanwhile, Campisi underlines that he does not recommend the money-weighted methodology for calculating the manager's return; he recommends money-weighting only for evaluating the contribution to return and attribution of return.

Illmer and Marty (2003) defend the money-weighted rate of return against the time-weighted rate of return (TWR). They decompose the money-weighted rate of return (MWR) into the three following effects: the benchmark effect, the management effect and the timing effect. The TWR of the portfolio is calculated by assuming no cash flows but considering the active asset allocations over the investment period. Adversely, the MWR of the portfolio reflects not only the active asset allocations but also the timing effects of the cash flow decisions. After calculating the overall returns of the benchmark and the portfolio, the benchmark effect equals the benchmark return, the management effect is the difference between the TWR of the portfolio and the benchmark return, and the timing effect is the difference between the MWR and TWR of the portfolio. Illmer and Marty show that neither the MWR calculation nor the MWR decomposition should be neglected but rather incorporated into the performance reporting and evaluation process. Not considering the MWR concept and ignoring the timing effects of cash flows bears the risk of misinterpretation and incorrect feedback in the investment process. The MWR concept still adds value and is by no means outdated. All participants are encouraged to reintroduce the MWR concept to the area of performance measurement as well as to the area of performance attribution.

1. Portfolio returns calculation

1.3. Evaluation over several periods

1.3.1. Arithmetic mean

The simplest calculation involves computing the arithmetic mean of the returns for the sub-periods, i.e. calculating:

$$\overline{R}_a = \frac{1}{T} \sum_{t=1}^{T} R_{Pt}$$

where the R_{Pt} are obtained arithmetically and T denotes the number of sub-periods. We thus obtain the mean return realised for a sub-period.

This mean overestimates the result, which can even be fairly far removed from the reality when the sub-period returns are very different from each other. The result also depends on the choice of sub-periods.

The arithmetic mean of the returns from past periods does, however, have one interesting interpretation. It provides an unbiased estimate of the return for the following period. It is therefore the expected return on the portfolio and can be used as a forecast of its future performance.

1.3.2. Geometric mean

The geometric mean (or compound geometric rate of return) allows us to link the arithmetic rates of return for the different periods, in order to obtain the real growth rate of the investment over the whole period. The calculation assumes that intermediate income is reinvested. The mean rate for the period is given by the following expression:

$$\overline{R}_g = \left[\prod_{t=1}^{T} (1 + R_{Pt}) \right]^{1/T} - 1$$

The geometric mean gives the real rate of return that is observed over the whole period, which is not true of the arithmetic mean.

In general, the return values for successive periods are not too different, and the arithmetic mean and geometric mean give similar results.

However, the arithmetic mean always gives a value that is greater than the geometric mean, unless the R_t returns are all equal, in which case the two means are identical. The greater the variation in R_t, the greater the difference between the two means.

We indicated that the arithmetic mean was interpreted as the expected return for the following period. However, if we are interested in the expected return over the long-term, and not just in the forthcoming period, it is better to consider the geometric rate. According to Filbeck and Tompkins (2004), geometric returns are the appropriate measure of historical performance because they accurately capture historic volatility. Assuming that past volatility is a predictor of future volatility, geometric returns provide a reasonable estimate of future returns.

1.3.3. Arithmetic mean versus geometric mean: what the literature says

Jacquier, Kane and Marcus (2003) investigated whether one should use arithmetic or geometric mean to forecast future fund performance. They explain that, as is generally noted in finance textbooks, an unbiased forecast of the terminal value of a portfolio requires compounding of its initial value at its arithmetic mean return for the length of the investment period. Despite this advice, many in the practitioner community seem to prefer geometric averages. They notice that compounding at the arithmetic average always produces an upwardly biased forecast of future portfolio value. This bias does not necessarily disappear even if the sample average return is itself an unbiased estimator of the true mean, the average is computed from a long data series, and returns are generated according to a stable distribution. In contrast, forecasts obtained by compounding at the geometric average will generally be biased downward. The biases are empirically significant. For investment horizon of 40 years, the difference in forecasts of cumulative performance can easily exceed a factor of 2. And the percentage difference in forecasts grows with the investment horizon, as well as with the imprecision in the estimate of

1. Portfolio returns calculation

the mean return. Indeed, the geometric average is unbiased, however, only in the special case when the sample period and the investment horizon are of equal length.

So they conclude that, when the arithmetic and geometric averages must be estimated subject to sampling error, neither approach yields unbiased forecasts. For typical investment horizons, the proper compounding rate is in between the arithmetic and geometric values. A weighted average of these two competing methods may allow an unbiased forecast. The proper weight for the geometric rate is the ratio of the investment horizon to the sample estimation period. Therefore, for short investment horizons, the arithmetic average is close to the "unbiased compounding rate", and as the horizon approaches the length of the estimation period, the weight on the geometric average approaches 1. For even longer horizons, both the geometric and arithmetic average forecasts will be upwardly biased. The percentage differences in forecast grow as the investment horizon and the imprecision in the estimate of the mean return grow.

1.4. Choice of frequency to evaluate performance

The improvements in technology have made it easier to monitor the performance of fund managers on a high frequency basis: quarterly, monthly or even daily. High frequency monitoring may have the positive effect of reducing perverse manager behaviour such as end-of-year window-dressing and tournament-induced changes in risk levels. However, more frequent investment performance monitoring also influences the distribution of observed excess returns. So an overly frequent measure of performance is not always the best choice, as has been underlined by some authors.

DiBartolomeo (2003) notices that in recent years it has become more and more commonplace for investment performance attribution analysis to be carried out with a daily observation periodicity. He explains that the justification given for

changing to daily observation frequency from longer periods (such as months) is that these analyses are believed to be better equipped to accurately reflect the actual investment returns on a fund. But, DiBartolomeo argues, such beliefs are based on a series of operational, mathematical and statistical assumptions that are demonstrably false. He asserts that applying typical attribution methods to daily data leads to analytical conclusions that are highly biased and unreliable and details this argument. For example, manager evaluation is normally performed using time-weighted returns (TWR) that are computed to remove the effect of cash flows. As the effect of cash flows in the data is removed, daily attribution analysis is not useful to investors in understanding their actual investment results. This argument is also developed by Darling and MacDougall (2002), who explain that there is information lost by using a TWR, and the more frequently the TWR is calculated, the more information may be lost. In that case, daily analysis can be regarded as less useful than monthly analysis. Moreover, lack of synchronization over a single day would cause an index fund to exhibit spurious active returns where none actually existed. Most problems of this type disappear in the case of monthly observation.

Another argument against measuring performance with excessively high frequency is related to the imperfections of the assumptions made upon the asset returns (investment returns are normally distributed; time series of returns are identically distributed; there is no serial correlation between investment returns). Academic literature illustrates that the imperfection of the assumptions with respect to quarterly or monthly return data is small, while for daily data these assumptions are rejected.

For example, Dimson and Jackson (2001) examined the impact that frequency of performance measurement has on the probability distribution of observed outcomes. With more frequent monitoring of rolling returns, there is a greatly increased probability of observing seemingly

1. Portfolio returns calculation

extreme observations. They demonstrated that if performance is appraised by focusing on returns to date, it is important to adjust the definition of extreme performance for the frequency with which returns are monitored. Failure to do so may lead to costly actions such as strategy revisions or manager terminations, which increase transaction costs and have detrimental effects on manager incentives. Marsh (1991) also points out that the danger with high-frequency monitoring is the way it might be used by investors who do not understand how to interpret such figures. Judgements about manager skill may be distorted by frequent monitoring. So it is important that investors recognize the impact of high frequency monitoring on the frequency with which they observe seemingly extreme performance events.

Performing industry-standard attribution procedures on a daily basis may lead to analytical conclusions that are likely to be biased and unreliable, leading to inappropriate management actions with respect to investment portfolios.

2. Absolute risk-adjusted performance measures

These measures evaluate funds' risk-adjusted returns, without any reference to a benchmark.

2.1. Sharpe ratio (1966)

This ratio, initially called the reward-to-variability ratio, is defined by:

$$S_P = \frac{E(R_P) - R_F}{\sigma(R_P)}$$

where:
$E(R_P)$ denotes the expected return of the portfolio;
R_F denotes the return on the risk-free asset;
$\sigma(R_P)$ denotes the standard deviation of the portfolio returns.

This ratio measures the return of a portfolio in excess of the risk-free rate, also called the risk premium, compared to the total risk of the portfolio, measured by its standard deviation. It is drawn from the capital market line, and not the Capital Asset Pricing Model (CAPM). It does not refer to a market index and is not therefore subject to Roll's (1977) criticism concerning the fact that the market portfolio is not observable.

Since this measure is based on the total risk of the portfolio, made up of the market risk and the unsystematic risk taken by the manager, it enables the performance of portfolios that are not very diversified to be evaluated. This measure is also suitable for evaluating the performance of a portfolio that represents an individual's total investment.

This ratio has been subject to generalisations since it was initially defined. It thus offers significant possibilities for evaluating portfolio performance, while remaining simple to calculate. One of the most common variations on this measure involves replacing the risk-free asset with a benchmark portfolio. The measure is then called the information ratio (cf. Sharpe, 1994) and will be presented in the next section describing relative risk-adjusted measures.

2.2. Treynor ratio (1965)

The Treynor ratio is defined by:

$$T_P = \frac{E(R_P) - R_F}{\beta_P}$$

where:
$E(R_P)$ denotes the expected return of the portfolio;
R_F denotes the return on the risk-free asset;
β_P denotes the beta of the portfolio.

This indicator measures the relationship between the return on the portfolio, above the risk-free rate, and its systematic risk. This ratio is drawn directly from the CAPM. Calculating this indicator requires a reference index to be chosen to estimate the beta of the portfolio. The results can then depend heavily on that choice, a fact that has been criticised by Roll.

The Treynor ratio is particularly appropriate for appreciating the performance of a well-diversified portfolio, since it only takes the systematic risk of the portfolio into account, i.e. the share of the risk that is not eliminated by diversification. It is also for this reason that the Treynor ratio is the most appropriate indicator for evaluating the performance of a portfolio that only constitutes a part of the investor's assets. Since the investor has diversified his investments, the systematic risk of his portfolio is all that matters.

Srivastava and Essayyad (1994) proposed Treynor's index, where beta is a composite measure generated by combining the expected asset returns from the traditional CAPM and the mean-lower partial moment CAPM. Their argument is that a composite forecast is more accurate than separate forecasts: valuable information missing from one model may be captured by the other model. They tested this measure on U.S.-based international funds and found that the composite beta is a statistically significant and meaningful parameter. They also ranked the performance of the funds using the Treynor index with three models (the CAPM, the mean-lower partial moment CAPM and a combination of the two), but their sample, which

2. Absolute risk–adjusted performance measures

was made up of 15 funds, was too small to test whether the difference in ranking obtained with the different models was significant.

2.3. Measure based on the VaR

The Value-at-Risk (VaR) is an indicator that enables to sum up the set of risks associated with a portfolio that is diversified over several asset classes in a single value. The VaR measures the risk of a portfolio as the maximum amount of the loss that the portfolio can sustain for a given level of confidence. This definition of risk can be used to calculate a risk-adjusted return indicator for evaluating the performance of a portfolio. In order to define a logical indicator, we divide the VaR by the initial value of the portfolio and thus obtain a percentage loss compared to the total value of the portfolio. We then calculate a Sharpe-like type of indicator in which the standard deviation is replaced with the risk indicator based on the VaR, as it was defined or:

$$\frac{R_P - R_F}{\dfrac{VaR_P}{V_P^0}}$$

where:
R_P denotes the return on the portfolio;
R_F denotes the return on the risk-free asset;
VaR_P denotes the VaR of the portfolio;
V_P^0 denotes the initial value of the portfolio.

Note that the calculation of VaR pre-supposes the choice of a confidence threshold. So the VaR-based ratios for different portfolios can only be compared for a same confidence level.

3. Relative risk-adjusted performance measures

These measures evaluate funds' risk-adjusted returns in reference to a benchmark.

3.1. Jensen's alpha (1968)

Jensen's alpha is defined as the differential between the return on the portfolio in excess of the risk-free rate and the return explained by the market model, or:

$$E(R_P) - R_F = \alpha_P + \beta_P (E(R_M) - R_F)$$

It is calculated by carrying out the following regression:

$$R_{Pt} - R_{Ft} = \alpha_P + \beta_P (R_{Mt} - R_{Ft}) + \varepsilon_{Pt}$$

The Jensen measure is based on the CAPM. The term $\beta_P (E(R_M) - R_F)$ measures the return on the portfolio forecast by the model. α_P measures the share of additional return that is due to the manager's choices.

The statistical significance of alpha can be evaluated by calculating the t-statistic of the regression, which is equal to the estimated value of the alpha divided by its standard deviation. This value is provided with the results of the regression. If the alpha values are assumed to be normally distributed, a t-statistic greater than two indicates that the probability of having obtained the result through luck, and not through skill, is strictly less than 5%. In this case, the average value of alpha is significantly different from zero.

Unlike the Sharpe and Treynor measures, the Jensen measure contains the benchmark. As with the Treynor measure, only the systematic risk is taken into account. This method, unlike the Sharpe and Treynor ratios, does not allow portfolios with different levels of risk to be compared. The value of alpha is actually proportional to the level of risk taken, measured by the beta. To compare portfolios with different levels of risk, we can calculate the Black-Treynor ratio[3] defined by:

$$\frac{\alpha_P}{\beta_P}$$

The Jensen alpha can be used to rank portfolios within peer groups. Peer groups group together portfolios that are managed in a similar manner and therefore have comparable levels of risk.

The Jensen measure is subject to the same criticism as the Treynor measure: the result depends on the choice of reference index. In addition, when managers practice a market timing strategy, which involves varying the beta according to anticipated movements in the market, the Jensen alpha often becomes negative, and does not then reflect the real performance of the manager. Performance analysis models taking variations in beta into account have been developed by Treynor and Mazuy and by Henriksson and Merton.

3.2. Extensions to Jensen's alpha

3.2.1. Jensen's alpha based on modified versions of the CAPM

3.2.1.1. Black's zero-beta model (1972)
This version of the CAPM was developed because two of the model's assumptions were called into question: the existence of a risk-free asset, and therefore the possibility of borrowing or lending at that rate, and the assumption of a single rate for borrowing and lending. Black showed that the CAPM theory was still valid without the existence of a risk-free asset, and developed a version of the model by replacing it with an asset or portfolio with a beta of zero. Instead of lending or borrowing at the risk-free rate, it is possible to take short positions on the risky assets.

With the Black model, the alpha is characterised by:

$$E(R_P) - E(R_Z) = \alpha_P + \beta_P (E(R_M) - E(R_Z))$$

3.2.1.2 Brennan's model (1970) taking taxes into account
The basic CAPM model assumes that there are no taxes. The investor is therefore indifferent to receiving income as a dividend or a capital

gain and investors all hold the same portfolio of risky assets. However, taxation of dividends and capital gains is generally different, and this is liable to influence the composition of the investors' portfolio of risky assets. Taking these taxes into account can therefore modify the equilibrium prices of the assets. As a response to this problem, Brennan developed a version of the CAPM that allows the impact of taxes on the model to be taken into account.

With the Brennan model, the alpha is characterised by:

$$E(R_P) - R_F = \alpha_P + \beta_P \left(E(R_M) - R_F - T(D_M - R_F) \right) + T(D_P - R_F)$$

with: $$T = \frac{T_d - T_g}{1 - T_g}$$

where:

T_d denotes the average taxation rate for dividends;

T_g denotes the average taxation rate for capital gains;

D_M denotes the dividend yield of the market portfolio;

D_P is equal to the weighted sum of the dividend yields of the assets in the portfolio, or

$$D_P = \sum_{i=1}^{n} x_i D_i$$

D_i denotes the dividend yield of asset i;

x_i denotes the weight of asset i in the portfolio.

3.2.2. Model where the risk premium is based on total risk

Elton and Gruber (1995) describe a performance measure using the same principle as the Jensen measure, namely measuring the differential between the managed portfolio and a theoretical reference portfolio. However, the risk considered is now the total risk and the reference portfolio is no longer a portfolio located on the Security Market Line, but a portfolio on the Capital Market Line, with the same total risk as the portfolio to be evaluated.

More specifically, this involves evaluating a manager who has to construct a portfolio with a total risk of σ_P. He can obtain this level of risk by splitting the investment between the market portfolio and the risk-free asset. Let A be the portfolio thereby obtained. This portfolio is situated on the Capital Market Line. Its return and risk respect the following relationship:

$$E(R_A) = R_F + \left(\frac{E(R_M) - R_F}{\sigma_M} \right) \sigma_P$$

since $\sigma_A = \sigma_P$. This portfolio is the reference portfolio.

If the manager thinks that he possesses particular stock-picking skills, he can attempt to construct a portfolio with a higher return for the fixed level of risk. Let P be his portfolio. The share of performance that results from the manager's choices is then given by:

$$E(R_P) - E(R_A) = E(R_P) - R_F - \left(\frac{E(R_M) - R_F}{\sigma_M} \right) \sigma_P$$

The return differential between portfolio P and portfolio A measures the manager's stock picking skills. The result can be negative if the manager does not obtain the expected result.

This measure is called total risk alpha (TRA) in Scholtz and Wilkens (2005), who notice that both this measure and the Jensen alpha can be easily manipulated by means of leverage.

In order to facilitate our understanding of the link between the total risk alpha and the Sharpe ratio, Gressis, Philippatos and Vlahos (1986) propose the following formulation for the total risk alpha: $TRA_i = \sigma_i (SR_i - SR_M)$, where SR refers to the Sharpe ratio.

3.2.3. Models suited to evaluating market timing strategy

The traditional Jensen alpha assumes that portfolio risk is stationary. It measures the additional return obtained, compared to the level of risk taken, by considering the average value

3. Relative risk–adjusted performance measures

of the risk over the evaluation period. The two first models presented below enable to take into account variations in the portfolio's beta over the investment period in portfolio performance evaluation. They actually involve statistical tests, which allow for qualitative evaluation of a market timing strategy, when that strategy is followed for the portfolio. These models allow us to measure the portfolio's Jensen alpha, and to assess whether the result was obtained through the right investment decisions being taken at the right time or through luck. The third model presents a decomposition of the Jensen measure, due to Grinblatt and Titman (1989b), and which enables timing to be evaluated.

