International Mutual Funds Performance and Persistence

across the Universe of Performance Measures

This version: September 22, 2019

Philippe Cogneau[1]

Georges Hübner[2]

---

[1] Associate Researcher, HEC Liège, Management School of the University of Liège, Belgium; Affiliate Professor, EDHEC Business School, Lille, France; E-mail: philippe.cogneau@uliege.be .

[2] Professor, HEC Liège, Management School of the University of Liège, Belgium; Affiliate Professor, EDHEC Business School, Lille, France E-mail: g.hubner@uliege.be .

Abstract.

We process an exhaustive set of 147 portfolio performance measures and their variations, and identify 18 relevant dimensions using a Principal Component Analysis on a sample of 1,625 international equity mutual funds. We isolate three of the seven most informative factors that uncover potential strong performance persistence. These factors reflect various forms of incremental return and preference-adjusted performance. Our paper is the first one that shows statistical and economic evidence that conditioning portfolio formation on past realizations of these factors may produce significant outperformance, from the point of view of naïve portfolio allocation as well as more classical selection criteria like the Sharpe ratio.

Keywords: performance measurement, persistence analysis, mutual fund, principal component analysis, portfolio optimization.

JEL Classification codes: G10, G11, G17.

Table of Contents

INTRODUCTION

The question of the risk-adjusted performance of actively managed portfolios was explicitly addressed for the first time by Sharpe (1966), through his eponymous ratio derived from the modern portfolio theory of Markowitz (1952). This measure refers to the CAPM equilibrium relationship, in which investors *ex ante* rationally choose the portfolio with the best reward-to-variability ratio. The notion of *ex post* portfolio performance, which is generically a measure of the efficiency of a transformation process, uses risk as an input and value creation as an output. When considering realized portfolio returns, the Sharpe ratio posits that risk can adequately be measured by the standard deviation and the ex *post* value creation of the portfolio by its mean excess return.

Many other measures have been proposed in the academic and practitioner's literature, be they complements (e.g. using alternative types of risks, like Jensen's alpha or the Information ratio), substitutes (e.g. replacing volatility as of measure of total risk, like the Sortino ratio), or totally remote from the Sharpe ratio. This blossoming phenomenon accelerated after the turn of the century. Recent surveys by Cogneau and Hübner (2009a,b) and Caporin *et al.* (2013) enumerate more than 100 measures, whose authors each defend their scientific and/or practical relevance. Furthermore, many of those measures are parametric, so that changing the value of the parameter(s) potentially induces an unlimited number of variations.[i]

With respect to this accumulation of measures, it is necessary to determine whether they are genuinely distinctive, or if they present a certain degree of redundancy. As we discuss in the next section, the literature on the subject demonstrates rather limited ambitions to sort out this universe of measures. Most studies take the Sharpe ratio as benchmark, and primarily aim to determine whether an arbitrary set of measures provide concordant rankings.

In this paper, we investigate the subject much deeper in comparison to previous research. We consider the dimensions of performance measurement from two perspectives. Firstly, we examine without prejudice the performance measures that capture the variations in realized risk-adjusted returns of actively managed portfolios. For this purpose, we use an exhaustive set of measures – 147 in total – covering almost all classes reported in the taxonomy of Cogneau and Hübner (2009a,b)[ii]. Thus, we introduce no *ex ante* bias in the selection of measures. Out of this set, we run an elimination process and perform a principal component analysis to show the main dimensions involved in the performance. To the best of our knowledge this topic has not been explored to date. Through an orthogonal rotation we remain with a reduced set of factors, simplifying its interpretation.

As such, finding out what set of measures span the dimensions of portfolio performance does not provide any insight regarding their relevance for portfolio management. A measure can represent a "dimension", in that it provides a portfolio ranking scheme that is totally remote from any other scheme, but it can be otherwise useless because this dimension bears no relevance for asset management decisions, be it for portfolio selection, manager compensation, or investment advice. This issue of relevance motivates our second, more important, perspective. Out of the set of identified dimensions, we aim to detect whether they display the ability to reflect genuine and sustained skills in active portfolio management through a complete persistence study.

Because the identification of dimensions in performance measurement is meant to be as large as possible, any measure can be considered in the final set irrespective of its popularity. We start with no *a priori* and simply attempt to uncover persistence if it exists. We assess persistence in two ways. First, we perform a classical analysis by analyzing the consistency over time of portfolio rankings using each considered dimension. The issue with such a method, which prevails in almost all existing persistence studies, is that we test the joint hypothesis of (i) the relevance of a given performance measure for portfolio rankings and (ii) the existence of persistence abilities among

managers. Such a joint test has weak power. Therefore, we carry out a second analysis by examining the ability of selected performance measures to generate persistence on other relevant dimensions, including total returns. In doing so, we alleviate the issue of the relevance of a considered performance measure by matching it with all candidates, and the test of pure persistence becomes more powerful.

Specifically, we build quintile portfolios constituted by ordering funds along with the main principal components, and follow these portfolios over 14 years. We emphasize that some of the factors indeed display some statistically as well as economically significant persistence. We also analyze the predictability of classical measures (Sharpe, total return, Jensen's alpha etc.) along the same factors and we identify three of them whose discriminant power is largely superior to the others. This lends support to the hypothesis of the ability of managers to yield relative performance that can be sustained over time, provided that the investor is able to identify the right signals for portfolio selection.

To finalize our study, we analyze how the power of the best three factors could permit investors to increase their rate of return by basing their investment in a pool of mutual funds selected according to this information. We build quintile portfolios by ordering funds according to a series of combinations of the three factors, and we compare those to portfolios randomly built or created according to simple criteria typically set by investors (Morningstar MRAR score, Sharpe ratio, total return). Our results report unequivocally much higher returns when the three factors are retained and combined together in almost equal fashion.

We take great care of the quality and diversity of the sample of funds in order to favorably compare with studies sharing a similar research question. First, we use weekly data instead of monthly data as used in almost all previous studies. Second, we examine a panel of international mutual funds. While previous papers often focus on a single market, mostly the US mutual funds, one has to acknowledge that competition in the mutual fund industry is global. Third, we perform

6

an initial rigorous manual analysis of the sample, to reduce as much as possible the sources of noise in the data.

The remainder of our paper is structured as follows. The next section discusses the literature devoted to the subject. In section 3, we present the data and methods used to build the performance measures. In section 4, we analyze the statistical relations between the measures. We process the principal component analysis in section 5. We also analyze the persistence of the factors and their predictability in comparison with classical measures in this section. In section 6, we examine the economic significance of our persistence results through the behavior of funds built using the most important factors. Section 7 concludes our findings.

1. LITERATURE REVIEW

To the best of our knowledge, only few papers, all published during the 21st century, investigate the research question of the underlying dimensions in performance measurement.[iii]

The earliest paper by Plantinga and De Groot (2001) studies a panel of 253 mutual funds over 26 years. They compute rank correlations for five classical measures (Sharpe ratio, Sortino ratio, Sharpe-Alpha, Fouse index and Upside potential ratio). The results range from 88% for Sharpe vs. Fouse to only 21% for Upside potential ratio vs. Sharpe. The correlation decreases when the risk aversion parameter increases. Several subsequent papers (Hwang and Salmon, 2002; Gemmill *et al.*, 2005; Razafitombo, 2010) reach a similar conclusion on smaller samples.

Pedersen and Rudholm-Afvin (2003) examine the performance of 400 financial and 254 small stocks between 1998 and 2003. They examine Spearman's rank correlations between the Sharpe ratio, Treynor ratio, Calmar, Sterling, Sharpe ratio based on the VaR, Sharpe ratio based on the CVaR and Sortino. Correlation levels are higher than 85% for financial firms, but lower for small firms. When they subdivide their sample in quartiles, the correlations present a broad spectrum of

values, reducing to as much as 55% for Treynor vs. Sharpe. The authors conclude that different approaches are needed, according to the considered assets, and that the mean-variance approach is not always applicable.

The first study based on a large sample of funds was performed by Frohlich *et al.* (2006). They consider a wide panel of 2,861 mutual funds between 1993 and 2004. Using monthly data, they compute the Sharpe ratio, Treynor ratio, Jensen's alpha, and the Fama-French and Carhart alphas. They observe Spearman's as well as Kendall's correlations that are largely inferior, by ca. 5%, compared to previous papers. The funds ranked in the top 5 (0.2 percentile) with a given measure are ranked #242 (8.5 percentile) on average according to any other measure. Results for the bottom 5 funds are more closely aligned, as the overall average is 2,839 (99.2 percentile). These five measures seem more concordant at ranking the worst performers than the best ones.

Eling and Schuhmacher (2005, 2007), in a pair of papers that are close to the topic of our study, document high observed correlations between ratios. Their first study is based on monthly returns of ten hedge fund indices between 1994 and 2003. They compute a series of ratios – Sharpe, Omega, Sortino, Kappa 3, Upside potential ratio, Calmar, Sterling, Burke, Sharpe ratios based on the CVaR and on the VaR – and observe large Spearman's rank correlations, very often higher than 95%, slightly lower for Calmar and Upside potential ratio. Their subsequent study reaches similar global conclusions regarding the performance of 2,763 hedge funds from 1985 to 2004. Correlations drop slightly by 1 or 2% if shorter periods (from 2 to 5 years) are considered. But the drop in the correlations is much higher, decreasing to 25% if the (monthly) reserve return selected in measures like Sortino is 1%, corresponding to aggressive investors. Eling (2008) examines the same 11 measures on a wide panel of 38,954 mutual funds (with 17,817 stock funds, 12,279 bonds funds, among others) with monthly prices between 1996 and 2005, and again notices similar results. Using daily data on a sample of 109 UK investment trusts, Adcock *et al.* (2010) get

analogous conclusions by using the Sharpe and Sortino ratios, then Sharpe ratios based on VaR and on CVaR.

Schuhmacher and Eling (2011, 2012) provide further theoretical support to these empirical findings. The ranking of investment funds with comparable return distributions, excluding location and scale[iv] parameters, is independent of the Sharpe ratio or another "admissible"[v] measure. The definition of admissible measures is large enough to include variations of the Sharpe ratio where the risk is evaluated by lower-partial moments or Value-at-Risk. It covers all measures considered in their previous empirical analysis.

Four recent papers significantly extend the list of tested performance measures.