3.2.3.1. The Treynor and Mazuy model (1966)

This model used a quadratic version of the CAPM, which provides us with a better framework for taking the adjustments made to the portfolio's beta into account, and thus for evaluating a manager's market timing capacity. Managers who anticipates market evolutions correctly will lower their portfolio's beta when the market falls. Their portfolio will thus depreciate less than if they had not made the adjustment. Similarly, when they anticipate a rise in the market, they increase their portfolio's beta, which enables them to make higher profits. The relationship between the portfolio return and the market return, in excess of the risk–free rate, should therefore be better approximated by a curve than by a straight line. The model is formulated as follows:

$$R_{Pt} - R_{Ft} = \alpha_P + \beta_P (R_{Mt} - R_{Ft}) + \delta_P (R_{Mt} - R_{Ft})^2 + \varepsilon_{Pt}$$

where:
R_{Pt} denotes the portfolio return vector for the period studied;
R_{Mt} denotes the vector of the market returns for the same period, measured with the same frequency as the portfolio returns;
R_{Ft} denotes the rate of the risk–free asset over the same period.

The α_P, β_P and δ_P coefficients in the equation are estimated through regression. If δ_P is positive and significantly different from

zero, we can conclude that the manager has successfully practised a market timing strategy.

This model was formulated empirically by Treynor and Mazuy (1966). It was then theoretically validated by Jensen (1972) and Bhattacharya and Pfleiderer (1983).

3.2.3.2. The Henriksson and Merton model (1981,1984)[4]

There are in fact two models: a non-parametric model and a parametric model. They are based on the same principle, but the parametric model seems to be more natural to implement. The non-parametric model is less frequently mentioned in the literature.

The non–parametric version of the model is older, and does not use the CAPM. It was developed by Merton (1981) and uses options theory. The principle is that of an investor who can split his portfolio between a risky asset and a risk-free asset, and who modifies the split over time, according to his anticipations on the relative performance of the two assets. If the strategy is perfect, the investor only holds stocks when their performance is better than that of the risk-free asset and only holds cash in the opposite case. The portfolio can be modelled by an investment in cash and a call on the better of the two assets. If the forecasts are not perfect, the manager will only hold a fraction of options f, situated between –1 and 1. The value of f allows us to evaluate the manager. To do so, we define two conditional probabilities:

P_1 denotes the probability of making an accurate forecast, given that the stocks beat the risk-free asset;
P_2 denotes the probability of making an accurate forecast, given that the risk-free asset beats the stocks.

We then have $f = P_1 + P_2 - 1$ and the manager has a market timing capacity if $f > 0$, i.e. if the sum of the two conditional probabilities is greater than one.

f can be estimated by using the following formula:

$$I_{t-1} = \alpha_0 + \alpha_1 y_t + \varepsilon_t$$

where:
$I_{t-1} = 1$, if the manager forecasts that the stocks will perform better than the risk-free asset during month t, otherwise 0;
$y_t = 1$, if the stocks actually did perform better than the risk-free asset, otherwise 0.

The coefficients in the equation are estimated through regression. α_0 gives the estimation of $1 - P_1$ and α_1 gives the estimation of $P_1 + P_2 - 1$. We then test the hypothesis $\alpha_1 > 0$.

Henriksson and Merton (1981) then developed a parametric model. The idea is still the same, but the formulation is different. It consists of a modified version of the CAPM which takes the manager's two risk objectives into account, depending on whether he forecasts that the market return will or will not be better than the risk-free asset return. The model is presented in the following form:

$$R_{Pt} - R_{Ft} = \alpha_P + \beta_{1P}(R_{Mt} - R_{Ft}) + \beta_{2P}D_t(R_{Mt} - R_{Ft}) + \varepsilon_{Pt}$$

with: $\quad D_t = 0$, if $R_{Mt} - R_{Ft} > 0$
$\quad\quad\quad D_t = -1$, if $R_{Mt} - R_{Ft} < 0$

The α_P , β_{1P} and β_{2P} coefficients in the equation are estimated through regression. The β_{2P} coefficient allows us to evaluate the manager's capacity to anticipate market evolution. If β_{2P} is positive and significantly different from zero, the manager has a good timing capacity.

These models have been presented while assuming that the portfolio was invested in stocks and cash. More generally, they are valid for a portfolio that is split between two categories of assets, with one riskier than the other, for example stocks and bonds, and for which we adjust the composition according to anticipations on their relative performance.

Goetzmann, Ingersoll and Ivkovic (2000) have studied the bias associated with this model used with monthly returns when market timers can make daily decisions. Their simulations suggest that this measure of timing skill is weak and biased downward when applied to the monthly returns of a daily timer. They propose an adjustment that mitigates this problem without the need to collect daily timer returns. Their approach consists in using daily returns to an index correlated to the timer's risky asset. Values of a daily put on the index are then cumulated over each month to form a regressor that captures timing skill.

3.2.3.3. Decomposition of Jensen measure: Grinblatt and Titman (1989b)

The Jensen measure has been subject to numerous criticisms, the main one being that a negative performance can be attributed to a manager who practices market timing. As we mentioned above, this comes from the fact that the model uses an average value for beta, which tends to overestimate the portfolio risk, while the manager varies his beta between a high beta and a low beta according to his expectations for the market. Grinblatt and Titman (1989b) present a decomposition of the Jensen measure in three terms: a term measuring the bias in the beta evaluation, a timing term and a selectivity term.

In order to establish this decomposition, we assume that there are n risky assets traded on a frictionless market, i.e. no transaction costs, no taxes and no restrictions on short selling. We assume that there is a risk-free asset. The assumptions are therefore those of the CAPM. We seek to evaluate the investor's performance over T time periods, by looking at the risk-adjusted returns of his portfolio.

We denote as:
r_{it} , the return on asset i in excess of the risk-free rate for period t;
x_{it} , the weight of asset i in the investor's portfolio for period t.

The return on the investor's portfolio for period t, in excess of the risk-free rate, is then given by:

3. Relative risk–adjusted performance measures

$$r_{Pt} = \sum_{i=1}^{n} x_{it}\, r_{it}$$

We denote as r_{Bt} the return in excess of the risk-free rate of a portfolio that is mean-variance efficient from an uninformed investor's viewpoint.

We can then write:

$$r_{it} = \beta_i\, r_{Bt} + \varepsilon_{it}$$

where:

$$\beta_i = \frac{cov(r_{it}, r_{Bt})}{var(r_{Bt})}$$

and:

$$E(\varepsilon_{it}) = 0$$

The portfolio return is then written as:

$$r_{Pt} = \beta_{Pt}\, r_{Bt} + \varepsilon_{Pt}$$

with:

$$\beta_{Pt} = \sum_{i=1}^{n} x_{it}\, \beta_i$$

and:

$$\varepsilon_{Pt} = \sum_{i=1}^{n} x_{it}\, \varepsilon_{it}$$

In order to establish the decomposition, we consider the limit, in the probabilistic sense, of the Jensen measure, which is written as follows:

$$J_P = \hat{r}_P - b_P\, \hat{r}_B$$

where:
b_p is the probability limit of the coefficient from the time-series regression of the portfolio returns against the reference portfolio series of returns;
$\hat{r}_P$ is the probability limit of the sample mean of the r_{Pt} series;
$\hat{r}_B$ is the probability limit of the sample mean of the r_{Bt} series.

Formally, the probability limit of a variable is defined as:

$$\hat{r}_P = p\lim\left[\frac{1}{T}\sum_{t=1}^{T} r_{Pt}\right]$$

It should be noted that b_P can be different from $\hat{\beta}_P$. This is the case when a manager practices market timing. $\hat{\beta}_P$ is then a weighted mean of the two betas used for the portfolio, while b_P is the regression coefficient obtained, without concerning oneself with the fact that the manager practices market timing.

We can write:

$$\hat{r}_P = p\lim\left[\frac{1}{T}\sum_{t=1}^{T} r_{Pt}\right]$$

or, by replacing r_{Pt} with its expression:

$$\hat{r}_P = p\lim\left[\frac{1}{T}\sum_{t=1}^{T}(\beta_{Pt}\, r_{Bt} + \varepsilon_{Pt})\right]$$

By arranging the terms in the expression, we obtain:

$$\hat{r}_P = \hat{\beta}_P\, \hat{r}_B + p\lim\left[\frac{1}{T}\sum_{t=1}^{T}\beta_{Pt}\,(r_{Bt} - \hat{r}_B)\right] + \hat{\varepsilon}_P$$

By using this formula in the Jensen measure expression, we obtain:

$$J_P = (\hat{\beta}_P - b_P)\hat{r}_B + p\lim\left[\frac{1}{T}\sum_{t=1}^{T}\beta_{Pt}\,(r_{Bt} - \hat{r}_B)\right] + \hat{\varepsilon}_P$$

This expression reveals three distinct terms:

- a term that results from the bias in estimated beta:

$$(\hat{\beta}_P - b_P)\hat{r}_B$$

- a term that measures timing:

$$p\lim\left[\frac{1}{T}\sum_{t=1}^{T}\beta_{Pt}\,(r_{Bt} - \hat{r}_B)\right]$$

- a term that measures selectivity:

$$\hat{\varepsilon}_P$$

If the weightings of the portfolio to be evaluated are known, the three terms can be evaluated separately. When the manager has no particular information in terms of timing, $\hat{\beta}_P = b_P$.

3. Relative risk–adjusted performance measures

3.2.4. Extensions to Jensen's alpha for international portfolios

3.2.4.1. McDonald's model (1973)

McDonald proposed a performance measure which is an extension to the Jensen measure. His model applies to a portfolio of stocks invested in the French and American markets. It is written as follows:

$$R_{Pt} - R_{Ft} = \Phi_P + \beta_{P1}^* (R_{M1,t} - R_{Ft}) + \beta_{P2}^* (R_{M2,t} - R_{Ft}) + e_{Pt}$$

where:

$R_{M1,t}$ denotes the rate of return of the French market in period t;

$R_{M2,t}$ denotes the rate of return of the American market in period t;

R_{Ft} denotes the rate of return of the risk-free asset in the French market in period t;

$\beta_{P1}^* = x_1 \beta_{P1}$ and $\beta_{P2}^* = x_2 \beta_{P2}$, with x_1 and x_2 being the proportions of the fund invested in each of the two markets and β_{P1} and β_{P2} the fund's coefficients of systematic risk compared to each of the two markets.

The overall excess performance of the fund Φ_P is broken down into:

$$\Phi_P = x_1 d_{P1} + x_2 d_{P2}$$

where d_{P1} and d_{P2} denote the excess performance of each of the two markets.

With this method we can attribute the contribution of each market to the total performance of the portfolio. This in turn allows us to evaluate the manager's capacity to select the best-performing international securities and to invest in the most profitable markets.

McDonald's model only considers investments in stocks and represents international investment as the American market alone. However, the model can be generalised for the case of investment in several international markets, and for portfolios containing several asset classes. This is what Pogue, Solnik and Rousselin propose.

3.2.4.2. Pogue, Solnik and Rousselin's model (1974)

Pogue, Solnik and Rousselin (1974) also proposed an extension to the Jensen measure for international portfolios. Their model measures the performance of funds invested in French and international stocks, without any limit on the number of countries, and in French bonds. The model is written as follows:

$$R_{Pt} = \alpha_P + x_{OF,P} \beta_{OF,P} (I_{OF,t} - R_{Ft}) + x_{AF,P} \beta_{AF,P} (I_{AF,t} - R_{Ft})$$
$$+ x_{WP} \beta_{WP} (I_{Wt} - R_{Wt}) + e_{Pt}$$

where:

R_{Ft} denotes the interest rate of the risk-free asset in the French market;

R_{Wt} denotes the eurodollar rate;

$I_{OF,t}, I_{AF,t}, I_{W,t}$ denote the returns on the three representative indices: the French bond market index, the French stock market index and the worldwide stock market index for period t;

$x_{OF,P}$, $x_{AF,P}$ and x_{WP} denote the proportion of the portfolio invested in each market;

$\beta_{OF,P}$, $\beta_{AF,P}$ and $\beta_{W,P}$ denote the systematic risk of each subset of the portfolio;

α_P denotes the portfolio's overall excess performance.

The result measures the manager's capacity to choose the most promising markets and his skill in selecting the best stocks in each market.

It is possible to go further in the analysis and breakdown of performance, by using multifactor models for international investment.

3.3. Information ratio

The information ratio, which is sometimes called the appraisal ratio, is defined by the residual return of the portfolio compared to its residual risk. The residual return of a portfolio corresponds to the share of the return that is not explained by the benchmark. It results from the choices made by the manager to overweight securities that he hopes will have a return greater than that of the benchmark. The residual, or diversifiable, risk measures the residual return

3. Relative risk-adjusted performance measures

variations. It is the tracking error of the portfolio and is defined by the standard deviation of the difference in return between the portfolio and its benchmark. The lower its value, the closer the risk of the portfolio to the risk of its benchmark. Sharpe (1994) presents the information ratio as a generalisation of his ratio, in which the risk-free asset is replaced by a benchmark portfolio. The information ratio is defined through the following relationship:

$$IR = \frac{E(R_P) - E(R_B)}{\sigma(R_P - R_B)}$$

where R_B denotes the return on the benchmark portfolio.

Managers seek to maximise its value, i.e. to reconcile a high residual return and a low tracking error. This ratio allows us to check that the risk taken by the manager, in deviating from the benchmark, is sufficiently rewarded. The information ratio is an indicator that allows us to evaluate the manager's level of information compared to the public information available, together with his skill in achieving a performance that is better than that of the average manager. As this ratio does not take the systematic portfolio risk into account, it is not appropriate for comparing the performance of a well-diversified portfolio with that of a portfolio with a low degree of diversification.

3.4. M² measure: Modigliani and Modigliani (1997)

Modigliani and Modigliani (1997) showed that the portfolio and its benchmark must have the same risk to be compared in terms of basis points of risk-adjusted performance. So they propose that the portfolio be leveraged or deleveraged using the risk-free asset. They defined the following measure:

$$RAP_P = \frac{\sigma_M}{\sigma_P}(R_P - R_F) + R_F$$

where:

$\dfrac{\sigma_M}{\sigma_P}$ is the leverage factor;

σ_M denotes the annualised standard deviation of the market returns;
σ_P denotes the annualised standard deviation of the returns of fund P;
R_P denotes the annualised return of fund P;
R_F denotes the risk-free rate.

This measure evaluates the annualised risk-adjusted performance (RAP) of a portfolio in relation to the market benchmark, expressed in percentage terms. According to Modigliani and Modigliani, this measure is easier to understand by the average investor than the Sharpe ratio. Modigliani and Modigliani propose the use of the standard deviation of a broad-based market index, such as the S&P 500, as the benchmark for risk comparison, but other benchmarks could also be used. For a fund with any given risk and return, the Modigliani measure is equivalent to the return the fund would have achieved if it had the same risk as the market index. The relationship therefore allows us to situate the performance of the fund in relation to that of the market. The most interesting funds are those with the highest RAP value.

The Modigliani measure is drawn directly from the capital market line. It can be expressed as the Sharpe ratio times the standard deviation of the benchmark index: the two measures are directly proportional. So Sharpe ratio and Modigliani measure lead to the same ranking of funds.

3.5. Market Risk-Adjusted Performance (MRAP) measure: Scholtz and Wilkens (2005)

Scholtz and Wilkens (2005) note that, as the RAP measure developed by Modigliani and Modigliani (1997) uses the standard deviation as risk measure, it is relevant only to investors who invest their entire savings in a single fund. So they propose a measure called market risk-adjusted performance (MRAP), following the same principle as Modigliani and Modigliani's measure, but measuring returns relative to market risk instead of total risk. As a result, the MRAP is suitable for investors who invest in many different assets.

3. Relative risk-adjusted performance measures

The idea is to compare funds on the basis of measure of market risk that is identical for all funds. The natural choice is the beta factor of the market index, $\beta_M = 1$. The market risk-adjusted performance for fund i is obtained by (de-)levering it in order to achieve a beta equal to one. If the fund's systematic risk exceeds that of the market ($\beta_i > 1$), this procedure can be interpreted as a fictitious sale of some fraction d_i of fund holdings and then an investment of the proceeds at the risk-free rate ($d_i < 0$). Similarly, if the fund's systematic risk falls below that of the market index ($\beta_i < 1$), the procedure corresponds to a fictitious loan at the risk-free rate, amounting to some fraction d_i, in order to increase investments into the fund ($d_i > 0$). The fraction d_i is calculated as follows:

$$d_i = \frac{1}{\beta_i} - 1$$

The market-risk-adjusted performance of fund i (MRAPi) is obtained by averaging the return of the market risk-adjusted fund (MRAF):

$$MRAP_i = \mu_{MRAFi} = (1 + d_i)\mu_i - d_i r_f = \frac{1}{\beta_i}(\mu_i - r_f) + r_f$$

On this basis, a fund, adjusted for market risk, outperformed the market index whenever its market risk-adjusted performance exceeds the return of the market index. Ranking funds according to their MRAPs corresponds to ranking them based on their Treynor Ratios, as:

$$MRAPi = TR_i + r_f$$

where TR refers to the Treynor Ratio.
The Treynor Ratio can also be expressed using Jensen Alpha (JA):

$$TR_i = \frac{JA_i}{\beta_i} + \mu_M - r_f = \frac{JA_i}{\beta_i} + TR_M$$

Then:

$$MRAP_i = \frac{JA_i}{\beta_i} + \mu_M = \frac{JA_i}{\beta_i} + TR_M + r_f$$

This implies that ranking based on MRAP is also equivalent to ranking based on alpha-beta ratio.

Like the M² measure, the MRAP measure is easy to interpret as it is expressed in basis points.

3.6. SRAP measure: Lobosco (1999)
This measure, described by Lobosco (1999), is a risk-adjusted performance measure that includes the management style as defined by Sharpe (1992). The SRAP (Style/Risk-Adjusted Performance is inspired by the work of Modigliani and Modigliani (1997). It is obtained as the difference between the RAP measure (or M²) for the portfolio and the RAP measure for the style benchmark representing the style of the portfolio. The first step to calculate the SRAP is to identify the combination of indices that best represents the manager's style. The use of a style benchmark instead of a broad market index enables a better and more accurate evaluation of managers' performance

3.7. Risk-adjusted performance measure in multimanagement: M3 — Muralidhar (2000, 2001)
Muralidhar has developed a new risk-adjusted performance measure that allows us to compare the performance of different managers within a group of funds with the same objectives (a peer group). This measure does contribute new elements compared to the Modigliani and Modigliani measure. It includes not only the standard deviations of each portfolio, but also the correlation of each portfolio with the benchmark and the correlations between the portfolios themselves. The method proposed by Muralidhar allows us to construct portfolios that are split optimally between a risk-free asset, a benchmark and several managers, while taking the investors' objectives into account, both in terms of risk and, above all, the relative risk compared to the benchmark.