Using a sample of 4,048 hedge funds between 1996 and 2005, Eling *et al.* (2011) consider recently introduced measures – Sortino-Satchell, Farinelli-Tibiletti, and Rachev ratios – on a large range of parameter values. Their main finding is the sensitivity of correlations to the choice of floating parameters for several performance measures. For the Farinelli-Tibiletti ratio, the choice of parameters p and q, corresponding to the investor's trade-off between high stakes and huge losses, induces concordance levels that can decay from 90% down to 59% for high parameter values of the two. Similarly, values for the Rachev ratios are higher than 90% when the thresholds are around 0.5, corresponding to a close to mean-variance investor; but when the thresholds are reduced to 0.05 or less, favoring the tails of the distribution (far from the centered values of Sharpe), correlations drop towards 50% too.

On a sample of 270 hedge funds, between 1997 and 2006, Zakamouline (2011) reaches even wider correlation ranges. He explains that the differences in terms of performance might not be significant between non-normal distributions if they exhibit similar kinds of deviation from normality. He shows that correlations are higher with lower Sharpe ratios and that differences

between rankings increase when distributions become more asymmetric – having high skewness. He also remarks that highest correlations do not imply identical rankings.

Caporin and Lisi (2011) use the widest panel of performance measures on a sample of stocks returns from the S&P 1500 index between 1990 and 2008. On sliding periods of 36, 60 and 120 months, they compute a set of 80 measures, including 20 Rachev ratios, 18 Farinelli-Tibiletti, 19 Loss aversion, 3 Morningstar, and the classical Treynor, Information ratio, MAD, Minimax, M2, Calmar, Sterling, Burke, Sortino, Kappa 3, Sharpe with VaR, Sharpe with CVaR. They observe that rank correlations are time variable, from less than 70% to more than 90% in the extreme cases, and that correlations on shorter periods are generally lower.

Finally, Ornelas *et al.* (2012) examine correlations with the Sharpe ratio on a set of 3,855 US mutual funds, between 1998 and 2008, and similar measures. They report most values around 90%, falling to 55% for Upside potential and Information ratios. The calculation of Kendall's correlations provides values lower than Spearman's. The authors also process computations with daily data and find slightly lower correlations.

Although many tested measures appear to be highly correlated with the Sharpe ratio, the number of exceptions is not marginal. In particular, parametric measures provide low levels of correlation when they characterize the profile of an extremely aggressive investor. From a temporal viewpoint, the duration of the performance measurement period is critical; furthermore, not all correlations remain stable over time. The shape of the return distributions of the observed funds is also a key driver of concordance between rankings. Kendall's correlations are 10% to 20% lower than Spearman's. Even with a high level of correlation, the best funds for one measure are not in the top list for others. Nevertheless, some stability between measures is observed when considering the worst funds.

This review identifies a clear case for extending the performance measurement universe to a wider set of measures and parameterizations. The dimensions of portfolio performance are likely to be neither restricted to the Sharpe ratio and its close substitutes, nor adequately encompassed by the set of classical measures.

Further, besides studies that purely focus on performance, a large stream of literature analyses the relationship between the main characteristics of the funds (age, size, management tenure and turnover) and their performance. There is a large consensus to say that new funds perform better (see among others Ferreira *et al.* (2013), Pastor *et al.*, (2015)), but the impact of size is not clear. Some studies, like Chen *et al.* (2004), Yan (2008) and Edelen *et al.* (2013), highlight that larger funds underperform their smaller peers. However, Elton *et al.* (2012), Ferreira *et al.* (2013) generally claim the opposite. Berk and Green (2004) propose a model that associates fund size with inverted persistence: better funds grow in size because investors chase performance, but larger funds suffer from various negative phenomena (diseconomies of scale, necessity to invest more money in less performing assets for diversification…) and an equilibrium is attained in steady state. Also contrasted are the relations between performance and the management tenure. Some authors, like Filbeck and Tompkins (2004) or Ferreira *et al.* (2013) report that more experienced managers perform better, but Chevalier and Ellison (1999) find no significant relationship. Peterson *et al.* (2001) even highlight an average negative alpha correlated to management tenure, because managers underperform the last years prior to their departure. Khorana (2001) documents significant improvements in post-replacement performance relative to the past performance of the fund, while Kostovetsky and Warner (2015) show that manager turnover has no effect on future return performance – but these authors draw the attention on the relevance of the performance measure which is used. Henceforth, we explicitly take age and size into account in our robustness checks and the study of ex-post persistence in performance.

2. DATA and PERFORMANCE COMPUTATION

We start with the description of the data sets used for the analysis and of their treatment in order to be workable in the scope of the study. Next, we describe the methods used for the construction of variables used as predictors of fund failure.

2.1. DATA

2.1.1. MUTUAL FUND DATA

We exploit a database of weekly[vi] returns from December 30th 1994 to January 8th 2010 (15 years) of 2,794 open-ended accumulation mutual funds[vii] with major or full allocation in equities on a worldwide basis, denominated in five major currency instruments (GBP, EUR, USD, CHF and JPY).[viii] All data are extracted from Thomson Reuters Datastream[ix].

The raw database is contaminated with a number of potential sources of interference and we take strong actions[x] to mitigate these. This leaves us with a final sample of 1,625 funds: 706 in GBP, 663 in EUR, 200 in USD, 30 in JPY and 26 in CHF. Considering the country of domiciliation, we have 696 funds issued from the United Kingdom, 405 from Luxembourg, 178 from France, 114 from Italy, 89 from Belgium and 58 from Austria. Other countries are represented in less than 50 funds. Summary statistics on the 1,625 funds are given in Table I.

[Insert Table I approximately here]

Average yearly returns are all positive except for funds denominated in JPY despite the 2007-08 crisis. They are higher for funds denominated in GBP or issued in the United Kingdom. Standard deviations are in the neighborhood of 20% for all currencies.

Skewness is always negative, and kurtosis is very positive. For more than 95% of the funds, the hypothesis of normally distributed returns can be rejected at the 5% confidence level using the Jarque-Bera statistic. Thus, performance measures based on the mean-variance framework are

likely to produce inaccurate outputs. The use of a larger array of performance measures for these funds is warranted.

The series of prices is uninterrupted for 973 funds (about 60% of the total). The remaining funds ceased to publish prices, mainly due to the absorption by or fusion with another fund, or to their liquidation. As our analysis focuses on the measures, not on the funds' performance, this survivorship rate is not a concern; nevertheless, we study in detail the consequences of fund liquidation in the section 5 devoted to performance persistence.

2.1.2. MARKET DATA

The 3-month Treasury Bill in the currency of the fund is taken as the risk-free rate. When available, we use the main stock index of the country where a fund was issued, as a proxy for the market. In the case of small countries, like Liechtenstein, we take the index of the most important neighboring country – or the average of the neighboring countries, in the case of Luxembourg. We smooth inflation rates of the involved countries by computing yearly moving averages of data retrieved from various official sources of information, mainly Central Banks and Eurostat.

Some performance measures, like the Information ratio, the Generalized Black-Treynor ratio or the Total risk alpha, require either the specification of a return generating process or the identification of a benchmark portfolio for the fund under review. We adopt the strong form of return-based style analysis (RBSA) framework proposed by Sharpe (1992). There are two reasons for this choice, also adopted in Cogneau and Hübner (2015). First, this approach leads to superior benchmark definition over self-reported benchmarks for many funds. Second, we can identify a replicating portfolio for every fund, including absolute or total return funds, which is necessary in order to produce a large number of performance measures.

Because we deal with international equity funds, we select a universe of 40 indexes that span the world markets (see Appendix 1). A first set is based on geographical delimitation, including

North America, Europe, Asia, and emerging markets. We also consider indexes dealing with company size and sectors, gold and oil.

To determine the benchmark of each fund, we then apply the RBSA on 41 indexes (the 40 market indexes with prices converted to the currency of the fund and a 41st index which is the domestic risk-free asset return). We select style indexes for each fund using the procedure described by Lobosco and DiBartolomeo (1997). We implement it through a three-step process: (i) we first regress the returns of the fund on the potential benchmarks, to determine 41 positive weights. We compute the standard deviation and set the 95% confidence interval for each weight. We retain all indexes having a strictly positive weight and an upper bound for the confidence interval greater than 10%; next, (ii) we reiterate the procedure with the selected set of indexes. For all further steps, we keep all potential benchmarks having a strictly positive weight and an upper bound greater than 20%; finally, (iii) we stop the process when no index exits from the list or when only two indexes remain – one of these being the risk-free.

2.2. COMPUTATION OF PERFORMANCE MEASURES

To date, there are more than a hundred performance measures for funds listed in the academic and practitioner's literature (Cogneau and Hübner, 2009a,b; Caporin *et al.*, 2013). We select and compute about 70 of those measures. For parametric measures, we consider multiple variations[xi]. We usually consider three variants when a reference return is needed: risk-free rate, inflation rate, or zero. We compute measures using the VaR or the Conditional VaR with different thresholds. Various investor-specific parameters are implemented for measures that are to reflect investor's risk or loss aversion, in measures like Farinelli-Tibiletti ratio, Sharpe's alpha or Aftalion and Poncet's index[xii]. For alphas based on conditional models, as described by Christopherson *et al.* (1999), we retain the yield spread (spread between the 10 year and the 3 month interest rates) and the credit spread (spread between BAA and AAA rated corporate bonds). As the third and fourth

variables, we consider the inflation rate (see Chen *et al.* 1986; Ferson and Harvey, 1995) and the market implied volatility (Bollerslev *et al.*, 2011) represented by the VIX index. This leads us to the computation of 147 performance measures, a detailed list of these can be found in Appendix 2. Each fund's performance is estimated over various time scales (annual, from one to five years), considering moving windows rolling every week over the full length of the data. This leads us to a total of 822,423 observations (resp. 738,052, 654,368, 571,740 and 492,417) (# funds × # performance measures × average # of windows per fund) for yearly (resp. two-year, three-year, four-year and five-year) measurement frequencies.

We finalize this computational step by centering (removing the mean) and standardizing (dividing by standard deviation) each measure, to get their normalized versions, which are used in the analysis.

3. The DIMENSIONS of PERFORMANCE

Our first objective consists of exploiting the universe of performance measures to identify a limited set of relevant dimensions to assess the skills of asset managers. We start with a correlation analysis, then perform a principal component analysis.

3.1. CORRELATION ANALYSIS

Much of the previous research reveals high correlation levels amongst ratio-based performance measures. We conduct a similar analysis considering our set of 147 measures. Our aim is to restrict this set to a thinner array whose dimensions can be further investigated.

We consider the sets of performance measures computed on the funds for consecutive durations of one year – i.e. excluding all overlapping periods. For each of the fifteen years of data, we build the matrix of Spearman's rank correlations among the 147 measures and take their

average. The process is repeated with Kendall correlations. We also build similar average matrices for longer durations – e.g. two, three, four, and five years.