The principle involves reducing the portfolios to those with the same risk in order to be able

3. Relative risk–adjusted performance measures

to compare their performance. This is the same idea as in Modigliani and Modigliani (1997) who compared the performance of a portfolio and its benchmark by defining transformations in such a way that the transformed portfolio and benchmark had the same standard deviation.

To create a correlation-adjusted performance measure, Muralidhar considers an investor who splits his portfolio between a risk-free asset, a benchmark and an investment fund. We assume that this investor accepts a certain level of annualised tracking error compared to his benchmark, which we call objective tracking error. The investor wishes to obtain the highest risk-adjusted value of alpha for a given portfolio tracking error and variance. We define as a, b and $(1 - a - b)$ the proportions invested respectively in the investment fund, the benchmark B and the risk-free asset F. The portfolio thereby obtained is said to be correlation-adjusted. It is denoted by the initials CAP (for correlation-adjusted portfolio). The return on this portfolio is given by:

$$R(CAP) = aR(manager) + bR(B) + (1 - a - b)R(F)$$

The proportions to be held must be chosen in an appropriate manner, so that the portfolio obtained has a tracking error equal to the objective tracking error and its standard deviation is equal to the standard deviation of the benchmark. The constraint on tracking error creates a unique target correlation between the CAP and the benchmark. This target correlation with that of the benchmark is given by:

$$\rho_{TB} = 1 - \frac{TE\,(Target)^2}{2\sigma_B^2}$$

The coefficients a and b are given by:

$$a = \frac{\sigma_B}{\sigma_P}\sqrt{\frac{(1 - \rho_{TB}^2)}{(1 - \rho_{PB}^2)}}$$

and: $$b = \rho_{TB} - a\frac{\sigma_P}{\sigma_B}\rho_{PB}$$

The search for the best return, in view of the constraints, leads to the calculation of optimal proportions that depend on the standard deviations and correlations of the different elements in the portfolio. The problem is considered here with a single fund, but it can be generalised to the case of several funds, to handle the case of portfolios split between several managers, and to find the optimal allocation between the different managers. The formulas that give the optimal weightings in the case of several managers have the same structure as those obtained in the case of a single manager, but they use the weightings attributed to each manager together with the correlations between the managers.

Once the optimal proportions have been calculated, the return on the correlation-adjusted portfolio has been fully determined. By carrying out the calculation for each fund being studied, we can rank the different funds.

The Muralidhar measure is certainly useful compared to the risk-adjusted performance measure that had been developed previously. We observe that the Sharpe ratio, the information ratio and the Modigliani and Modigliani measure turn out to be insufficient to allow investors to rank different funds and to construct their optimal portfolio. These risk-adjusted measures only include the standard deviations of the portfolios and the benchmark, even though it is also necessary to include the correlations between the portfolios and between the portfolios and the benchmark. The Muralidhar model therefore provides a more appropriate risk-adjusted performance measure, because it takes into account both the differences in standard deviation and the differences in correlations between the portfolios. It produces a ranking of funds that is different from that obtained with the other measures. In addition, neither the information ratio nor the Sharpe ratio indicates how to construct portfolios in order to produce the objective tracking error, while the Muralidhar measure provides the composition of the portfolios that satisfy the investors' objectives.

3. Relative risk-adjusted performance measures

The composition of the portfolio obtained through the Muralidhar method enables us to solve the problem of an institutional investor's optimal allocation between active and passive management, with the possible use of a leverage effect to improve the risk-adjusted performance.

3.8. SHARAD: Muralidhar (2001, 2002)

Muralidhar underlines that the M² and M³ measures do not take into account differences in data history among portfolios, which requires the use of the same data period to compare their results, namely the lowest common data period. Muralidhar explains that the longer the history, the higher the degree of confidence in the manager's skill. So he proposes a new measure with all the properties of the M3 measure, but which also allows differences in data history to be taken into account. He names this measure SHARAD for Skill, History and Risk-Adjusted.

Ambarish and Seigel (1996) demonstrate that the minimum number of data points, or time History H, required for skill to emerge from the noise is given by the following relation:

$$H > \frac{S^2(\sigma_P^2 - 2\rho\sigma_P\sigma_B + \sigma_B^2)}{\left[\left(R_P - \frac{\sigma_P^2}{2}\right) - \left(R_B - \frac{\sigma_B^2}{2}\right)\right]^2}$$

where:
R_P is the return of the manager's portfolio;
R_B is the return of the benchmark;
σ_P is the standard deviation of the manager's portfolio;
σ_B is the standard deviation of the benchmark;
ρ is the correlation of returns between the manager's portfolio and the benchmark;
S is the number of standard deviations for a given confidence level.

As H is given by performance history, Muralidhar solves for the degree of confidence S. Using the information ratio (IR) and tracking error (TE)

formula, the expression can be rewritten in terms of S, where S is a function of IR.

$$S < \sqrt{H}\left[IR(P) - \left(\frac{\sigma_P^2 - \sigma_B^2}{2TE(P)}\right)\right]$$

The confidence in skill is derived from converting S to percentage terms for a normal distribution, which is equivalent to computing the cumulative probability of a unit normal distribution with a standard deviation S. If one defines $C(S)$ as the cumulative probability of a unit normal with standard deviation of S for fund P, $C(S)$ will be the measure of confidence in skill.

When the term $\left(\dfrac{\sigma_P^2 - \sigma_B^2}{2TE(P)}\right)$ is generally small

or insignificant, the IR and length of data history will largely determine the confidence in skill. This is the case when tracking error is substantial and driven largely by low correlation between the portfolio and the benchmark (i.e. $\sigma_P \cong \sigma_B$). As a result, two portfolios with identical variances, information ratios and tracking errors, but differing only in length of history, will have different confidence in skill.

The SHARAD measure for portfolio P is a probability-adjusted measure, defined as:

$$SHARAD_P = C(S_P) * R(CAP_P)$$

This measure has all the properties of M³ and, in addition, it accounts for data period in a manner that is consistent with the skill evaluation.

3.9. AP Index: Aftalion and Poncet (1991)

This performance indicator is defined as the difference between the annual average expected return of the portfolio and that of its benchmark, from which is deduced the product of the difference between the portfolio risk (σ_P) and the benchmark risk (σ_B), multiplied by the price of risk PXR:

3. Relative risk–adjusted performance measures

$$AP = [E(R_P) - E(R_B)] - PXR[\sigma_P - \sigma_B]$$

The excess average return of the portfolio compared to its benchmark contributes positively in this index, while the excess risk contributes negatively. The price of risk, which has the same dimension as an average expected return divided by a standard deviation, allows the two terms in the AP index to have the same dimension. It represents the additional return (in percent) that investors require on average for each additional point of risk. It can be estimated with econometric methods using historical data for five to ten years.

AP index has the same dimension as Jensen alpha. It allows portfolios with the same benchmark to be ranked by decreasing AP index. This index is an alternative to the Sharpe ratio when risk premiums are negative, making negative Sharpe ratios difficult to interpret.

The AP index has a form relatively similar to the Sharpe's alpha described by Plantinga and de Groot (2001, 2002) and which is given by the following formula:

$$\alpha = E(R_P) - A\sigma_P^2$$

where:
$E(R_P)$ is the expected rate of return of the portfolio;
σ_P is the standard deviation;
A is the parameter driving the level of risk aversion.

3.10. Graham–Harvey (1997) measures

Graham and Harvey have developed two measures to make up for two problems encountered with the Sharpe ratio. First, the estimates are not precise enough when fund volatilities are too different. Second, the calculation of the Sharpe ratio is made assuming that the risk-free rate is constant and not correlated to risky asset returns.

The two measures provide different perspectives. The first measure (GH1) is obtained by drawing an efficient frontier using a reference index and cash. This results in a hyperbola as the variations of short-term interest rates are correlated with market return. Searching for the point with the same volatility as the fund under analysis and calculating the difference between the return of this portfolio and that of the portfolio being analysed provides us with the GH1 measure. The second measure (GH2) is obtained by searching for the set of portfolios that combines a given fund with cash. The difference between the return of the portfolio with the same volatility as the market index and the market index return provides us with the GH2 measure.

The GH2 measure is similar to the M² measure proposed by Modigliani and Modigliani (1997). However, Modigliani and Modigliani do not allow for curvature in the efficient frontier. That is, they assume that the cash return has zero variance and zero covariance with other assets.

3.11. Efficiency ratio: Cantaluppi and Hug (2000)

While the relative methods of performance measurement tend to answer the question "What is the performance of a portfolio relative to other portfolios?", the efficiency ratio methodology proposed by Cantaluppi and Hug tends to answer the question "Which performance could have been achieved by the portfolio?".

To explain how this measure works, Cantaluppi and Hug consider two portfolios, named *A* and *B*, with portfolio *A* having a higher Sharpe ratio than portfolio *B*. However, portfolio *B* is on the efficient frontier, while portfolio *A* is not. The efficiency ratio is computed as the distance to the ex post efficient frontier. The efficiency ratio of portfolio *A* is obtained by dividing its return by that of a portfolio with similar volatility, but located on the efficient frontier. The efficiency ratio of portfolio *B* is equal to 100%, as it is located on the efficient frontier, while that of portfolio *A* is strictly lower than 100% and

therefore lower than the portfolio B efficiency ratio. A portfolio ranking based on the efficiency ratio is thus different from one obtained using the Sharpe ratio.

3.12. Investor-Specific Performance Measurement (ISM): Scholtz and Wilkens (2004)

Scholtz and Wilkens consider the situation of an investor holding a portfolio P and wanting to invest additional money without changing his initial portfolio. The additional amount will be put in a portfolio Di. The overall portfolio of the investor will then be made up of portfolio P in proportion $(1 - w_D)$, and portfolio Di. in proportion w_D. Portfolio Di is made up of a fund i in proportion w_i, and the risk-free rate in proportion $(1 - w_i)$.

The ISM performance measure is based on classic dominance considerations. The starting point is that at a predetermined expected return of the overall portfolio, the portfolio with the lowest variance dominates all the other portfolios with higher variance. Given the expected return of the overall portfolio μ_{G^+}, an investor can build an appropriate overall portfolio for each fund and then identify the fund which dominates the other ones. The ISM measure is defined as:

$$ISM_i = -w_D \left(\frac{\frac{\mu_{D^+} - r_f}{\sigma_P}}{\frac{\mu_i - r_f}{\sigma_P}} \right)^2 - 2(1 - w_D) \frac{\mu_{D^+} - r_f}{\sigma_{iP}/\sigma_P^2}$$

with:

$$\mu_{D^+} = \frac{\mu_{G^+} - \mu_P}{w_D} + \mu_P$$

μ_P as the expected return of the portfolio P;
r_f as the risk free rate;

Investors can compare funds based on the ISM measure. This measure depends on the investor-specific portfolio structure and the investor-specific expected return of the overall portfolio.

Hence it is referred to as the investor-specific performance measure. A fund j with a higher ISM is superior to a fund k with a lower ISM at a given portfolio structure and a predetermined expected return of the overall portfolio. The lower the ISM of a fund, the higher the variance of the returns of the overall portfolio for a given expected return μ_{G^+}.

If the portfolio P is the market index the formula can be rewritten in the following form:

$$ISM_i = -w_D \left(\frac{\frac{\mu_{D^+} - r_f}{\sigma_M}}{S_i} \right)^2 - 2(1 - w_D) \frac{\mu_{D^+} - r_f}{T_i}$$

with: $\quad \mu_{D^+} = \frac{\mu_{G^+} - \mu_M}{w_D} + \mu_M$

where:
S_i is the Sharpe ratio of fund i;
T_i is the Treynor ratio of fund i.

It appears that the value of ISMi for different expected returns of the overall portfolio is determined by the Sharpe ratio and the Treynor ratio of fund i. No further fund specific information is needed to assess the performance of the particular fund. According to the formula above, the higher the Sharpe ratio and the Treynor ratio of a fund i, the higher the fund's ISM.

4. Some new research on the Sharpe ratio

4.1. Critics and limitations of the Sharpe ratio

The CAPM assumes either that all asset returns are normally distributed and thus symmetrical or that investors have mean-variance preferences and thus ignore skewness. Assuming only that the rate of return on the market portfolio is independently and identically distributed and that markets are perfect, Leland (1999) shows that the CAPM and its risk measures are invalid: the market portfolio is mean-variance inefficient, and the CAPM alpha mismeasures the value added by investment managers.

Cvitanic, Lazrak and Wang (2004) show that the typical mean-variance efficiency justification for using the Sharpe ratio, valid in a static setting, typically fails in a multi-period setting. The trading strategy that leads to the most desirable portfolio for each quarter and for four consecutive quarters is not the same as the strategy that gives the highest Sharpe ratio for a year. As a consequence, unless the investor's investment horizon exactly matches the performance measurement period of the portfolio manager, the portfolio with the highest Sharpe ratio is not necessarily the most desirable from the investor's point of view.

4.2. "Double" Sharpe Ratio: Vinod and Morey (2001)

One problem with the Sharpe ratio is that its denominator is random, as it is computed using a data sample of returns on a given history and not the whole population of returns. So it is difficult to evaluate its risk estimation. Vinod and Morey (2001) proposed a modified version of the Sharpe ratio, called the Double Sharpe ratio, to take into account estimation risk. This ratio is defined as follows:

$$DS_P = \frac{S_P}{\sigma(S_P)}$$

where $\sigma(S_P)$ is the standard deviation of the Sharpe ratio estimate, or the estimation risk.

To calculate this standard deviation they use bootstrap methodology to generate a great number of resamples from the original returns sample and derive a series of Sharpe ratios. Using the 30 largest-growth mutual funds, Vinod and Morey found that the ranking of mutual funds by the Sharpe and Double Sharpe ratios can be quite different.

4.3. Generalised Sharpe ratio: Dowd (2000)

Dowd proposes an approach based on the VaR to evaluate an investment decision. Dowd considers the case of an investor who holds a portfolio that he is thinking of modifying, by introducing, for example, a new asset. He will study the risk and return possibilities linked to a modification of the portfolio and choose the situation for which the risk-return balance seems to be sufficiently favourable. To do that, he could decide to define the risk in terms of the increase in the portfolio's VaR. He will change the portfolio if the resulting incremental VaR (IVaR) is sufficiently low compared to the return that he can expect. This can be formalised as a decision rule based on Sharpe's decision rule.

Sharpe's rule states that the most interesting asset in a set of assets is the one that has the highest Sharpe ratio. By calculating the existing Sharpe ratio and the Sharpe ratio for the modified portfolio and comparing the results, we can then judge whether the planned modification of the portfolio is desirable.

By using the definition of the Sharpe ratio, we find that it is useful to modify the portfolio if the returns and standard deviations of the portfolio before and after the modification are linked by the following relationship:

$$\frac{R_P^{new}}{\sigma_{R_P^{new}}} \geq \frac{R_P^{old}}{\sigma_{R_P^{old}}}$$

where:
R_P^{old} and R_P^{new} denote, respectively, the return on the portfolio before and after the modification;

$\sigma_{R_P^{old}}$ and $\sigma_{R_P^{new}}$ denote, respectively, the standard deviation of the portfolio before and after the modification.

4. Some new research on the Sharpe ratio

We assume that part of the new portfolio is made up of the existing portfolio, in proportion $(1 - a)$, and the other part is made up of asset A in proportion a.

The return on this portfolio is written as follows:

$$R_P^{new} = aR_A + (1 - a)R_P^{old}$$

where R_A denotes the return on asset A.

By replacing R_P^{new} with its expression in the inequality between the Sharpe ratios, we obtain:

$$\frac{aR_A + (1 - a)R_P^{old}}{\sigma_{R_P^{new}}} \geq \frac{R_P^{old}}{\sigma_{R_P^{old}}}$$

which finally gives:

$$R_A \geq R_P^{old} + \frac{R_P^{old}}{a}\left(\frac{\sigma_{R_P^{new}}}{\sigma_{R_P^{old}}} - 1\right)$$

This relationship indicates the inequality that the return on asset A must respect for it to be advantageous to introduce it into the portfolio. The relationship depends on proportion a. It shows that the return on asset A must be at least equal to the return on the portfolio before the modification, to which is added a factor that depends on the risk associated with the acquisition of asset A. The higher the risk, the higher the adjustment factor and the higher the return on asset A will have to be.

Under certain assumptions, this relationship can be expressed through the VaR instead of the standard deviation. If the portfolio returns are normally distributed, the VaR of the portfolio is proportional to its standard deviation, or:

$$VaR = -\alpha \sigma_{R_P} W$$

where:
α denotes the confidence parameter for which the VaR is estimated;
W is a parameter that represents the size of the portfolio;
σ_{R_P} is the standard deviation of the portfolio returns.

By using this expression of the VaR, we can calculate:

$$\frac{VaR^{new}}{VaR^{old}} = \frac{W^{new}\sigma_{R_P^{new}}}{W^{old}\sigma_{R_P^{old}}}$$

which enables us to obtain the following relationship:

$$\frac{\sigma_{R_P^{new}}}{\sigma_{R_P^{old}}} = \frac{VaR^{new}}{VaR^{old}} \frac{W^{old}}{W^{new}}$$

We assume that the size of the portfolio is conserved. We therefore have $W^{old} = W^{new}$. We therefore obtain simply, after substituting into the return on A relationship:

$$R_A \geq R_P^{old} + \frac{R_P^{old}}{a}\left(\frac{VaR^{new}}{VaR^{old}} - 1\right)$$

The incremental VaR between the new portfolio and the old portfolio, denoted by IVaR, is equal to the difference between the old and new value, or $IVaR = VaR^{new} - VaR^{old}$.

By replacing the term $\left(\frac{VaR^{new}}{VaR^{old}} - 1\right)$ in the

inequality according to the IVaR, we obtain:

$$R_A \geq R_P^{old} + \frac{R_P^{old}}{a}\left(\frac{IVaR}{VaR^{old}}\right) = R_P^{old}\left(1 + \frac{1}{a}\frac{IVaR}{VaR^{old}}\right)$$

By defining function η_A as:

$$\eta_A(VaR) = \frac{1}{a}\frac{IVaR}{VaR^{old}}$$

we can write:

$$R_A \geq (1 + \eta_A(VaR))R_P^{old}$$

where $\eta_A(VaR)$ denotes the percentage increase in the VaR occasioned by the acquisition of asset A, divided by the proportion invested in asset A.