In accordance with analyses presented in previous papers, we first consider[xiii] the correlations between measures and the Sharpe ratio. Evidence confirms the observation of Frolich *et al.* (2006) and Razafitombo (2010) that Kendall correlations are on average 10 to 11% lower than Spearman, which is notoriously less conservative. We observe on average a slight decrease when the duration increases – going from 1 to 5 years. There are however differences between measures and even inside classes of measures. The results based on three-year windows enable us to retrieve once again the findings of Eling and Schuhmacher (2005, 2007) and Eling *et al.* (2011) on similar sets of measures. In particular, we encounter 37 measures with correlations lower than 50%, 21 of which have correlations lower than 25%. Most of these are Rachev ratios and all Farinelli-Tibiletti ratios when their parameters are high, indicating high conservativeness risk aversion.

On a second step, we consider market timing measures: correlations are particularly low, and we even observe some negative values.

3.2. REDUCTION OF THE SET OF MEASURES

Cogneau and Hübner (2009a,b) and Caporin *et al.* (2013) show that more than one hundred measures proposed in the literature aim to address various aspects of portfolio performance. Our goal is to determine their main underlying dimensions. As a first step[xiv], we have to reduce this large quantity of measures, to remain with a smaller set facilitating a more precise analysis. To eliminate the superfluous measures amongst the 147, we remove all collinear measures based on the average of the 15 yearly Spearman correlation matrices through a stepwise elimination procedure.

- At each step, we build the list of all measures whose correlation with another remaining measure is maximum; this list contains at least two measures;

- For all measures in this list, we eliminate the measure whose correlations with all other measures is minimum;

- We keep the set of measures whose two-by-two correlations do not exceed 90%.[xv]

  Our final data set reduces to 36 measures. Table II reports their correlation matrix.

[Insert Table II approximately here]

This set contains only seven of the measures exploited by the previous papers mentioned in the literature review: two members of Sortino-Satchell, two of the Farinelli-Tibiletti families, and three Rachev ratios.

We retrieve at least one member of each class of measures, according to the decomposition presented in Cogneau and Hübner (2009a,b):

1. The list features twelve "Market timing" measures. There are five distinct variations of the alpha (four alphas from the classical models of Treynor-Mazuy and Henriksson-Merton, the conditional variation of the Treynor-Mazuy, and the multifactor variant of Henriksson-Merton; then the market timing alpha of Treynor-Mazuy). In parallel, we retain a similar list of six variations of the gamma, and the delta from the cubic model.

2. The Stutzer convergence index represents the class of "Preference-based" measures. It appears twice, with the risk-free rate and the inflation rate as reserve return, with a correlation close to 80%.

3. The Fouse index with risk-free rate and a risk aversion of 1 represents the class of "Gain-based difference" measures.

4. Seven "Return-based difference" measures remain: SRAP, eSDAR, Aftalion and Poncet index with market price of risk equal to 2, and four variations of regression alphas (Jensen, Fama-French, Hwang-Satchell, and Total risk alphas).

5. Seven "Gain-based ratios" are included: two Farinelli-Tibelitti ratios (the Upside potential ratio, with low values of the parameters reflecting a moderately risk-averse investor; another with large values p = 2 and q = 4, reflecting a highly defensive investor), three Rachev ratios having progressive values for the threshold of the CVaR (5% and 5%; then 20% and 5%; then 50% and 20%), corresponding to various levels of aggressiveness; two ratios of the drawups on the drawdowns (one with the averages, the other with the maxima).

6. Finally, seven measures cover a large spectrum in the class of "Return-based ratios". Two measures use total risk, namely the Sortino ratio and its variation with the ratio of the skewness by the kurtosis. The modified Treynor ratio and the Generalized Black-Treynor ratio refer to systematic risk. Finally, the Moses-Cheney-Veit measure, the Israelsen information ratio and its multifactor generalization as the fourth Standardized version of the Information ratio use a notion of specific risk.

   The final list misses some classical measures. Noteworthy absences feature, for instance: the Sharpe ratio, Morningstar MRAR indexes, Sharpe with VaR and CvaR, Calmar and Sterling (all are very correlated to Sortino-Satchell ratio with inflation and parameter equal to 2), Treynor ratio (extremely correlated with the Modified Treynor ratio), Carhart alpha (correlated at more than 95% with the Fama-French alpha). In particular, the very high correlation of Sharpe ratio with Sortino-Satchell reveals that considering the semi-variance instead of the variance does not provide significant information.

3.3. PRINCIPAL COMPONENT ANALYSIS

We carry the Principal Component Analysis (PCA) over a one-year investment horizon.

[Insert Figure 1 approximately here]

Figure 1 displays the proportion of variance explained by the principal components (PCs). The first eigenvalue is 11.379, meaning that the first eigenvector explains 31.6% (= 11.379 / 36) of the variance. The proportion of variance explained by the first ten factors is already higher than 75%. If we consider the first eighteen factors, this proportion reaches 93.9%. Out of this analysis, very few eigenvectors bear an intuitive interpretation. We thus execute an orthogonal (Varimax) rotation. With this transformation, each original variable tends to be associated with a reduced number of components, and each component represents only a small number of variables. Figure 2 reports these results, with a presentation similar to Figure 1.

[Insert Figure 2 approximately here]

The first factor explains less of the total variance, but the added explained variance decreases more slowly with the first factors in Figure 2 than in Figure 1, when considering the following factors. Consequently, after ten components we can already explain about 70%, and at the eighteenth we explain more than 92%. We narrow our further analysis to the first eighteen factors, and discard the rest. We justify this choice on three findings: (i) 90% of total variance is explained; (ii) the next eigenvalue is much lower than 1; and (iii) after the eighteenth, no factor contains a measure with a loading higher than 75%.

Table III reports the composition of the first eighteen factors after the Varimax rotation[xvi]. We display the loadings of the measures only when these are greater than 75%, and do not report measures that do not load to any factor with a loading greater than 75%.

[Insert Table III approximately here]

The rotated factors lend themselves to a much easier interpretation. Factor 1, built with two ratios of drawups by drawdowns and two utility functions, represents a generic efficiency measure based on the returns. Factor 2, standing for incremental return, is a mix of an alpha corrected for market timing and an alpha relative to a weighted benchmark. Factor 3 represents the ability to anticipate the evolution of the market in general, while factor 4 is based on three incremental return measures. Factor 5 typically represents a measure adapted to an investor that is strongly risk-averse (a ratio of gain by losses, with a high power at the denominator). The next factor, 6, is a combination of the two alphas: the original and simple Jensen's alpha, and a more sophisticated one, computed with varying betas and where market timing effects are excluded. Factor 7 is a pair of ratios of expected shortfalls, for which the thresholds are relatively high, corresponding to investors with medium risk tolerance. The following factors usually load to a single measure, and so their interpretation is straightforward[xvii].

Beyond the seventh factor, the proportion of explained variance of each factor falls[xviii]. Figure 2 also shows that 59.6% of the total variance is explained by the first 7 factors. The pattern is relatively close to the corresponding one in Figure 1: before rotation, factors 1 to 7 altogether explain 66.5% of total variance of performance measures. Thus, less than 7% of explained variance is lost with the Varimax rotation.

3.4. ROBUSTNESS CHECKS

We proceed to a series of three types of robustness checks on the results of the principal component analysis. These checks are meant to address estimation, specification and methodological issues. The first check, which is related to performance estimation issues, aims at verifying the sensitivity of the factor perimeters to a change in the performance measurement window. As a second robustness check, we examine whether the identification and relevance of factors is altered by splitting the universe of funds is subsamples. We conduct the split according

to both naïve (random) and informed (according to a potential classification variable) rules. We also split the sample in two distinct time periods. Finally, we conduct a totally different sample partitioning technique through cluster analysis to determine the concordance of the families of factors with the resulting clusters.

Our first check targets potential performance estimation issues. We examine whether our proposed decomposition of performance dimensions also holds for durations longer than one year for past performance computation. Table IV reports the aggregate results of two additional principal component analyses, considering performance computed over three and five years. The performance measures that load to each rotated factor are compared with the one-year outputs.

[Insert Table IV approximately here]

Examining one year and three years in parallel, we see that factors are extremely similar. Two other alphas complete the second factor, and the Fouse index disappears from the fourth factor. For the rest, we only observe a small reordering of the lowest factors – but the levels of explained variance remain at neighboring levels. Considering a duration of five years, the differences remain marginal. The first factor is split as the Stutzer index initiates its own factor, while the Standardized Information ratio shifts from factor 4 to factor 1, and eSDAR is no longer a component of the factors. Other factors are simply reordered. We can infer from Table IV that the factors we have built are relatively insensitive to the duration of the computed performance – while the slight differences should be part of further analysis.

The second check focuses on the robustness of the decomposition to a change in the mutual fund universe. We therefore analyze the stability of our results across subsamples of funds. As a basic approach, we first consider a naïve split by randomly creating subsamples of equal size. Next, we examine how sensitive our results are to other parameters often highlighted by previous research, as mentioned in our literature review. According to the sample of data we use, i.e. a large

and international panel of funds, we focalize on the age and the size of the funds. On one hand, we divide our sample into three classes of funds, according to their age tercile. On the other hand, we divide the universe into three classes according to their assets under management, still using terciles. Finally, we consider distinctly the nine first years of performance and the five following years.

On each of those subsamples, we rerun the principal component analysis with the orthogonal (Varimax) rotation. The results of the PCA for both the naïve, the informed, and the temporal subsamples are reported in Table V.

[Insert Table V approximately here]

Table V highlights a noticeable stability of the factors. In most of the cases, the composition of the seven factors remain exactly the same, with very similar weights of the components. It happens that a measure is added due to a weight slightly higher than 75%, but we observe that the added measures do not materially change the interpretation of the factors. Conversely, a few measures are excluded of their owning factors because their weight is slightly lower than 75% - but again, this does not affect the substance of the factor according to the remaining measures[xix].

Finally, we conduct a clustering analysis to assess the similarities among the 147 computed performance measures. This check enables us to assess the dependence of the performance decomposition to the statistical technique used. We adopt a top down approach based on the eigenvalues, to group measures in clusters[xx].

[Insert Figure 3 approximately here]

The results are very consistent with the output of the principal component analysis. The cardinality of the clusters on the last line of the dendrogram is almost equal to the number of factors that we have highlighted in Table III. In general, the relation is one-to-one between the eighteen

clusters and the factors. This process has been executed for durations equal to one year, three years and five years, with the same results.