4. Some new research on the Sharpe ratio

Dowd also considers the case where the reference is no more the risk-free asset, but a benchmark portfolio. In that case, the standard deviation of the difference between the portfolio and its benchmark is no longer equal to σ_{Rp}, but is given by:

$$\sigma_d = \sqrt{\sigma_{Rp}^2 + \sigma_{Rb}^2 - 2\rho_{RpRb}\sigma_{Rp}\sigma_{Rb}}$$

The decision rule is now to acquire the new position if:

(1)

$$R_A - R_b \geq (R_p^{old} - R_b) + (\sigma_d^{new}/\sigma_d^{old} - 1)(R_p^{old} - R_b)/a$$

Since d is the difference between the relevant (i.e., old or new) portfolio return and the benchmark return, we can regard the standard deviation of d as the standard deviation of the return to a combined position that is long the relevant portfolio and short the benchmark. This combined position has its own VaR, which Dembo (1997) calls the benchmark-VaR, or BVaR. Assuming normality, the ratio of standard deviations in (1) is then equal to the ratio of the new to old BVaRs, as given by the following equation (2):

(2)

$$\sigma_d^{new}/\sigma_d^{old} = BVaR^{new}/BVaR^{old}$$

Substituting (2) in (1) and rewriting it in its BVaR form, we obtain:

$$R_A - R_b \geq (1 + \eta_A(BVaR))(R_p^{old} - R_b)$$

This rule is an exact analogue of the previous rule, but with $R_A - R_b$ and $R_p^{old} - R_b$ instead of R_A and R_p^{old}, and the BVaR elasticity instead of the earlier VaR elasticity.

The generalized Sharpe ratio is superior to the standard Sharpe ratio because it is valid regardless of the correlations of the investments being considered with the rest of the portfolio. Since it is derived in a mean-variance world, it should be used cautiously where departures from normality are considerable.

4.4. Negative excess returns: Israelsen (2005)

Israelsen (2005) notices that the Sharpe ratio and information ratio, two performance indicators often used to rank mutual funds, may lead to spurious ranking when fund excess returns are negative. In that case, the fund with the higher ratio is not always the best one. This can be easily seen in the following example. The argument below concerns the information ratio, but is similar in the case of the Sharpe ratio.

	Excess return over the S&P 500	Tracking error	Information ratio
Fund A	-6.96	13.86	-0.50
Fund B	-3.62	5.03	-0.72

The table shows that the information ratio of fund A is higher than that of fund B, though fund B is preferable to fund A as its excess return is higher and its tracking error lower.

Israelsen proposes to correct this anomaly by modifying the standard information ratio and Sharpe ratio. He introduces an exponent to the denominator of these ratios, equal to the fund excess return divided by its absolute value. Using the previous notations, the modified Sharpe ratio is defined as:

$$S_P^{modified} = \frac{E(R_P) - R_F}{\sigma(R_P)^{(E(R_P)-R_F)/abs(E(R_P)-R_F)}}$$

and the modified information ratio is defined as:

$$IR_P^{modified} = \frac{E(R_P) - E(R_B)}{\sigma(R_P - R_B)^{(E(R_P)-R_F)/abs(E(R_P)-R_F)}}$$

We note that these modified ratios coincide with the standard ones, when excess returns are positive.

Applying the modified information ratio to the example leads to a value of -96.47 for fund A and a value of -18.21 for fund B, which reverse the ranking comparatively to the standard

ratio. The ratios proposed by Israelsen allow us to consistently rank funds, whether the fund excess returns are positive or negative. As the modification in the ratios causes enormous range in its size, Israelsen points out that their values give no useful information and should only be used as a ranking criterion.

5. Measures based on downside risk and higher moments

5.1. Actuarial approach: Melnikoff (1998)

In this approach, the investor's aversion to risk is characterised by a constant (W) which measures his gain-shortfall equilibrium, i.e. the relationship between the expected gain desired by the investor to make up for a fixed shortfall risk. The average annual risk-adjusted return is then given by:

$$RAR = R - (W - 1)S$$

where:

S denotes the average annual shortfall rate;
W denotes the weight of the gain-shortfall aversion;
R denotes the average annual rate of return obtained by taking all the observed returns.

For an average individual, W is equal to two, which means that the individual will agree to invest if the expected amount of his gain is double the shortfall. In this case, we have simply:

$$RAR = R - S$$

5.2. Sortino ratio[5]

An indicator such as the Sharpe ratio, based on the standard deviation, does not allow us to know whether the differentials compared to the mean were produced above or below the mean. The notion of semi-variance brings a solution to this problem by taking into account the asymmetry of risk. The calculation principle is the same as that of the variance, apart from the fact that only the returns that are lower than the mean are taken into account. It therefore provides a skewed measure of the risk, which corresponds to the needs of investors, who are only interested in the risk of their portfolio losing value. It is written as follows:

$$\frac{1}{T} \sum_{\substack{0 \le t \le T \\ R_{Pt} < \bar{R}_P}} (R_{Pt} - \bar{R}_P)^2$$

where R_{Pt} denotes the return on portfolio P for sub-period t;

$\bar{R}_P$ denotes the mean return on asset P over the whole period;
T denotes the number of sub-periods.

The lower partial moment generalises the notion of semi-variance. It measures the risk of falling below a target return set by the investor. The mean return is replaced in this formula by the value of the target return below which the investor does not wish to drop. This notion can then be used to calculate the risk-adjusted return indicators that are more specifically appropriate for asymmetrical return distributions. The best known indicator is the Sortino ratio. It is defined on the same principle as the Sharpe ratio. However, the risk-free rate is replaced with the minimum acceptable return (MAR), i.e. the return below which the investor does not wish to drop, and the standard deviation of the returns is replaced with the standard deviation of the returns that are below the MAR, or:

$$\text{Sortino Ratio} = \frac{E(R_P) - MAR}{\sqrt{\frac{1}{T} \sum_{\substack{t=0 \\ R_{Pt} < MAR}}^{T} (R_{Pt} - MAR)^2}}$$

This measure allows a distinction between "good" and "bad" volatility: it does not penalise portfolios with returns that are far from their mean return, but higher than this mean, contrary to the Sharpe ratio.

5.3. Fouse index

Sortino and Price (1994) described a measure using utility theory in a mean-downside risk environment — the Fouse index:

$$Fouse = E(R) - B\delta^2$$

where:
B is a parameter representing the degree of risk aversion of the investor;
δ is the downside risk with respect to the minimal acceptable rate of return.
This index is equivalent to Sharpe's alpha[6] in a mean-downside risk environment.

5 - Cf. Sortino and Van der Meer (1991).

6 - Cf. Plantinga and de Groot (2001).

5. Measures based on downside risk and higher moments

5.4. Upside potential ratio: Sortino, Van der Meer and Plantinga (1999)

This ratio, developed by Sortino, Van der Meer and Plantinga, is the probability-weighted average of returns above the reference rate. It is defined as:

$$UPR = \frac{\sum_{t=1}^{T} \iota^{+} \frac{1}{T}(R_t - MAR)}{\left[\sum_{t=1}^{T} \iota^{-} \frac{1}{T}(R_t - MAR)^2\right]^{1/2}}$$

where T is the number of periods in the sample, R_t is the return of an investment in period t, $\iota^{+} = 1$ if $R_t > MAR$, $\iota^{+} = 0$ if $R_t \leq MAR$, $\iota^{-} = 1$ if $R_t \leq MAR$ and $\iota^{-} = 0$ if $R_t > MAR$.

The numerator of the Upside Potential ratio is the expected return above the MAR and can be thought of as the potential for success. The denominator is downside risk as calculated in Sortino and van der Meer (1991) and can be thought of as the risk of failure. An important advantage of using the upside potential ratio rather than the Sortino ratio is the consistency in the use of the reference rate for evaluating both profits and losses.

According to Sortino, Miller and Messina (1997), more stable estimates of risk are possible by employing style analysis. Sharpe (1992) developed a procedure for identifying a manager's style in terms of a set of passive indexes which enables to construct a style benchmark for the manager. Using the distribution of returns of the style benchmark, instead of the manager's return distribution, it is possible to calculate downside risk using a longer data history than that of the manager.

5.5. Symmetric downside–risk Sharpe ratio: Ziemba (2005)

The Sharpe ratio relies on mean-variance theory, so it is only suited for quadratic preferences or normal distributions. Lo (2002) points out that care must be used in Sharpe ratio estimations when the investment returns are not independent and identically distributed (iid). In order to penalize a fund manager for losing, but not for winning, Ziemba calculates a Sharpe ratio using downside variance instead of variance. He defines the downside variance as:

$$\sigma_{x-}^{2} = \frac{\sum_{i=1}^{n} (x_i)_{-}^{2}}{n-1}$$

where the x_i taken are those below zero. The reference is zero instead of the mean of the returns, so it measures the downside risk.

The total variance is computed as twice the downside variance. And the corresponding Sharpe ratio is given by:

$$S_{-} = \frac{\bar{R} - R_F}{\sqrt{2}\sigma_{x-}}$$

This measure is closely related to the Sortino ratio, which considers downside risk only.

5.6. Higher moment measure of Hwang and Satchell (1998)

When portfolios returns are not normally distributed, higher moments such as skewness and kurtosis need to be considered to adjust for the non-normality and to account for the failure of variance to measure risk accurately. In these cases, a higher-moment CAPM should prove more suitable than the traditional CAPM and so a performance measure based on higher moments may also be more accurate. Assuming the validity of the three-moment CAPM and a quadratic return generating process of the form:

$$r_{Pt} - r_f = a_{0p} + a_{1p}(r_{mt} - r_f) + a_{2p}(r_{mt} - E(r_m))^2 + \varepsilon_{pt}$$

we can define a performance measure of a portfolio under the three-moment CAPM as :

$$a_p = \mu_p - \lambda_1 \mu_m - \lambda_2 (\beta_{pm} - \gamma_{pm})$$

where:

$$\lambda_1 = \frac{\gamma_m^2 \gamma_{pm} - (\theta_m - 1)\beta_{pm}}{\gamma_m^2 - (\theta_m - 1)}$$

$$\lambda_2 = \frac{\gamma_m \sigma_m}{\gamma_m^2 - (\theta_m - 1)}$$

5. Measures based on downside risk and higher moments

with:

$$\mu_p = E(r_{pt} - r_f)$$

$$\mu_m = E(r_{mt} - r_f)$$

$$\sigma_m = E[(r_m - rE(r_m))^2]^{1/2}$$

and:

$$\gamma_m = \frac{E[(r_m - E(r_m))^3]}{\sigma_m^3}$$

$$\theta_m = \frac{E[(r_m - E(r_m))^4]}{\sigma_m^4}$$

$$\beta_{pm} = \frac{E[(r_p - E(r_p))(r_m - E(r_m))]}{E[(r_m - E(r_m))^2]}$$

$$\gamma_{pm} = \frac{E[(r_p - E(r_p))(r_m - E(r_m))^2]}{E[(r_m - E(r_m))^3]}$$

γ_m and θ_m are the skewness and kurtosis of the market returns, and β_{pm} and γ_{pm} are beta and coskewness respectively. If the market returns are normal, then $\lambda_1 = \beta_{pm}$ and $\lambda_2 = 0$ and the alpha measure is therefore equivalent to Jensen's alpha. This measure suffers from the same limitations as Jensen's alpha but does account for non-gaussiannity.

5.7. Omega measure: Keating and Shadwick (2002)

As notified by their authors, the analysis underlying the omega measure development is to be related with downside risk, lower partial moments and gain-loss literature. Keating and Shadwick observe that an assumption that the two first moments, i.e. mean and variance, fully describe a distribution of returns causes inaccuracies in performance measurement. According to them, performance measurement also requires higher moments. They also advocate the usefulness of a return level reference, aside from the mean return in the description of the

risk-reward characteristics of a portfolio. In response to these observations, they introduce a performance evaluation measure called omega which incorporates all of the higher moments of a returns distribution. Omega also takes into account the level of return against which a given outcome will be viewed as a gain or a loss, which is additional information, even in the case where returns are normally distributed.

The principle of the measure consists in partitioning returns into loss and gain relative to a return threshold corresponding to the minimum acceptable return (MAR) for an investor, and then considering the probability weighted ratio of returns above and below the partitioning. The Omega measure is defined as a function of the MAR threshold in the following way:

$$\Omega(MAR) = \frac{\int_{MAR}^{b} (1 - F(x))dx}{\int_{a}^{MAR} F(x)dx}$$

where:
(a, b) is the interval of possible returns;
F is the cumulative distribution function for the returns.

Omega may be used to rank manager performance. The rankings will depend on the interval of returns under consideration and will incorporate all higher moment effects. Because of the additional information it employs, omega is expected to produce significant different rankings of portfolios compared to those derived with Sharpe ratios, alphas or value-at-risk.

This measure is specifically recommended for evaluating portfolios that do not exhibit normally distributed return distributions. For this reason, it usually appears in a setting of hedge fund portfolios. Meanwhile, the issue of not normal distribution also exists in the context of traditional investment, though to a lesser extent. Note that in the cases where higher moments are of little significance, the omega measure is in accordance with traditional measure and avoids the need to estimate means and variances.

6. Performance measurement method using a conditional beta: Ferson and Schadt (1996)

6.1. The model

The method is based on a conditional version of the CAPM, which is consistent with the semi-strong form of market efficiency as interpreted by Fama (1970).

The conditional formulation of the CAPM allows the return of each asset i to be written as follows:

$$r_{i,t+1} = \beta_{im}(I_t)r_{m,t+1} + u_{i,t+1} \qquad (1a)$$

with:

$$E(u_{i,t+1}/I_t) = 0 \qquad (1b)$$

and:

$$E(u_{i,t+1}r_{m,t+1}/I_t) = 0 \qquad (1c)$$

r_{it} denotes the return on asset i in excess of the risk-free rate, or:

$$r_{it} = R_{it} - R_{Ft}$$

where R_{Ft} denotes the risk-free interest rate for period t.

In the same way, r_{mt} denotes the return on the market in excess of the risk-free rate, or:

$$r_{mt} = R_{mt} - R_{Ft}$$

These relationships are valid for $i = 0,...,n$, where n denotes the number of assets, and for $t = 0,...,T-1$, where T denotes the number of periods.

I_t denotes the vector that represents the public information at time t. The beta of the regression, $\beta_{im}(I_t)$, is a conditional beta, i.e. it depends on the information vector I_t. Beta will therefore vary over time depending on a certain number of factors. When I_t is the only information used, no alpha term appears in the regression equation, because the latter is null. The error term in the regression is independent from the information, which is translated by relationship (1b). This corresponds to the efficient market hypothesis.

Using asset return relationships, we can establish a portfolio return relationship. By hypothesising that the investor uses no information other than the public information, we deduce that the investor's portfolio beta β_{Pm} only depends on I_t. By using a development from Taylor, we can approximate this beta through a linear function, or:

$$\beta_{Pm}(I_t) = b_{0P} + B_P' i_t$$

In this relationship, b_{0P} can be interpreted as an average beta. It corresponds to the unconditional mean of the conditional beta, or:

$$b_{0P} = E(\beta_{Pm}(I_t))$$

The elements of vector B_p are the response coefficients of the conditional beta with respect to the information variables I_t.

i_t denotes the vector of the differentials of I_t compared to its mean, or:

$$i_t = I_t - E(I)$$

From this we deduce a conditional formulation of the portfolio return:

$$r_{P,t+1} = b_{0P} r_{m,t+1} + B_P' i_t r_{m,t+1} + u_{P,t+1}$$

with:

$$E(u_{P,t+1}/I_t) = 0$$

and:

$$E(u_{P,t+1}r_{m,t+1}/I_t) = 0$$

The model's stochastic factor is a linear function of the market return, in excess of the risk-free rate, the coefficients of which depend linearly on I_t.

The model thereby developed enables the traditional performance measures, which came from the CAPM, to be adapted by integrating a time component. These applications are discussed in the following section.

6. Performance measurement method using a conditional beta: Ferson and Schadt (1996)

6.2. Application to performance measurement[7]

6.2.1. The Jensen measure

The traditional Jensen measure does not provide satisfactory results when the risk and return are not constant over time. The model proposed enables this problem to be solved.

To evaluate the performance of portfolios, we employ an empirical formulation of the model which uses the term α_{CP}, or:

$$r_{P,t+1} = \alpha_{CP} + b_{0P}\, r_{m,t+1} + B_P^{'}\, i_t\, r_{m,t+1} + e_{P,t+1}$$

α_{CP} represents the average difference between the excess return of the managed portfolio and the excess return of a dynamic reference strategy. This model provides a better forecast of alpha. A manager with a positive conditional alpha is a manager who has a higher return than the average return of the dynamic reference strategy.

The first step involves determining the content of the information to be used. This is the same as using explanatory factors. Ferson and Schadt (1996) propose linking the portfolio risk to market indicators, such as the market index dividend yield (DY_t) and the return on short-term T-bills (TB_t), lagged by one period compared to the estimation period.

The dy_t and tb_t variables denote the differentials compared to the average of the variables DY_t and TB_t, or:

$$\begin{cases} dy_t = DY_t - E(DY) \\ tb_t = TB_t - E(TB) \end{cases}$$

We therefore have:

$$i_t = \begin{bmatrix} dy_t \\ tb_t \end{bmatrix}$$

or:

$$B_P = \begin{pmatrix} b_{1P} \\ b_{2P} \end{pmatrix}$$

The conditional beta is then written as follows:

$$\beta_P = b_0 + b_1\, dy_t + b_2\, tb_t$$

from which we have the conditional formulation of the Jensen model:

$$r_{P,t+1} = \alpha_{cp} + b_{0p}\, r_{m,t+1} + b_{1p}\, dy_t\, r_{m,t+1}$$

$$+ b_{2p}\, tb_t\, r_{m,t+1} + e_{p,t+1}$$

where α_{cp} represents the conditional performance measure, b_{0p} denotes the conditional beta and b_{1p} and b_{2p} measure the variations in conditional beta compared to the dividend yield and the return on the T-bills.

The coefficients are evaluated through regression from the time-series of the variables.

6.2.2. The Treynor and Mazuy model

The non-conditional approach does not draw a distinction between the skill in using macroeconomic information that is available to everybody and a manager's specific stock-picking skill. The conditional approach allows these to be separated.

The conditional formulation, applied to the Treynor and Mazuy model, involves adding a conditional term to the first order, or:

$$r_{P,t+1} = \alpha_{CP} + b_{0P}\, r_{m,t+1} + B_P^{'}\, i_t\, r_{m,t+1}$$
$$+ \gamma_P\, r_{m,t+1}^2 + e_{P,t+1}$$

where γ_P denotes the market timing coefficient. The conditional formulation is only used in the part that is shared with the Jensen measure and not in the model's additional term.