4. The PERSISTENCE of PRINCIPAL COMPONENTS

4.1. PREDICTABILITY OF THE PRINCIPAL COMPONENTS

Cogneau and Hübner (2015) have reported that an adequate selection of performance measures can partially predict, after controlling for various effects such as size or age, the subsequent disappearance of mutual funds. In this paper, we intend to go one extra mile, and try to detect whether past performance, when carefully measured, is also adequate in order to predict future performance.

Based upon the results of the previous section, we restrict the analysis to the persistence and the robustness of the first seven rotated factors. However, for practical and commercial reasons, it is likely that the use of factors built by the combination of different measures is not usable by practitioners. We tackle this issue by processing the following analysis on periods of one year, replacing each factor by only one measure which most significantly loads on it (i.e. which is the most correlated to the factor, as reported in Table III)[xxi].

We compute them on yearly sliding periods for our sample of 1,625 mutual funds[xxii].

For each possible pair of measures (7 x 6 = 42), we construct twenty-five portfolios that are rebalanced each year in the following way. We rank our funds in quintiles, according to the value of the first measure of the pair recorded during the previous year. Each of these quintiles is, in turn, subdivided into five sub-quintiles, according to the ranking of the second measure of the pair. We materialize twenty-five portfolios by initially investing $1,000 to each, which is equally split into all funds from the corresponding class. We track these portfolios during the subsequent year

and rebalance their composition at each anniversary date according to the same criteria. If a fund disappears during the year, its last value is equally invested in all the remaining funds of the portfolio, on basis of their values at that moment. As a variant, we consider applying a fixed percentage penalty[xxiii] each time such an event occurs.

With the sample window of 15 years of data at our disposal, we follow these 1,050 (42 pairs x 25 sub-quintile portfolios) rebalanced portfolios for 14 years. At the end of each year, we compute their realized performance, according to the same seven measures and to six classical performance measures: total return, Sharpe ratio, Treynor ratio, Jensen's alpha, Fama-French alpha and Morningstar's MRAR with a risk aversion coefficient of 2 (MRAR2). We then compute their average realized values over the 14 years. For all thirteen measures, we build the ranking with the 1,050 portfolios and associate each of them to its achieved percentile in the classification, from the best (100.00%) to the worst (0.10%).

To assess the predictability of each individual dimension of performance, we adopt an original approach based on percentile analysis. We consider each factor individually: for each quintile, we compute the average percentile of the performance reached by all portfolios constructed with the second factor of each pair, i.e. the average of 30 portfolios (the 6 remaining PCs times the 5 sub-quintile portfolios for each PC) out of the 1,050 sub-quintile portfolios. For instance, for factor 1, the Q5 percentile represents the average of ranking percentiles computed for all 30 portfolios whose only common feature is that all funds are in the first quintile for factor 1. For this analysis, the use of sub-quintile portfolios enables us to get a high number of observations, while introducing no selection bias in portfolio compositions. This procedure enables us to observe how the quintiles for one measure perform the following year, according to the different selected measures.

[Insert Table VI approximately here]

Table VI reports the results of this procedure when the penalty for closing a fund is zero and returns are computed in the domestic currency of each fund[xxiv]. To get a synthetic view of the results, we report the Fano coefficient and the hit rate. The former is a dispersion measure computed as the ratio of the variance to the average. For a given PC, it shows the dispersion of the quintile portfolios for the analyzed measure – the higher, the better. The latter measure is the proportion of quintile portfolios that are in the same quintile the following year for the observed measure.

Consider first the figures on the diagonal axis, which project the persistence of each principal component on itself. The best results are provided by the fifth measure, which corresponds to one measure, the Farinelli-Tibiletti ratio for a risk-averse investor. Each quintile portfolio is almost exactly in the corresponding quintiles of all portfolios during the next year. It presents the highest Fano, highlighting the largest dispersion of the quintile portfolios, and a hit rate higher than 63% emphasizing a very good stability. Measures representing factors 2 and 4 (respectively Treynor Mazuy variant of market timing alpha and the fourth version of the Standardized Information Ratio) present good persistence, mainly for the worst funds. None of the other measures present substantial evidence of persistence. Persistence is also much less pronounced for figures outside of the diagonal axis. We observe some reversals for the link between factor 3, which represents the market timing, and the first two measures: this is emphasized by the small values of the hit rate.

4.2. PREDICTABILITY OF THE CLASSICAL MEASURES

We now draw our attention on the predictive power of the fundamental factors towards a set of classical performance measures. We try to determine which of the highlighted factors convey better information about future performance according to total realized returns, Sharpe ratio,

Treynor ratio, Jensen's alpha, Fama & French alpha, and MRAR2. This is of practical interest, because these measures are widely published, reported, and used for fund screening and selection.

Table VII shows how the quintile portfolios perform on average the following year, according to each of these classical measures, in a construction similar to that of Table VI.

[Insert Table VII approximately here]

Regarding the Sharpe ratio, the first and second measures present some predictability for both the best and the worst performers. Here again, predictability is higher for the worst funds. The third measure is peculiar, in the sense that the order of ranking for Sharpe ratio is almost the reverse of the quintiles, telling once more that market timers are bad performers in general. Similar results, while less contrasted, are observed for the Treynor ratio. For an investor only interested in obtaining a higher total return, the fourth factor is the best choice.