By using an information vector with two components, we obtain:

$$r_{P,t+1} = \alpha_{CP} + b_{0P}\, r_{m,t+1} + b_{1P}\, dy_t\, r_{m,t+1}$$

$$+ b_{2P}\, tb_t\, r_{m,t+1} + \gamma_P\, r_{m,t+1}^2 + e_{P,t+1}$$

The coefficients of the relationship are estimated through ordinary regressions.

6. Performance measurement method using a conditional beta: Ferson and Schadt (1996)

6.2.3. The Henriksson and Merton model

The manager seeks to forecast the differential between the market return and the expectation of the return that is conditional on the available information, or:

$$u_{m,t+1} = r_{m,t+1} - E(r_{m,t+1}/I_t)$$

Depending on whether the result of this forecast is positive or negative, the manager chooses a different value for the conditional beta of his portfolio.

If the forecast is positive, then:

$$\beta_{up}(I_t) = b_{0up} + B'_{up}i_t$$

If the forecast is negative, then:

$$\beta_d(I_t) = b_{0d} + B'_d i_t$$

Henriksson and Merton's conditional model is written as follows:

$$r_{P,t+1} = \alpha_{CP} + b_{0d}r_{m,t+1} + B'_d i_t r_{m,t+1} + \gamma_c r^*_{m,t+1}$$
$$+ \Delta' i_t r^*_{m,t+1} + u_{P,t+1}$$

with:
$$\gamma_c = b_{0up} - b_{0d}$$

$$\Delta = B_{up} - B_d$$

and:

$$r^*_{m,t+1} = r_{m,t+1}I\{r_{m,t+1} - E(r_{m,t+1}/I_t) > 0\}$$

where $I\{\}$ denotes the indicator function.

More explicitly, if $r_{m,t+1} - E(r_{m,t+1}/I_t) > 0$,

then: $r_{P,t+1} = \alpha_{CP} + b_{0up}r_{m,t+1} + B'_{up}i_t r_{m,t+1} + u_{P,t+1}$

and if $r_{m,t+1} - E(r_{m,t+1}/I_t) \le 0$,

then: $r_{P,t+1} = \alpha_{CP} + b_{0d}r_{m,t+1} + B'_d i_t r_{m,t+1} + u_{P,t+1}$

By again taking our example of an information vector with two components, the model is written:

$$r_{P,t+1} = \alpha_{CP} + b_{0d}r_{m,t+1} + b_{1d}dy_t r_{m,t+1} + b_{2d}tb_t r_{m,t+1}$$
$$+ \gamma_c r^*_{m,t+1} + \delta_1 dy_t r^*_{m,t+1} + \delta_2 tb_t r^*_{m,t+1} + u_{P,t+1}$$

with:
$$B_{up} = \begin{pmatrix} b_{1up} \\ b_{2up} \end{pmatrix}$$

$$B_d = \begin{pmatrix} b_{1d} \\ b_{2d} \end{pmatrix}$$

$$\Delta = \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} = \begin{pmatrix} b_{1up} - b_{1d} \\ b_{2up} - b_{2d} \end{pmatrix}$$

The market timing strategy is evaluated by determining the coefficients of the equation through regression. In the absence of market timing, γ_c and the components of Δ are null. If the manager successfully practices market timing, we must have $\gamma_c + \Delta' i_t > 0$, which means that the conditional beta is higher when the market is above its conditional mean, given the public information, than when it is below its conditional mean.

6.3. Model with a conditional alpha[8]

The evaluation of conditional performance enables the portfolio risk and return to be forecast with more accuracy. A better estimation of the beta leads to a better estimation of the alpha. But to be more specific in evaluating portfolio performance, we can assume that the alpha also follows a conditional process. This allows us to evaluate excess performance that varies over time, instead of assuming that it is constant. The relationship given by the conditional alpha is written as follows:

$$\alpha_{CP} = a_P(i_t) = a_{0P} + A'_P i_t$$

6. Performance measurement method using a conditional beta: Ferson and Schadt (1996)

The regression equation that enables the Jensen alpha to be evaluated is then written:

$$r_{P,t+1} = a_{0P} + \dot{A_P}\dot{i_t} + b_{0P}r_{m,t+1} + \dot{B_P}\dot{i_t}r_{m,t+1} + u_{P,t+1}$$

By again taking the information model that is made up of two variables, the alpha component is written:

$$\alpha_{CP} = a_{0P} + a_{1P}dy_t + a_{2P}tb_t$$

with:

$$\dot{A_P} = \begin{pmatrix} a_{1P} \\ a_{2P} \end{pmatrix}$$

The model is then written

$$r_{P,t+1} = a_{0P} + a_{1P}dy_t + a_{2P}tb_t + b_{0P}r_{m,t+1}$$

$$+ b_{1P}r_{m,t+1}dy_t + b_{2P}r_{m,t+1}tb_t + u_{P,t+1}$$

The coefficients of the equation are estimated through regression.

6.4. The contribution of conditional models

The study of mutual funds shows that their exposure to risk changes in line with available information on the economy. The use of a conditional measure eliminates the negative Jensen alphas. Their value is brought back to around zero. The viewpoint developed in Christopherson, Ferson and Turner (1999) is that a strategy that only uses public information should not generate superior performance. The methods for measuring the performance of market timing strategies, such as Treynor and Mazuy's and Henriksson and Merton's, are also improved by introducing a conditional component into the model.

7. Performance analysis methods that are not dependent on the market model

The Roll criticism, by underlining the impossibility of measuring the true market portfolio, cast doubt over the performance measurement models that refer to the market portfolio. Measures that were independent from the market model were therefore developed to respond to the criticisms of the model and propose an alternative. These measures are mainly used for evaluating a manager's market timing strategy.

7.1. The Cornell measure (1979)[9]

The Cornell measure involves evaluating a manager's superiority as his capacity to pick stocks that have a higher return than their normal return. This measure does not use the market portfolio. The asset returns are the direct references used. The difficulty is to define the return that is considered to be "normal" for each asset.

In practice, the Cornell measure is calculated as the average difference between the return on the investor's portfolio, during the period in which the portfolio is held, and the return on a reference portfolio with the same weightings, but considered for a different period than the investor's holding period. The calculation can therefore only be carried out when the securities are no longer held in the investor's portfolio, i.e. at the end of the investment management period. The limitations of this measure relate to the number of calculations required to implement it and the possibility that certain securities will disappear during the period.

Formally, by using the notation from section 3.2.3.3., presenting the decomposition of the Jensen measure, the asymptotic value of the Cornell measure can be written as follows:

$$C = \hat{r}_P - \hat{\beta}_P \hat{r}_B$$

By replacing $\hat{r}_P$ with its expression established in section 3.2.3.3, or:

$$\hat{r}_P = \hat{\beta}_P \hat{r}_B + plim\left[\frac{1}{T}\sum_{t=1}^{T}\beta_{Pt}(r_{Bt} - \hat{r}_B)\right] + \hat{\varepsilon}_P$$

we obtain:

$$C = plim\left[\frac{1}{T}\sum_{t=1}^{T}\beta_{Pt}(r_{Bt} - \hat{r}_B)\right] + \hat{\varepsilon}_P$$

i.e. the sum of the selectivity and timing components from the decomposition of the Jensen measure.

The Jensen and Cornell measures both attribute a null performance to an investor who has no particular skill in terms of timing or in terms of selectivity.

7.2. The Grinblatt and Titman measure (1989a, b): Positive Period Weighting Measure

The Cornell measure does correct the problem of the Jensen measure, which wrongly attributes a negative performance to managers who practice market timing. But this measure requires the weightings of the assets that make up the managed portfolio to be known. Grinblatt and Titman proposed a measure that is an improvement on the Jensen measure, enabling the performance of market timers to be evaluated correctly, but which does not require information on portfolio weightings.

This model is based on the following principle. When a manager truly possesses market-timing skills, his performance should tend to repeat over several periods. The method therefore involves taking portfolio returns over several periods, and attributing a positive weighting to each of them. The weighted average of the reference portfolio returns in excess of the risk-free rate must be null. This condition translates the fact that the measure attributes a null performance to uninformed investors.

The Grinblatt and Titman measure is thus defined by:

$$GB = \sum_{t=1}^{T} w_t (R_{Pt} - R_{Ft})$$

7. Performance analysis methods that are not dependent on the market model

with:

$$\sum_{t=1}^{T} w_t = 1$$

and:

$$\sum_{t=1}^{T} w_t (R_{Bt} - R_{Ft}) = 0$$

where:

R_{Pt} denotes the return on the portfolio for period t;

R_{Bt} denotes the return on the reference portfolio for period t;

R_{Ft} denotes the risk-free rate for period t;

w_t denotes the weighting attributed to the return for period t.

A positive Grinblatt and Titman measure indicates that the manager accurately forecasted the evolution of the market.

This method presents the disadvantage of not being very intuitive. In addition, in order to implement it we need to determine the weightings to be assigned to the portfolio returns for each period.

7.3. Performance measure based on the composition of the portfolio: Grinblatt and Titman study (1993)

Grinblatt and Titman also proposed a method for evaluating market timing based on studying the evolution of the portfolio's composition. The method is therefore fairly different from most other performance measurement methods. The methodology is similar to Cornell's (1979). The measure is based on the study of changes in the composition of the portfolio. It relies on the principle that an informed investor changes the weightings in his portfolio according to his forecast on the evolution of the returns. He overweights the stocks for which he expects a high return and lowers the weightings of the other stocks. A non-null covariance between the weightings of the assets in the portfolio and the returns on the same assets must ensue. The measure is put together by aggregating the covariances.

It is defined by:

$$\sum_{i=1}^{n} \sum_{t=1}^{T} (r_{it} (x_{it} - x_{i,t-k}))/T$$

where:

r_{it} denotes the return on security i, in excess of the risk-free rate, for period t;

x_{it} and $x_{i,t-k}$ denote the weighting of security i at the beginning of each of the periods t and $t - k$.

The expectation of this measure will be null if an uninformed manager modifies the portfolio. It will be positive if the manager is informed.

This measure does not use reference portfolios. It requires the returns on the assets and their weightings within the portfolio to be known. Like the Cornell measure, this method is limited by the significant number of calculations and data required to implement it.

7.4. Measure based on levels of holdings and measure based on changes in holdings: Cohen, Coval and Pastor (2005)

Cohen, Coval and Pastor (2005) observe that the traditional measures that rely solely on historical returns are imprecise, because return histories are often short. They develop a performance evaluation approach in which a fund manager's skill is judged by the extent to which the manager's investment decisions resemble the decisions of managers with distinguished performance records. They proposed two performance measures that use historical returns and holdings of many funds to evaluate the performance of a single fund. The first measure is based on level of holdings, while the second one is based on changes in holdings. They compare their new measures with those proposed by Grinblatt and Titman (1993), which also rely on fund and note that these measures do not exploit the information contained in the holdings and returns of other funds. This specific point is the innovation of their new measures.

7. Performance analysis methods that are not dependent on the market model

7.4.1. Measure based on levels of holdings

For each stock *n*, Cohen, Coval and Pastor define a quality measure as the average skill of all managers who hold stock n in their portfolios, weighted by how much of the stock they hold, i.e.

$$\bar{\delta}_n = \sum_{m=1}^{M} v_{mn} \alpha_m$$

with:

$$v_{mn} = \frac{w_{mn}}{\sum_{m=1}^{M} w_{mn}}$$

where:

α_m denotes the reference measure of skill for manager *m*. It is supposed to be measured against a benchmark taking into account any style effects for which the manager should not be rewarded (the authors notice that several choices of skill measures are possible);

w_{mn} denotes the current weight on stock *n* in manager *m*'s portfolio;

M is the total number of managers;

N is the total number of stocks.

Stocks with high quality are those that are held mostly by highly skilled managers. Managers who hold stocks of high quality are likely to be skilled because their investment decisions are similar to those of other skilled managers.

The measure of a manager's performance is then given by:

$$\delta_m^* = \sum_{m=1}^{M} w_{mn} \bar{\delta}_m$$

This is the average quality of all stocks in the manager's portfolio, where each stock contributes according to its portfolio weight. This is a weighted average of the usual skill measure across all managers.

The corresponding estimated value is obtained by replacing α_m by its estimator $\hat{\alpha}_m$.

$$\hat{\delta}_m^* = \frac{\sum_{j=1}^{M} \sum_{n=1}^{N} w_{mn} w_{jn} \hat{\alpha}_j}{\sum_{m=1}^{M} w_{mn}}$$

The weight assigned to the performance of manager *j* is a loose measure of covariance between the weights of managers *m* and *j*.

7.4.2. Measure based on changes in holdings

Cohen, Coval and Pastor also propose to compare managers' trades. Their trade-based performance measure judges a manager's skill by the extent to which recent changes in his holdings match those of managers with outstanding past performance. This measure is also a weighted average of the traditional skill measures, but now the weights are essentially the covariances between the concurrent changes in the manager's portfolio weights and those of the other managers. According to the trade-based measure, the manager is skilled if he tends to buy stocks that are concurrently purchased by other managers who have performed well, and if he tends to sell stocks that are concurrently purchased by managers who have performed poorly. This performance measure exploits similarities between changes in the managers' holdings, rather than their levels.

The authors underline that their approach adds value only if there is some commonality in the managers' investment decisions. They argue that their measures are particularly useful for funds with relatively short return histories. A vast majority of real-world mutual funds have return histories shorter than 20 years. They also found that their measures are well-suited for empirical applications that involve ranking managers.

They have conducted an empirical study, successively using the CAPM alpha, the Fama-French (1993) alpha, and the four-factor alpha following Carhart (1997). Using their measures to rank managers, the authors found strong predictability in the returns of U.S. equity funds. They observe that the persistence in performance weakens when the momentum

7. Performance analysis methods that are not dependent on the market model

factor is included. They compared the predictive power of alpha and their two new measures and found that these three measures seem capable of predicting fund returns, with an advantage for the measure based on levels of holdings. They also investigated whether their measures contain useful information for forecasting fund returns not contained in alpha and found that their measure provides information about future fund returns that is not contained in the standard measures. Their results suggest that the measure based on levels of holdings contains significant information about future fund returns above and beyond alpha and that most of the information contained in alpha is already in the measure based on levels of holdings. The measure based on changes in holdings also adds incremental information about future fund returns over and above alpha. However, alpha seems to contain some incremental information beyond this measure. As a result, mutual fund portfolio strategies would benefit from combining the information in these measures.

They notice that their measures of manager's skill rely on the manager's most recent holdings or trades, without considering his historical holdings. The idea is that a manager's current decisions should be more informative than his past decisions about future performance. The authors suggest that historical holdings could contain useful information about managerial skill and that it would be interesting to design performance measures that exploit similarities in historical holdings or trade across managers, and perhaps also the correlation between historical holdings and subsequent holding returns as in Grinblatt and Titman (1993). Since such measures use yet more information, they might be able to predict fund returns even more effectively than the simple measures proposed here.

8. Factor models: more precise methods for evaluating alphas

Factor models have been developed as an alternative to the CAPM, following Roll's (1977) criticism. As they rely on fewer hypotheses than the CAPM, they may be validated empirically. These models enable us to explain portfolio returns with a set of factors (various market indexes, macroeconomic factors, fundamental factors), instead of just the theoretical and non observable market portfolio, and thus provide more specific information on risk analysis and evaluation of managerial performance. These models generalised Jensen's alpha. Their general formulation is as follows:

$$R_{it} = \alpha_i + \sum_{k=1}^{K} b_{ik} F_{kt} + \varepsilon_{it}$$

where:

R_{it} denotes the rate of return for asset i;

α_i denotes the expected return for asset i;

b_{ik} denotes the sensitivity (or exposure) of asset i to factor k;

F_{kt} denotes the return of factor k with $E(F_k) = 0$;

ε_{it} denotes the residual (or specific) return of asset i, i.e. the share of the return that is not explained by the factors, with $E(\varepsilon_i) = 0$. The residual returns of the different assets are independent from each other and independent from the factors. We therefore have: $cov(\varepsilon_i, \varepsilon_j) = 0$, for $i \neq j$ and $cov(\varepsilon_i, F_k) = 0$, for all i and k.

There are several types of factor models.

8.1. Explicit factor models based on macroeconomic variables

These models are derived directly from Arbitrage Pricing Theory (APT) developed by Ross (1976). The risk factors that affect asset returns are approximated by observable macroeconomic variables that can be forecasted by economists. The choice of the number of factors, namely five macroeconomic factors and the market factor, comes from the first empirical tests carried out by Roll and Ross with the help of a factor analysis method. The classic factors in the APT models are industrial production, interest rates,

oil prices, differences in bond ratings and the market factor. These factors are described in Chen, Roll and Ross (1986).

8.2. Explicit factor models based on microeconomic factors (also called fundamental factors)

This approach is much more pragmatic. The aim now is to explain the returns on the assets with the help of variables that depend on the characteristics of the firms themselves, and no longer from identical economic factors for all assets. The modelling no longer uses any theoretical assumptions but considers a factor breakdown of the average asset returns directly. The model assumes that the factor loadings of the assets are functions of the firms' attributes, called fundamental factors. The realisations of the factors are then estimated by regression. Here again, the choice of explanatory variables is not unique. The factors used are, among others, the size, the country, the industrial sector, etc. Below are some examples of this kind of models, among the most popular.

8.2.1. Fama and French's three-factor model[10]

Fama and French have highlighted two important factors that characterise a company's risk, as a complement to the market beta: the book-to-market ratio and the company's size measured by its market capitalisation. They therefore propose a three-factor model, which is formulated as follows:

$$E(R_i) - R_F = b_{i1}(E(R_M) - R_F)$$
$$+ b_{i2}E(SMB) + b_{i3}E(HML)$$

where:

$E(R_i)$ denotes the expected return of asset i;

R_F denotes the rate of return of the risk-free asset;

$E(R_M)$ denotes the expected return of the market portfolio;

SMB (small minus big) denotes the difference between returns on two portfolios: a small-capitalisation portfolio and a large-capitalisation portfolio;

8. Factor models: more precise methods for evaluating alphas

HML (high minus low) denotes the difference between returns on two portfolios: a portfolio with a high book-to-market ratio and a portfolio with a low book-to-market ratio;

b_{ik} denotes the factor loadings.

8.2.2. Carhart's four-factor model (1997)

This model is an extension of Fama and French's three-factor model. The additional factor is momentum, which enables the persistence of the returns to be measured. This factor was added to take the anomaly revealed by Jegadeesh and Titman (1993) into account. With the same notation as above, this model is written:

$$E(R_i) - R_F = b_{i1}(E(R_M) - R_F)$$
$$+ b_{i2}E(SMB) + b_{i3}E(HML) + b_{i4}(PR1YR)$$

where *PR1YR* denotes the difference between the average of the highest returns and the average of the lowest returns from the previous year.

8.2.3. The Barra model

The Barra multifactor model is the best known example of commercial application of a fundamental factor model. The model uses thirteen risk indices[11].

The returns are characterised by the following factor structure:

$$R_{it} = \sum_{k=1}^{K} b_{ikt}\alpha_{kt} + u_{it}$$

where:

R_{it} denotes the return on security *i* in excess of the risk-free rate;

b_{ik} denotes the factor loading or exposure of asset *i* to factor *k*;

α_k denotes the return on factor *k*;

u_i denotes the specific return on asset *i*.

This model assumes that asset returns are determined by the fundamental characteristics of the firm. These characteristics constitute the exposures or betas of the assets. The approach therefore assumes that the exposures are known and then calculates the factors.

8.3. Implicit or endogenous factor models

The idea behind this approach is to use the asset returns to characterise the unobservable factors. It is natural to assume that the factors which influence the returns leave an identifiable trace. These factors are therefore extracted from the asset return database through a factor analysis method and the factor loadings are jointly calculated. To do this, we perform a principal component analysis which enables us to explain the behaviour of the observed variables using a smaller set of non observed implicit variables. From a mathematical point of view, this consists in turning out a set of n correlated variables in a set of orthogonal variables (the implicit factors), which reproduce the original information that was in the correlation structure. Each implicit factor is defined as a linear combination of the initial variables. As the implicit variables are chosen for their explaining power, it seems natural that a given number of explicit factors may explain a larger part of the variance-covariance matrix of asset returns than the same number of explicit factors. This approach was originally used for the first tests on the APT model. This type of model is used by the firms Quantal and Advanced Portfolio Technology (APT). However, the search of implicit factors has the drawback of not allowing us to identify the nature of the factors, except the first one which exhibits a strong correlation with the market index.

The explicit factor models appear, at least in theory, to be simpler to use, but they assume that the factors that generate the asset returns are known and that they can be observed and measured without errors. As multifactor model theory does not specify the number or nature of the factors, their choice results from empirical studies and there is no unicity. Implicit factor models solve the problem of the choice of factors, since the model does not make any prior assumptions about the number and nature of the factors. As they are directly extracted from asset returns, it therefore enables the true factors to be used: there is no risk of including bad factors, or omitting good ones. However, factors are thus

11 - A detailed list of the factors used can be found in Amenc and Le Sourd (2003).

8. Factor models: more precise methods for evaluating alphas

mute variables and it may be difficult to give them an economic significance.

8.4. Application to performance measure

The multifactor models have a direct application in investment fund performance measurement. In analysing portfolio risk according to various dimensions, it is possible to identify the sources of risk to which the portfolio is submitted and to evaluate the associated reward. The result is a better control of portfolio management and an orientation of this one toward the good sources of risk, which lead to an improvement of its performance. These models contribute more information to performance analysis than the Sharpe, Treynor and Jensen indices. The asset returns could be decomposed linearly according to several risk factors common to all the assets, but with specific sensitivity to each. Once the model has been determined, we can attribute the contribution of each factor to the overall portfolio performance. This is easily done when the factors are known, which is the case for models that use macroeconomic factors or fundamental factors, but becomes more difficult when the nature of the factors has not been identified. Performance analysis then consists of evaluating whether the manager was able to orient the portfolio towards the most rewarding risk factors.

Practically speaking, the implementation of factor models is carried out in two stages. First, betas are estimated through regression of asset returns on factors returns:

$$R_{it} = \beta_{i0} + \sum_{k=1}^{K} \beta_{ik} F_{kt} + \varepsilon_{it}$$

Lambdas are then estimated through cross-sectional regression for each date t. The dependent variables are the returns in excess of the risk-free rate $R_{it} - R_f$, for $i = 1,..., n$, assuming there are n assets (or funds, or portfolios). The dependent variables are the estimated $\hat{\beta}_{ik}$. The following regression is performed for each t:

$$R_{it} - R_f = \hat{\alpha} + \sum_{k=1}^{K} \hat{\beta}_{ik} \hat{\lambda}_{kt} + \zeta_{it}$$

The first step is not necessary for factor models based on explicit microeconomic factors, where the sensitivity is an observed variable. In the case of implicit factor models, the sensitivity is one of the results calculated by the ACP.

In the equation above, $\hat{\alpha}$ is an estimation of the excess return coming from the manager's skill and $\hat{\lambda}_{kt}$ is an estimation of the risk premium associated to the kth risk factor at time t. The $\hat{\lambda}_{kt}$ allows a calculation the average risk premium:

$$\lambda_k = \frac{1}{T} \sum_{t=1}^{T} \lambda_{kt}$$

If the value of λ_k is significantly positive, the factor is kept as a rewarding factor. If the value of λ_k is not significantly different from zero, the factor is discarded. The two step analysis is carried out again with the remaining factors.

When the list of factors is established and the risk premium calculated, the fund performance is given by:

$$\alpha_i = \overline{R}_i - \overline{R}_f - \sum_{k=1}^{K} \hat{\beta}_{ik} \lambda_k$$

The APT-based performance measure was formulated by Connor and Korajczyk (1986). It should be noted that the estimation procedure of factor models contains some difficulties. There are several methods for estimating the factor sensitivities of individual securities and several portfolio-formation procedures that use the estimated factor loadings and idiosyncratic variances. In addition, there are important data-analytic choices including the number of securities to include in the first-stage estimation as well as the periodicity of data appropriate for estimating the factor loadings. Lehmann and Modest (1986) examined whether different methods for constructing reference portfolios lead to different conclusions about the relative performance of mutual funds and showed that alternative APT implementations often suggested substantially different absolute and relative

8. Factor models: more precise methods for evaluating alphas

mutual fund rankings. The fund ranking based on alpha is very sensitive to the method used to construct the APT benchmark.

8.5. Multi-index models

8.5.1. Elton, Gruber, Das and Hlavka's model (1993)

The Elton, Gruber, Das and Hlavka model is a three-index model that was developed in response to a study by Ippolito (1989) which shows that performance evaluated in comparison with an index that badly represents the diversity of the assets in the fund can give a biased result. Their model is presented in the following form:

$$R_{Pt} - R_{Ft} = \alpha_P + \beta_{PL} (R_{Lt} - R_{Ft})$$
$$+ \beta_{PS} (R_{St} - R_{Ft}) + \beta_{PB} (R_{Bt} - R_{Ft}) + \varepsilon_{Pt}$$

where:

R_{Lt} denotes the return on the index that represents large-cap securities;
R_{St} denotes the return on the index that represents small-cap securities;
R_{Bt} denotes the return on a bond index;
ε_{Pt} denotes the residual portfolio return that is not explained by the model.

This model is a generalisation of the single index model. It uses indices quoted on the markets, specialised by asset type. The use of several indices therefore gives a better description of the different types of assets contained in a fund, such as stocks or bonds, but also, at a more detailed level, the large or small market capitalisation securities and the assets from different countries. The multi-index model is simple to use because the factors are known and easily available.

8.5.2. Sharpe's (1992) style analysis model

The theory developed by Sharpe stipulates that a manager's investment style can be determined by comparing the returns on his portfolio with those of a certain number of selected indices. Intuitively, the simplest technique for identifying the style of a portfolio involves successively comparing his returns to those of the different style indices. The goodness of fit between the portfolio returns and the returns on the index is measured with the help of a quantity called R^2 which measures the proportion of variance explained by the model. If the value of R^2 is high, the proportion of unexplained variance is minimal. The index for which the R^2 is highest is therefore the one that best characterises the style of the portfolio. But managers rarely have a pure style, hence Sharpe's idea to propose a method that would enable us to find the combination of style indices which gives the highest R^2 with the returns on the portfolio being studied.

The Sharpe model is a generalisation of the multifactor models, where the factors are asset classes. Sharpe presents his model with twelve asset classes. These asset classes include several categories of domestic stocks, i.e. American in the case of the model: value stocks, growth stocks, large-cap stocks, mid-cap stocks and small-cap stocks. They also include one category for European stocks and one category for Japanese stocks, along with several major bond categories. Each of these classes, in a broad sense, corresponds to a management style and is represented by a specialised index.

The model is written as follows:

$$R_{it} = \sum_{k=1}^{K} b_{ik} F_{kt} + \varepsilon_{it}$$

where:

F_{kt} denotes the return on index k;
b_{ik} denotes the sensitivity of the portfolio to index k and is interpreted as the weighting of class k in the portfolio;
ε_{it} represents the portfolio's residual return term for period t.

Unlike ordinary multifactor models, where the values of the coefficients can be arbitrary, they represent here the distribution of the different asset groups in the portfolio, without the possibility of short selling, and must therefore respect the following constraints:

8. Factor models: more precise methods for evaluating alphas

$$0 \leq b_{ik} \leq 1$$

and:

$$\sum_{k=1}^{K} b_{ik} = 1$$

These constraints enable us to interpret the coefficients as weightings. These weightings are determined by a quadratic program, which consists of minimising the variance of the portfolio's residual return. A customised benchmark, fitted to the portfolio style, is then constructed by taking the weighted linear combination of the various asset classes. Once the benchmark has been constructed for a representative period, the manager's performance is calculated as being the difference between the return on his portfolio and the return on the benchmark. We thereby isolate the share of performance that comes from asset allocation and is explained by the benchmark. The residual share of performance not explained by the benchmark constitutes the management's value-added and comes from the stock picking, within each category, that is different from that of the benchmark. It is the manager's active return. The proportion of the variance not explained by the model, i.e. the

quantity $\quad 1 - R^2 = \dfrac{var(\varepsilon_{it})}{var(R_{it})}\quad$, measures the

importance of stock picking quantitatively.

The Sharpe model uses an analysis that is called return-based, i.e. based solely on the returns. The advantage of this method is that it is simple to implement. It does not require any particular knowledge about the composition of the portfolio. The information on the style is obtained simply by analysing the monthly or quarterly returns of the portfolio through multiple regression. But the major disadvantage of this method lies in the fact that it is based on the past composition of the portfolio and does not therefore allow us to correctly evaluate the modifications in style to which it may have been subjected during the evaluation period. Another possibility for analysing portfolio style consists in using a portfolio-based analysis, based on

portfolio characteristics and which consists in analysing each of the securities that make up the portfolio. The securities are studied and ranked according to the different characteristics that allow their style to be described. The results are then aggregated at the portfolio level to obtain the style of the portfolio as a whole. This method therefore requires the present and historical composition of the portfolio, together with the weightings of the different securities that it contains, to be known with precision (cf. Daniel, Grinblatt, Titman and Wermers, 1997). As an up-to-date composition of funds is not often available, this second method is more difficult to use and Sharpe's method remains the most used.

It is tempting to interpret the "skill" or total excess return ε_{it} in style analysis as an abnormal return measure. There are however two important drawbacks to this. First, introducing the constraints on the factor weightings (they must be positive and sum up to one) into style analysis distorts the results of the standard regression. As a result, the standard properties desirable in linear regression models are not respected. In particular, the correlation between the error term and the benchmark can be non-null (Deroon, Nijman, ter Horst, 2000). Moreover, an analysis of that sort does not provide an explanation for the abnormal return on a risk-adjusted basis. In order to bring a solution to this problem, it is possible to use a multi-index model, where the market indices are used as factors. This model is written in the following way (cf. Amenc, Curtis and Martellini, 2003):

$$R_{it} - R_{ft} = \alpha_i + \sum_{k=1}^{K} \beta_{ik}(F_{kt} - R_{ft}) + \zeta_{it}$$

This factor model generalises the CAPM Security Market Line. It is in the same vein as the one used by Elton et al. (1993) to evaluate managers' fund performance. This equation can be seen as a weak form of style analysis consisting of relaxing coefficient constraints and including a constant term in the regression. Excess returns are used. From a practical point of view, this

8. Factor models: more precise methods for evaluating alphas

approach enables one to consider benchmark construction and performance measurement in a unified setting: once the suited indices have been selected, they can be used both for returns-based style analysis (strong form of style analysis with constraints on coefficients) and for measuring portfolio abnormal return (weak form of style analysis applied to returns in excess of risk-free rate). The performance is then given by the following formula:

$$\alpha_i = \overline{R_i} - R_f - \sum_{k=1}^{K} \beta_{ik} \lambda_k$$

9. Performance persistence

The question of performance persistence in funds is often addressed in two ways. The first is linked to the notion of market efficiency. If we admit that markets are efficient, the stability of fund performance cannot be guaranteed over time. Nevertheless, according to MacKinlay and Lo (1998), the validity of the random market theory is now being called into question, with studies showing that weekly returns are, to a certain extent, predictable for stocks quoted in the United States[12]. This type of affirmation is, however, contested by other university research, which continues to promote the theory of market efficiency, according to which prices take all available information into account, and as a result of which active portfolio management cannot create added value.

The second part of the problem posed by the existence or non-existence of performance persistence is intended to be less theoretical or axiomatic and more pragmatic: Are the winners always the same? Are certain managers more skilful than others? Of course, if certain managers beat the market regularly, over a statistically significant period, they will prove de facto that active investment makes sense and cast doubt over the market efficiency paradigm. But that is not the purpose of the question. A manager who beats the market regularly by taking advantage of arbitrage opportunities from very temporary inefficiencies will not prove that the market is inefficient over a long period.

Professionals speak more willingly of checking whether an investment performance is the fruit of the real skill of the manager, and not just luck, rather than showing that the markets in which they invest are inefficient. In practice, one is often tempted to believe that a manager who has performed well one year is more likely to perform well the following year than a manager who has performed poorly. The publication of fund rankings by the financial press is based on that idea. But the results of studies that tend to verify this assumption are contradictory and do not allow us to affirm that past performance is a good indicator of future performance. The results

depend on the period studied, but generally it would seem that the poorest performances have more of a tendency to persist than the best performances. The results are also different depending on whether equity funds or bond funds are involved. The literature describes two phenomena that depend on the length of the period studied. In the long term (three to five years) and the short term (one month or less) we observe a reversal of trends: past losers become winners and vice versa. Over the medium term (six to twelve months), the opposite effect is observed: winners and losers conserve their characteristics over the following periods and in this case there is performance stability.

Empirical studies carried out to study the phenomenon of performance persistence have enabled performance measurement models to be developed and improved.

A large amount of both academic and professional research is devoted to performance persistence in American mutual funds. The results seem to suggest that there is a certain amount of performance persistence, especially for the worst funds. But parts of these studies also suggest that managers who perform consistently better than the market do exist. In what follows we summarise the results of a certain number of studies. Kahn and Rudd (1995) present a fairly thorough study of the subject, in which they also refer to earlier basic research. The earliest observations generally lead to the conclusion that there is no performance persistence, while the most recent articles conclude that a certain amount of performance persistence exists. The authors, for their part, observed slight performance persistence for bond funds, but not for equity funds. Their study takes into account style effects, management fees and database errors. They conclude that it is more profitable to invest in index funds than in funds that have performed well in the past.

Among the studies that concluded that there was an absence of manager skill in stock picking, we can cite Jensen (1968) and Gruber

9. Performance persistence

(1996). Carhart (1997) shows that performance persistence in mutual funds is not a reflection of the manager's superior stock picking skills. Instead, the common asset return factors and the differences in fees and transaction costs explain the predictable character of fund returns. In addition, he observes that the ranking of funds from one year to another is random. The funds at the top of the rankings one year may perhaps have a slightly greater chance of remaining there than the others. In the same way, the worst ranked funds are very likely to be badly placed again or even disappear. However, the ranking can vary greatly from one year to the next and the winning funds of one year could be the losing funds of the following year and vice versa.

Other studies brought to light persistence in the performance of mutual funds. This is the case of Hendricks, Patel and Zeckhauser (1993), who highlighted a phenomenon of performance persistence for both good managers and bad managers. Malkiel (1995) observed significant performance persistence for good managers in the 1970s, but no consistency in fund returns in the 1980s. His results also suggest that one should invest in funds that have performed best in the past. These funds perform better than the average funds over certain periods, and their performance is not worse than that of the average funds for other periods. However, he qualifies his results slightly with several remarks: the results obtained are not robust, the returns calculated must be reduced by the amount of the fees and the survivorship bias must be taken into account. In addition, the performance of the funds for the period studied is worse than that of the reference portfolios over the same period, both before and after deducting management fees. He also analyses fund fees to determine whether high fees result in better performance. The study finds no relationship between the amount of fees and the value of returns before those fees are deducted. He also concludes, like Kahn and Rudd (1995), that it can be much more profitable for investors to buy index funds with reduced fees, rather than trying to select an active fund manager who seems to be particularly skilful.

Carhart (1997) observed performance persistence for managers whose performance was negative.

Brown, Goetzmann, Ibbotson and Ross (1992) showed that short-term performance persisted, but that the survivorship bias attached to the database (i.e. the fact that funds that perform badly tend to disappear) could significantly affect the results of performance studies and could in particular give an appearance of significant persistence. Malkiel (1995) and Carhart (1997) also show that the persistence they identified could be attributed either to survivorship bias or to a poor choice of benchmark. Malkiel (1995) observes that around 3% of mutual funds disappear every year. As a result, performance statistics in the long run do not contain the results of the bad funds that have disappeared. So the survivorship bias is much more important than previous studies suggested. More recent studies have thus used databases that are corrected for survivorship bias. Malkiel therefore concludes that the investment strategy must not be based on a belief in return persistence over the long-term. A study by Lenormand-Touchais (1998), carried out on French equity mutual funds for the period from January 1 1990 to December 31 1995, shows that there is no long-term performance persistence, unless a slight persistence in negative performance is counted. In the short term, on the other hand, a certain amount of performance persistence can be observed, which is more significant when the performance measurement technique used integrates a risk criterion.

Jegadeesh and Titman (1993) show, with NYSE and AMEX securities over the period 1965-1989, that a momentum strategy that consists of buying the winners from the previous six months, i.e. the assets at the top of the rankings, and selling the losers from the previous six months, i.e. the assets at the bottom of the rankings, earns around 1% per month over the following six months. This shows that asset returns exhibit momentum, which means that the winners of the past continue to perform well and the losers of the past continue to perform badly. Rouwenhorst

9. Performance persistence

(1998) obtains similar results with a sample of 12 European countries for the period 1980-1995.

Although the earliest studies were only based on performance measures drawn from the CAPM, such as Jensen's alpha, the more recent studies used models that took factors other than market factors into account. These factors are size, book-to-market ratio and momentum. Fama and French are responsible for the model that uses three factors (market factor, size and book-to-market ratio). In an article from 1996, Fama and French stress that their model does not explain the short-term persistence of returns highlighted by Jegadeesh and Titman (1993) and suggest that research could be directed towards a model integrating an additional risk factor. It was Carhart (1997) who introduced momentum, which allows short-term performance persistence to be measured, as an additional factor. He suggests that the "hot hands" phenomenon (i.e. a manager's ability to pick the best performing stocks) is principally due to the momentum effect over one year described by Jegadeesh and Titman (1993). Using a four-factor model, Daniel, Grinblatt, Titman and Wermers (1997) studied fund performance to see whether the manager's stock picking skill compensated for the management fees. The authors conclude that performance persistence in funds is due to the use of momentum strategies by the fund managers, rather than the managers being particularly skilful at picking winning stocks.

Brown and Goetzmann (1995) studied performance persistence for equity funds. Their results indicate that relative (i.e. measured in relation to a benchmark) risk-adjusted performance persists. Poor performance also tends to increase the probability that the fund will disappear. Blake and Timmermann (1998) analysed the performance of mutual funds in the United Kingdom, underlining the fact that most performance studies concern American funds and that there are very few on European funds. As it happens, the "equity" mutual fund management industry in the United Kingdom is very advanced and is the one in Europe for which

we have the most historical data. The study shows that equity funds perform slightly worse than the market on a risk-adjusted basis. Performance seems to persist to the extent that, on average, a portfolio made up of funds that have performed best in historical terms will perform better in the following period than a portfolio made up of funds that have performed worst in historical terms.

Elton, Gruber and Blake (1996) confirmed the hot hands result previously described by Hendricks, Patel and Zeckhauser – that high return can predict high return in the short run. However, using risk-adjusted returns to rank funds, they found that past performance is predictive of future risk-adjusted performance in both the short term and long term. Moreover, they found that there is still predictability even after the major impacts of expenses have been removed.

Jan and Hung (2004) found that short-run mutual fund performance is likely to persist in the long run. Subsequent-year performance is predicted not only by past short-run performance, but also by past long-run performance. Their study reveals that in the subsequent year the best funds significantly outperformed the worst funds. Moreover, funds with strong both short- and long-run performance significantly outperform funds with weak both short- and long-run performance. According to them, mutual fund investors can likely benefit from selecting funds on the basis of not only past short-run performance but also past long-run performance.

Bollen and Busse (2005) considered persistence in mutual fund performance on a short-term horizon. Observing that superior performance is short-lived, they suggest that a short measurement horizon provides a more precise method of identifying top performers. So they propose to use three-month measurement periods with daily returns. They not only investigate performance persistence in stock selection but also in market timing strategy, which is new compared to previous studies. They found that

9. Performance persistence

the top decile of funds generates a statistically significant abnormal return in the post-ranking quarter. Increasing the length of time over which they measure risk-adjusted returns, they found that the abnormal return of the top decile disappears. They also observed that the superiority of the top decile over the bottom decile is more pronounced when they used risk-adjusted returns rather than raw returns. They thus concluded that superior performance appears to be a short-lived phenomenon that is not detectable using annual measurement windows. They also notice that, although their findings are statistically significant and robust to a battery of diagnostic tests, the economic significance of persistence in mutual fund abnormal returns is questionable. After taking into account transaction costs and taxes, investors may generate superior returns by following a naïve buy-and-hold approach rather than a performance-chasing strategy, even if short-term performance is predictable.

The different results observed for performance persistence according to the periods studied can be linked to the fact that more market trends, such as seasonal effects and day of the week effects, have been observed in recent years. However, if performance persistence exists in the short term, it is seldom seen over the long term and, as most studies stress, only performance persistence that is observed over a number of years would really allow us to conclude that it is statistically significant. In the absence of a period that is sufficiently long, it is not possible to distinguish luck from skill.

Finally, the studies that seek to check whether it is possible for the manager to add value within the framework of an efficient market were carried out on funds that were invested in a single asset class, generally equities or bonds. While the contribution of stock picking to performance in an efficient market is questionable, the same cannot be said for the contribution of asset allocation to performance. All the studies conclude that asset allocation is important in building performance and often the question of persistence cannot be separated from the asset allocation choices.

Moreover, we can observe that stock markets are subject to cycles. Therefore, certain investment styles produce better performances during certain periods, and worse performances during others. The existence of these cycles can thus explain the performance of a specialised manager persisting over a certain period, if the cycle is favourable, and then suffering from a reversal in the trend when the cycle becomes unfavourable.

9. Performance persistence

The table below summarises the results from the main studies presented in this section.

Authors	Type of data/Period/Models	Results
Jensen (1968)	1945 to 1964 115 mutual funds	No evidence of performance persistence.
Brown, Goetzmann, Ibbotson, Ross (1992)	1976 to 1987 Investigation of the survivorship bias problem.	Short-term performance persistence. The survivorship bias attached to the database could significantly affect the result of performance studies and could in particular give an appearance of significant persistence.
Hendricks, Patel, Zeckhauser (1993)	1974 to 1988 165 of Wiesenberger's equity mutual funds.	Performance persistence for both good and bad managers.
Jegadeesh, Titman (1993)	1965 to1989 Funds made up of NYSE and AMEX securities. Three-factor model (the momentum factor is not included in the model).	Performance persistence for both good and bad managers. Assets returns exhibit momentum: the winners of the past continue to perform well and the losers of the past continue to perform badly. Performance persistence is due to the use of momentum strategies.
Brown, Goetzmann (1995)	1976 to 1988 Wiesenberger's equity mutual funds. Sample free of survivorship bias.	Performance persistence for equity funds on a risk-adjusted basis. Poor performance tends to increase the probability that the fund will disappear.
Kahn, Rudd (1995)	1983 to 1993 for the equity funds. 1988 to 1993 for the bond funds.	Slight performance persistence for bond funds, but not for equity funds. The analysis takes into account style effects, management fees and database errors.
Malkiel (1995)	1971 to 1991 Analysis of fund fees. Study of the survivorship bias.	Significant performance persistence for good managers in the 1970s, but no consistency in fund returns in the 1980s. No long-term persistence. The persistence identified could be due to survivorship bias.
Fama, French (1996)	1963 to 1993 NYSE, AMEX and NASDAQ stocks. Three-factor model (market factor, size and book-to-market ratio).	Their model does not explain the short-term persistence of returns highlighted by Jegadeesh and Titman (1993). Suggest that research could be directed towards a model integrating an additional risk factor.
Gruber (1996)	1985 to 1994 270 of Wiesenberger's equity mutual funds. Sample free from survivorship bias. Single index and four index model.	Evidence of persistence in performance.

9. Performance persistence

Elton, Gruber and Blake (1996)	1977-1993 188 of Wiesenberger's "common stock" funds Model including the three factors of Fama and French plus an index to account for growth versus value.	High return can predict high return in the short run. Past performance is predictive of future risk-adjusted performance, in both the short run and long run.
Carhart (1997)	1962 to 1993 Equity funds made up of NYSE, AMEX and NASDAQ stocks. Free from survivorship bias. Four-factor model (Fama and French's three-factor model with momentum as additional factor).	Performance persistence for bad managers. Short-term performance persistence is due to the use of momentum strategies. Ranking of fund from one year to another is random.
Daniel, Grinblatt, Titman, Wermers (1997)	1975 to 1994 2500 equity funds made up of stocks from NYSE, AMEX and NASDAQ. Four-factor model. Study of management fees.	Performance persistence is due to the use of momentum strategies, rather than the managers being particularly skilful at picking winning stocks.
Blake, Timmermann (1998)	1972 to 1995 Mutual funds in the United Kingdom. Three-factor model.	Performance persistence for equity funds: on average, a portfolio made up of funds that have performed best in historical terms will perform better in the following period than a portfolio made up of funds that have performed worst in historical terms.
Lenormand-Touchais (1998)	1990 to 1995 French equity mutual funds.	Short-term performance persistence, more significant when the performance measurement technique used integrates a risk criterion. No long-term performance persistence, unless a slight persistence in negative performance is counted.
Rouwenhorst (1998)	1980 to 1995 A sample of funds from 12 European countries.	Performance persistence for both good and bad managers. Asset returns exhibit momentum.
Jan and Hung (2004)	January 1961 to June 2000 3316 Equity funds from the CRSP Survivor-Bias Free US Mutual Fund Database. Carhart's four-factor model.	Short-term mutual fund performance is likely to persist in the long run.
Bollen and Busse (2005)	1985 to 1995 230 of Wiesenberger's "common stock" mutual funds with a "maximum capital gain", "growth" or "growth and income" investment objective. Carhart's four-factor model.	Superior performance is a short-lived phenomenon that is observable only when funds are evaluated several times a year.

9. Performance persistence

These performance persistence studies do not give very conclusive results as to whether persistence really exists. Over a long period, there is a greater tendency to observe under-performance persistence on the part of poor managers than over-performance persistence from good managers. However, the studies do not take the investment style followed by the managers into account. We do, nevertheless, observe that different investment styles are not all simultaneously favoured by the market. Markets are subject to economic cycles and a style that is favourable for one period, i.e. which offers a performance that is better than that of the market, can be less favourable over another period and lead to under-performance compared to the market. This can be measured by comparing the returns of the different style indices with the returns of a broad market index. The fact that an investment style performs well or badly should not be confused with the manager's skill in picking the right stocks within the style that he has chosen. As we mentioned a little earlier, a manager's skill in practising a well-defined style should be evaluated in comparison with a benchmark that is adapted to that style.

Few studies have addressed the subject of performance persistence for managers who specialise in a specific style. The results of the studies that have been performed are contradictory and do not allow us to conclude that persistence exists. For example, Coggin, Fabozzi and Rahman (1993) carried out a study on American pension funds over a period from 1983 to 1990. Their study relates to identification of both the market timing effect and the selectivity effect. They used two broad indices: the S&P 500 index and the Russell 3000 index, and four specialised indices: the Russell 1000 index for large-cap stocks, the Russell 2000 index for small-cap stocks, a Russell index specialised in value stocks and a Russell index specialised in growth stocks. They showed that the timing effect and the selectivity effect were both sensitive to the choice of benchmark and the period of the study. They found a positive selectivity effect compared to the specific indices, while that effect was negative compared to the broad indices. However, they found a negative market timing effect in both cases. The study shows, therefore, that specialisation is a source of value-added. Managers succeed in performing better than their reference style index, even if they do not manage to beat the market as a whole. Over the period studied, the different style indices did not all perform in line with the market. The performance of the value stock index was approximately equal to that of the market, which implies that the study period was favourable for value stocks. The performance of the growth stock index was slightly worse. As far as the small-cap stock index was concerned, its performance was half as good as that of the market index.

However, Kahn and Rudd (1995, 1997) concluded that fund performance was not persistent for a sample of 300 US funds over a period from October 1988 to October 1995.

Another interesting study is that of Chan, Chen and Lakonishok (1999). This study concerns Morningstar funds. The study shows that on the whole there is a certain consistency in the style of the funds. Nevertheless, funds that have performed badly in the past are more liable to modify their style than others. This study shows that it is preferable to avoid managers who change style regularly. They make it more difficult to optimise a portfolio that is shared between several managers and produce worse performances than managers whose style is consistent.

Finally, Ibbotson and Patel (2002) investigated U.S. domestic equity funds performance persistence after adjusting for the investment style of the funds. They measured the skill of managers against a benchmark that adjusts for the style of the fund. The style adjustment was made by using returns-based style analysis to construct customized benchmarks. Their results indicate that winning funds do repeat good performance.

9. Performance persistence

When fund style is in question, the problem of fund misclassification has to be considered. DiBartolomeo and Witkowski (1997) note that a large proportion of mutual funds are misclassified, rendering performance comparisons inadequate. Mutual fund managers sometimes misclassify their investment strategy in order to show more competitive results. DiBartolomeo and Witkowski find that 40% of mutual funds are misclassified, and 9% seriously so. They cite ambiguity of classification systems and competitive pressures as the major reasons for misclassification. Kim, Shukla and Tomas (2000) agree that a majority of mutual funds are misclassified (one-third seriously misclassified), but they do not find evidence that fund managers are gaming their objectives (i.e., diverging from stated objectives in order to achieve a higher ranking).

Conclusion

Throughout this paper, we have presented the main research available in the area of performance evaluation and developed since the end of the 1950s. We have seen the evolution of performance evaluation from elementary measures of returns to more sophisticated methods that include the various aspects of risk through multifactor models and also take into account the non stationarity of risk through dynamic evaluation.

Selecting an investment manager is a matter of choosing the manager who can produce the best numbers in the future. Arnott and Darnell (2003) underline that the same set of numbers drawn from the past can often present two very different pictures. Changing the benchmark, changing the fiscal year, risk-adjusting the performance can all make a bad product look good or a good product look bad. He concludes that the quest for a single, simple measure of performance often leads to an overly simplistic view of the past, which can lead to poor choices for the future.

Beside the performance measurement itself, we must not forget that the choice of a benchmark for the portfolio to be evaluated and the design of this benchmark are important elements in performance evaluation. Portfolio performance is mostly presented as being relative to a benchmark, even if the portfolio management is said to be benchmark-free. In this specific area, some improvements are still possible, in order to choose the most accurate benchmark to evaluate performance. In particular, we observe that most managers do not give all the attention required to this choice, and often use a market index as benchmark. It is not appropriate to compare portfolio performance to broad market indexes, which usually constitute inefficient investments[13]. It is necessary to derive benchmarks that mimic the portfolio to be evaluated in the best possible way, and specifically benchmarks that take the manager's skill into account. This choice of benchmark defines the level and the kind of risk supported by the portfolio during the investment period and thus its future performance.

For this purpose, Kuenzi (2003) proposes the use of strategy benchmarks. He chooses this term "strategy benchmarks" instead of the more common term "custom benchmarks" to emphasise the fact that these benchmarks are related to a manager's specific strategy and universe of securities. Kuenzi explains that the choice of an inappropriate benchmark may distort the portfolio risk and performance analysis and does not ensure the integrity of performance measures. Kuenzi underlines that while investors are prepared to bear the benchmark risk, managers are supposed to bear the active risk. Consequently, the concept of risk controls becomes distorted if the manager employs a benchmark that is not representative of his portfolio's true neutral weights. Using an inappropriate benchmark makes manager evaluation more difficult.

More attention could also be given to performance persistence evaluation, specifically the persistence of a portfolio manager's skill.

13 - For more details on this subject see N. Amenc, F. Goltz and V. Le Sourd, "Assessing the Quality of Stock Market Indices: Requirements for Asset Allocation and Performance Measurement", EDHEC Risk and Asset Management Research Centre publication, 2006.

Bibliography

• Arnott R. D. and Darnell M., "What's Behind the Numbers?", *Journal of Investing*, vol. 12, n°2, summer 2003.

• Ambarish R., Seigel L., "Time is the essence", *Risk*, vol. 9, n°8, August 1996.

• Amenc N., Curtis S. and Martellini L., "The Alpha and Omega of Hedge Fund Performance Measurement", Working Paper, EDHEC Risk and Asset Management Research Centre, January 2003.

• Amenc N. and Le Sourd V., *Portfolio Theory and Performance Analysis*, Wiley, 2003.

• Bhattacharya S., Pfleiderer P., "A Note on Performance Evaluation", *Technical Report 714*, Graduate School of Business, Stanford University, 1983.

• Black F., "Capital Market Equilibrium with Restricted Borrowing", *Journal of Business*, n°45, July 1972, pp. 444-455.

• Blake D. and Timmermann A., "Mutual Fund Performance: Evidence from the UK", *European Finance Review 2*, 1998, pp. 57-77.

• Bollen N. P. B. and Busse J. A., "Short-Term Persistence in Mutual Fund Performance", *Review of Financial Studies*, vol. 18, n°2, summer 2005.

• Brennan M., "Taxes, Market Valuation and Corporate Financial Policy", *National Tax Journal 25*, 1970, pp. 417-427.

• Brown S. J. and Goetzmann W. N., "Performance Persistence", *Journal of Finance*, vol. 50, n°2, June 1995, pp. 679-698.

• Brown S. J., Goetzmann W., Ibbotson R. G. and Ross S. A., "Survivorship Bias in Performance Studies", *Review of Financial Studies*, vol. 5, 1992, pp. 553-580.

• Campisi S., "The Case for Money-Weighted Performance Attribution", *Journal of Performance Measurement*, vol. 8, n°3, spring 2004.

• Cantaluppi L. and Hug R., "Efficiency Ratio: A New Methodology for Performance Measurement", *Journal of Investing*, vol. 9, n°2, summer 2000.

• Carhart M. M., "On Persistence in Mutual Fund Performance", *Journal of Finance*, vol. 52, n°1, March 1997, pp. 57-82.

• Chan L. K. C., Chen H.-L. and Lakonishok J., "On Mutual Fund Investment Styles", Working Paper n°7215, National Bureau of Economic Research, 1999.

• Chestopalov I. and Beliaev S., "A Simplified Method for Calculating the Money-Weighted Rate of Return", *Journal of Performance Measurement*, vol. 9, n°2, winter 2004/2005.

• Christopherson J. A., Ferson W. E. and Turner A. L., "Performance Evaluation Using Conditional Alphas and Betas", *Journal of Portfolio Management*, vol. 26, n°1, fall 1999, pp.59-72.

Bibliography

• Coggin T. D., Fabozzi F. J. and Rahman S., "The Investment Performance of U.S. Equity Pension Fund Managers: An Empirical Investigation", *Journal of Finance*, vol. 48, n°3, July 1993, pp. 1039-1055.

• Cohen R. B., Coval J. D. and Pastor L., "Judging Fund Managers by the Company they Keep", *Journal of Finance*, vol. 60, n°3, June 2005.

• Connor G. and Korajczyk R., "Performance Measurement with the Arbitrage Pricing Theory: A New Framework for Analysis", *Journal of Financial Economics*, vol. 15, 1986, pp. 374-394.

• Cornell B., "Asymmetric Information and Portfolio Performance Measurement", *Journal of Financial Economics*, vol. 7(4), December 1979, pp. 381-390.

• Cvitanic J., Lazrak A. and Wang T., "Sharpe Ratio as a Performance Measure in a Multi-Period Setting", Working Paper, July 2004.

• Daniel K., Grinblatt M., Titman S. and Wermers R., "Measuring Mutual Fund Performance with Characteristic-Based Benchmarks", *Journal of Finance*, vol. 52, n°3, July 1997, pp. 1035-1058.

• Darling R. and MacDougall A., "Using Performance Statistics: Have the Measurers Lost the Plot?", WM Company, Working Paper, September 2002.

• Dembo R. S., "Value-at-Risk and Return", *Net exposure: The Electronic Journal of Financial Risk*, vol. 1, October 1997, pp. 1-11.

• DeRoon F., Nijman T. and ter Horst J., "Evaluating Style Analysis", Working Paper, Quantitative Investment Research Europe, 2000.

• DiBartolomeo D., "Just Because We Can Doesn't Mean We Should – Why Daily Observation Frequency in Performance Attribution is Not Better", *Journal of Performance Measurement*, vol. 7, n°3, spring 2003.

• DiBartolomeo D. and Witkowski E., "Mutual Fund Misclassification: Evidence Based on Style Analysis", *Financial Analysts Journal*, vol. 53, September/October 1997, pp. 32-43.

• Dimson E. and Jackson A., "High-Frequency Performance Monitoring", *Journal of Portfolio Management*, vol. 28, n°1, fall 2001.

• Dowd K., "Adjusting for Risk: An Improved Sharpe Ratio", *International Review of Economics and Finance*, vol. 9, n°3, 2000, pp. 209-222.

• Elton E. J. and Gruber M. J., *Modern Portfolio Theory and Investment Analysis*, 5th ed., Wiley, 1995.

• Elton E. J., Gruber M. J. and Blake C. R., "The Persistence of Risk-Adjusted Mutual Fund Performance", *Journal of Business*, vol. 69, n°2, 1996.

• Elton E. J., Gruber M. J., Das S. and Hlavka M., "Efficiency with Costly Information: A Reinterpretation of Evidence from Managed Portfolios", *Review of Financial Studies*, vol. 6, n°1, 1993, pp. 1-22.

Bibliography

• Fama E. F., "Efficient Capital Markets: A Review of Theory and Empirical Work", *Journal of Finance*, vol. 25, n°2, March 1970, pp. 383-417.

• Fama E. F. and French K. R., "Multifactor Explanations of Asset Pricing Anomalies", *Journal of Finance*, vol. 51, n°1, March 1996, pp. 55-81.

• Ferson W. E. and Schadt R. W., "Measuring Fund Strategy and Performance in Changing Economic Conditions", *Journal of Finance*, vol. 51, n°2, June 1996, pp. 425-461.

• Filbeck G. and Tompkins D. L., "Management Tenure and Risk-Adjusted Performance of Mutual Funds", *Journal of Investing*, vol. 13, n°2, summer 2004.

• Goetzmann W., N., Ingersoll Jr. J. and Ivkovic Z., "Monthly Measurement of Daily Timers", *Journal of Financial and Quantitative Analysis*, vol. 35, n°3, September 2000.

• Graham J. R. and Harvey C. R., "Grading the Performance Market-Timing Newsletters", *Financial Analysts Journal*, November/December 1997.

• Gray Jr. K. B. and Dewar R. B. K., "Axiomatic Characterization of the Time-Weighted Rate of Return", *Management Science*, vol. 18, n°2, October 1971.

• Gressis N., Philippatos G. C. and Vlahos G., "Net Selectivity as a Component Measure of Investment Performance", *Financial Review*, vol. 21, n°1, 1986.

• Grinblatt M. and Titman S., "Mutual Fund Performance: an Analysis of Quarterly Portfolio Holdings", *Journal of Business*, vol. 62 n°3, 1989 a, pp. 393-416.

• Grinblatt M. and Titman S., "Portfolio Performance Evaluation: Old Issues and New Insights", *Review of Financial Studies*, vol. 2, 1989 b, pp. 393-421.

• Grinblatt M. and Titman S., "Performance Measurement without Benchmarks: an Examination of Mutual Fund Returns", *Journal of Business*, vol. 66, n°1, 1993, pp. 47-68.

• Gruber M. J., "Another Puzzle: The Growth in Actively Managed Mutual Funds", *Journal of Finance*, vol. 51, 1996, pp. 783-810.

• Hendricks D., Patel J. and Zeckhauser R., "Hot Hands in Mutual Funds: Short-Run Persistence of Relative Performance, 1974-1988", *Journal of Finance*, vol. 48, March 1993, pp. 93-130.

• Henriksson R. D., "Market Timing and Mutual Fund Performance: an Empirical Investigation", *Journal of Business*, vol. 57, n°1, 1984, pp. 73-96.

• Henriksson R. D. and Merton R. C., "On Market Timing and Investment Performance II: Statistical Procedures for Evaluating Forecasting Skills", *Journal of Business*, vol. 54, n°4, 1981, pp. 513-533.

• Hwang S. and Satchell S., "Evaluation of Mutual Fund Performance in Emerging Markets", *Emerging Markets Quarterly*, vol. 2, n°3, 1998, pp. 39-50.

Bibliography

• Ibbotson R. G. and Patel A. K., "Do Winners Repeat with Style?", Working Paper, February 2002.

• Illmer S. and Marty W., "Decomposing the Money-Weighted Rate of Return", *Journal of Performance Measurement*, vol. 7, n°4, summer 2003.

• Ippolito R., "Efficiency with Costly Information: a Study of Mutual Fund Performance, 1965-84", *Quarterly Journal of Economics*, vol. 104, 1989, pp. 1-23.

• Israelsen C. L., "A Refinement to the Sharpe Ratio and Information Ratio", *Journal of Asset Management*, vol. 5, n°6, April 2005.

• Jacquier E., Kane A. and Marcus A. J., "Geometric or Arithmetic Mean: A Reconsideration", *Financial Analysts Journal*, vol. 59, n°6, November/December 2003.

• Jan Y.-C. and Hung M.-W., "Short-Run and Long-Run Persistence in Mutual Funds", *Journal of Investing*, vol. 13, n°1, spring 2004.

• Jegadeesh N. and Titman S., "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency", *Journal of Finance*, vol. 48, 1993, pp. 65-91.

• Jensen M. C., "The Performance of Mutual Funds in the Period 1945-1964", *Journal of Finance*, vol. 23, May 1968, pp. 389-419.

• Jensen M., "Optimal Utilization of Market Forecasts and the Evaluation of Investment Performance", in *Mathematical Methods in Investment and Finance*, G. P. Szego and K. Shell eds., Elsevier, 1972.

• Kahn R. N. and Rudd A., "Does Historical Performance Predict Future Performance?" *Barra Newsletter*, spring 1995.

• Kahn R. N. and Rudd A., "The Persistence of Equity Style Performance: Evidence from Mutual Fund Data", in *The Handbook of Equity Style Management*, T. Daniel Coggin, Frank J. Fabozzi, Robert D. Arnott, eds., 2nd ed., Frank J. Fabozzi Associates, 1997.

• Keating C. and Shadwick W.F., "A Universal Performance Measure", *Journal of Performance Measurement*, vol. 6, n°3, 2002.

• Kim M., Shukla R. and Tomas M., "Mutual Fund Objective Misclassification", *Journal of Economics and Business*, vol. 52, July/August 2000, pp. 309-323.

• Kuenzi D.E., "Strategy Benchmarks", *Journal of Portfolio Management*, vol. 29, n°2, winter 2003, pp. 46-56.

• Lehmann B. and Modest D., "Mutual Fund Performance Evaluation: A Comparison of Benchmarks and Benchmark Comparisons", *Journal of Finance*, vol. 21, 1987, pp. 233-265.

• Leland H. E., "Beyond Mean-Variance: Performance Measurement in a Nonsymmetrical World", *Financial Analysts Journal*, January/February 1999.

Bibliography

• Lenormand-Touchais G., "Etude de la stabilité des performances: le cas des Sicav actions françaises", *Banque et Marchés* n°36, September-October 1998.

• Lo A. W., "The Statistics of Sharpe ratios", *Financial Analysts Journal*, vol. 56, 2002, pp. 36-52.

• Lobosco A., "Style/Risk-Adjusted Performance", *Journal of Portfolio Management*, vol. 25, n°3, spring 1999, pp.65-68.

• McDonald J., "French Mutual Fund Performance: Evaluation of Internationally Diversified Portfolios", *Journal of Finance*, vol. 28, n°5, 1973, pp. 1161-1180.

• Malkiel B. G., "Returns from Investing in Equity Mutual Funds 1971 to 1991", *Journal of Finance*, vol. 50, n°2, June 1995, pp. 549-572.

• Marsh P., "Fund Managers and Quarterly Performance Measurement", in *Short Termism on Trial*, London: IFMA, 1991.

• Melnikoff M., "Investment Performance Analysis for Investors", *Journal of Portfolio Management*, vol. 25, n°1, fall 1998, pp. 95-107.

• Merton R. C., "On Market Timing and Investment Performance I: an Equilibrium Theory of Value for Market Forecasts", *Journal of Business*, vol. 54, n°3, 1981.

• Modigliani F. and Modigliani L., "Risk-Adjusted Performance", *Journal of Portfolio Management*, winter 1997, pp.45-54.

• Muralidhar A. S., "Risk-Adjusted Performance: The Correlation Correction", *Financial Analysts Journal*, vol. 56, n°5, September-October 2000, pp.63-71

• Muralidhar A. S., "Optimal Risk-Adjusted Portfolios with Multiple Managers", *Journal of Portfolio Management*, vol. 27, n°3, spring 2001, pp. 97-104.

• Muralidhar A. S., "Skill, History and Risk-Adjusted Performance", *Journal of Performance Measurement*, winter 2001-2002.

• Plantinga A. and de Groot S., "Risk-Adjusted Performance Measures and Implied Risk-Attitudes", *Journal of Performance Measurement*, vol. 6, n°2, winter 2001/2002, pp. 9-19.

• Pogue G., Solnik B. and Rousselin A., "International Diversification: A Study of the French Mutual Funds", MIT, Working Paper, 1974.

• Roll R., "A Critique of the Asset Pricing Theory's Tests", *Journal of Financial Economics*, March 1977, pp. 129-176.

• Rouwenhorst K. G., "International Momentum Strategies", *Journal of Finance*, vol. 53, n°1, February 1998, pp. 267-284.

Bibliography

• Scholtz H. and Wilkens M., "Investor Specific Performance Measurement – A Justification of Sharpe Ratio and Treynor Ratio", Working Paper, 2004.

• Scholtz H. and Wilkens M., "A Jigsaw Puzzle of Basic Risk-adjusted Performance Measures", *Journal of Performance Measurement*, spring 2005.

• Sharpe W. F., "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk", *Journal of Finance*, vol. 19, September 1964, pp.425-442.

• Sharpe W. F., "Mutual Fund Performance", *Journal of Business*, January 1966, pp. 119-138.

• Sharpe W. F., "Asset Allocation: Management Style and Performance Measurement", *Journal of Portfolio Management*, vol. 18, winter 1992, pp. 7-19.

• Sharpe W. F., "The Sharpe Ratio", *Journal of Portfolio Management*, fall 1994.

• Sortino F.A., Miller G. and Messina J., "Short Term Risk-Adjusted Performance : A Style Based Analysis", *Journal of Investing*, summer 1997.

• Sortino F. A. and Price L. N., "Performance Measurement in a Downside Risk Framework", *Journal of Investing*, vol. 3, n°3, 1994.

• Sortino F. A. and van der Meer R., "Downside Risk", *Journal of Portfolio Management*, 1991, pp. 27-32.

• Sortino F. A., van der Meer R. and Plantinga A., "The Dutch Triangle", *Journal of Portfolio Management*, vol. 18, summer 1999, pp. 50-59.

• Spaulding D., "Is the Modified Dietz Formula Money-Weighted or Time-Weighted?", *Journal of Performance Measurement*, vol. 7, n°3, spring 2003.

• Srivastava S. C. and Essayyad M., "Investigating a New methodology for Ranking International Mutual Funds", *Journal of Economics and Finance*, vol. 18, n°3, fall 1994.

• Treynor J. L., "How to Rate Management of Investment Funds", *Harvard Business Review* 43, January-February 1965, pp. 63-75.

• Treynor J. L. and Black F., "How to Use Security Analysis to Improve Portfolio Selection", *Journal of Business*, vol. 46 n°1, 1973, pp. 61-86.

• Treynor J. and Mazuy K., "Can Mutual Funds Outguess the Market?", *Harvard Business Review* 44, July-August 1966, pp. 131-136.

• Vinod H. D. and Morey M. R., "A 'Double' Sharpe Ratio", Working Paper, 2001.

• Ziemba W. T., "The Symmetric Downside-Risk Sharpe Ratio", Working Paper, April 2005.

About the EDHEC Risk and Asset Management Research Centre

EDHEC is one of the top five business schools in France and was ranked 7th in the Financial Times Masters in Management Rankings 2006 owing to the high quality of its academic staff (over 100 permanent lecturers from France and abroad) and its privileged relationship with professionals that the school has been developing since it was established 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 is one of the few business schools in Europe to have received the triple international accreditation: AACSB (US-Global), Equis (Europe-Global) and AMBA (UK-Global). EDHEC pursues an active research policy in the field of finance. Its **Risk and Asset Management Research Centre** carries out numerous research programmes in the areas of asset allocation and risk management in both the traditional and alternative investment universes.

The choice of asset allocation

The EDHEC Risk and Asset Management Research Centre structures all of its research work around asset allocation. This issue corresponds to a genuine expectation from the market. On the one hand, the prevailing stock market situation in recent years has shown the limitations of active management based solely on stock picking as a source of performance.

Percentage of variation between funds

3.5% Commissions
11% Stock Picking
40% Allocation Stratégique
45.5% Allocation tactique
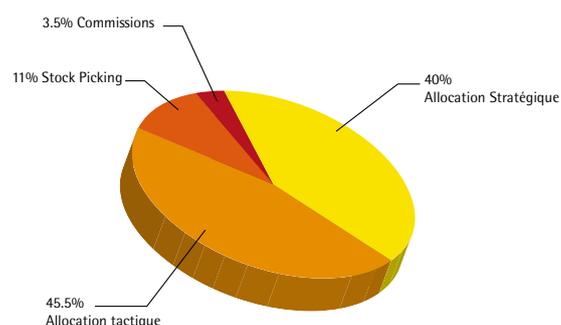
Source: EDHEC (2002) and Ibbotson, Kaplan (2000)

On the other, the appearance of new asset classes (hedge funds, private equity), with risk profiles that are very different from those of the traditional investment universe, constitutes a new opportunity in both conceptual and operational terms. This strategic choice is applied to all of the centre's research programmes, whether they involve proposing new methods of strategic allocation, which integrate the alternative class; measuring the performance of funds while taking the tactical allocation dimension of the alphas into account; taking extreme risks into account in the allocation; or studying the usefulness of derivatives in constructing the portfolio.

An applied research approach

In a desire to ensure that the research it carries out is truly applicable in practice, EDHEC has implemented a dual validation system for the work of the EDHEC Risk and Asset Management Research Centre. All research work must be part of a research programme, the relevance and goals of which have been validated from both an academic and a business viewpoint by the centre's advisory board. This board is made up of both internationally recognised researchers and the centre's business partners. The management of the research programmes respects a rigorous validation process, which guarantees both the scientific quality and the operational usefulness of the programmes.

To date, the centre has implemented six research programmes:

Multi-style/multi-class allocation

This research programme has received the support of Misys Asset Management Systems, SG Asset Management and FIMAT. The research carried out focuses on the benefits, risks and integration methods of the alternative class in asset allocation. From that perspective, EDHEC is making a significant contribution to the research conducted in the area of multi-style/multi-class portfolio construction.

About the EDHEC Risk and Asset Management Research Centre

Performance and style analysis

The scientific goal of the research is to adapt the portfolio performance and style analysis models and methods to tactical allocation. The results of the research carried out by EDHEC thereby allow portfolio alphas to be measured not only for stock picking but also for style timing. This programme is part of a business partnership with the firm EuroPerformance (part of the Fininfo group).

Indices and benchmarking

EDHEC carries out analyses of the quality of indices and the criteria for choosing indices for institutional investors. EDHEC also proposes an original proprietary style index construction methodology for both the traditional and alternative universes. These indices are intended to be a response to the critiques relating to the lack of representativity of the style indices that are available on the market. EDHEC was the first to launch composite hedge fund strategy indices as early as 2003. The indices and benchmarking research programme is supported by AF2I, Euronext, BGI, BNP Paribas Asset Management and UBS Global Asset Management.

Asset allocation and extreme risks

This research programme relates to a significant concern for institutional investors and their managers – that of minimising extreme risks. It notably involves adapting the current tools for measuring extreme risks (VaR) and constructing portfolios (stochastic check) to the issue of the long-term allocation of pension funds. This programme has been designed in co-operation with Inria's Omega laboratory. This research programme also intends to cover other potential sources of extreme risks such as liquidity and operations. The objective is to allow for better measurement and modelling of such risks in order to take them into consideration as part of the portfolio allocation process.

Asset allocation and derivative instruments

This research programme focuses on the usefulness of employing derivative instruments in the area of portfolio construction, whether it involves implementing active portfolio allocation or replicating indices. "Passive" replication of "active" hedge fund indices through portfolios of derivative instruments is a key area in the research carried out by EDHEC. This programme is supported by Eurex and Lyxor.

ALM and asset management

This programme concentrates on the application of recent research in the area of asset-liability management for pension plans and insurance companies. The research centre is working on the idea that improving asset management techniques and particularly strategic allocation techniques has a positive impact on the performance of Asset-Liability Management programmes. The programme includes research on the benefits of alternative investments, such as hedge funds, in long-term portfolio management. Particular attention is given to the institutional context of ALM and notably the integration of the impact of the IFRS standards and the Solvency II directive project. This programme is sponsored by AXA IM.

Research for business

To optimise exchanges between the academic and business worlds, the **EDHEC Risk and Asset Management Research Centre** maintains a website devoted to asset management research for the industry: www.edhec-risk.com, circulates a monthly newsletter to over 75,000 practitioners, conducts regular industry surveys and consultations, and organises annual conferences for the benefit of institutional investors and asset managers. The centre's activities have also given rise to the business offshoots **EDHEC Investment Research** and **EDHEC Asset Management Education**. **EDHEC Investment Research** supports institutional investors and asset managers in the implementation of the centre's research results and proposes asset allocation services in the context of a 'core-satellite' approach encompassing alternative investments.
EDHEC Asset Management Education helps investment professionals to upgrade their skills with advanced risk and asset management training across traditional and alternative classes.

Notes

EDHEC
BUSINESS SCHOOL
Lille - Nice