Regarding Jensen's alpha, the link is much more difficult to draw. Only the worst performers can be foreseen, thanks to factor 2. Better predictions are obtained for Fama-French alpha, often similar than those we observe for Sharpe, Treynor and the total return. The best predictor for the MRAR2 is clearly the fourth factor, for which the scale is fully maintained. Measure 2 can also be used, but only to detect the future worst performers.

The Fano coefficient suggests that two measures emerge. Measure 2 (market timing alpha) and 4 (Standardized Information Ratio 4) provide the most pronounced distinction between the best and the worst performers for all six classical measures.

We also perform a similar analysis, considering the second principal components in the pair as described above. We obtain[xxv] analogous results, but less contrasting: this is clearly due to the fact that this factor appears second when building the portfolios, so its importance is lowered in the construction of the portfolios.

To fine-tune our findings, we directly examine the 1,050 individual portfolios, focusing on the tails of the list, for each of the classical measures. We report in Figure 4 the proportion of occurrences of each quintile according to the first factor of the pair, in the Top-50 (left part) or Bottom-50 (right part) funds for each of the six classical measures. Visually, the presence of a larger bubble in one coordinate (X,Y) indicates the overrepresentation of good (left graphs) or bad (right graphs) funds, as measures with the corresponding measure, from funds drawn from quintile Y with factor X in the previous period. The theoretical proportion under the null hypothesis of absence of predictability should be $1 / (7 \times 5) = 2.86\%$.= This would correspond to bubbles of equal size scattered across each graph. Perfect predictability would induce large bubbles in the bottom row for the Top-50 (left graphs) and in the top row for the Bottom-50 (right graphs)

[Insert Figure 4 approximately here]

It clearly appears that predictability related to the worst performers exceeds that of the best performers, a phenomenon often observed (see for instance Hendricks, Patel and Zeckhauser (1993), Brown and Goetzmann (1995), Carhart (1997), Cogneau and Hübner (2015)). Figure 4 also emphasizes that factors 2 and 4 are always very good predictors for the worst, far ahead of the others – confirming their high Fano index.

For the best performers, the choice of the factor depends on the classical measure. For an investor interested in the Treynor ratio or in one of the two alphas, factor 5 is the best choice, while it is factor 4 for MorningStar. If the investor only cares for the total return without attention to the incurred risk, factors 4 and 5 are the best predictors. For the highest Sharpe ratios, the most persistent quintile is the last for the factor 3, which is market timing driven. After it, the portfolios based on the two best quintiles of factor 2 are the most persistent, as well as the best quintile of factors 1 and 4 which are also in the top of the list.

One can object that previous results on persistence and predictability of classical measures are due to a momentum effect – a large stream of literature has reported that the stocks having higher past performance continue to outperform stocks with poor past performance in the next period[xxvi]. To address this potential concern, we carry out the same persistence analysis, but with subsequent performance computed over durations of 3 and 5 years.

[Insert Table VIII approximately here]

Panel A of Table VIII reports the persistence of principal measures for each factor and their predictability of classical measures when the performance is computed over 3 years. The results provide very similar evidence to that of Tables VI and VII for a duration of 1 year. The predictability of factor 2 supersedes that of factor 4.

When considering Panel B, which reports results for a performance duration of 5 years, one must keep in mind the factor adjustment explained in Table IV. The factor typical of high risk reluctance (high weight associated to extreme losses) becomes the fourth factor and it is again highly persistent. Factors 2 and 1 – which is a combination of factors 1 and 4 for shorter durations – still display high persistence.

The prediction of classical measures is best provided by factors 1 and 2, but by contrast to the results for the shortest duration, predicting which funds will be the weakest in the future is not much easier than predicting the best funds[xxvii].

In conclusion, our results show that the higher persistence and the good predictive power of some emphasized factors are not due to a momentum effect. Rather, these performance measures present fundamental properties regarding their predictability. This result is potentially of interest to researchers as well as practitioners, as it shows consideration of only two or three carefully selected measures permits the building of a powerful tool, to encompass different relevant

dimensions, and to uncover persistence. It remains to be tested in an economic context through the creation of rule-based portfolios, which is the topic of the next section.

5. INVESTMENT IMPLICATIONS

The analysis conducted in the statistical PCA framework has highlighted three factors, namely the second, the fourth and the fifth factors that can altogether indicate the ability of managers to get some persistence in performance. We now illustrate how investors could realistically increase their rate of return by conditioning their investment in a pool of mutual funds selected on such information.

5.1. FUNDS OF ALL AVAILABLE FUNDS

We first build a series of portfolios that are yearly rebalanced using a set of rules that are easily available, observed and implemented by investors:

- Consistent with the findings of Cogneau and Hübner (2015), we retain the previous 3 years as the performance period used to classify funds in quintiles.

- The performance measures consider four classical measures: total return, Sharpe ratio and Morningstar's MRAR with risk aversion parameter γ equal to 2 and 3. Then we consider different linear combinations of factors 2, 4 and 5 (complete, or reduced to one measure as described in section 5.4), by weights of multiples of 10%. We also consider an equally-weighted combination of these three factors. As a benchmark, we also consider one series of randomly rebalanced portfolios.

- When a fund closes, we apply a penalty of 0% and 5%. We have checked that other levels deliver even stronger results than reported here.

[Insert Figures 5.a to 5.f approximately here]

Figures 5.a to 5.c report the average returns for the portfolios built with the original factors, while Figures 5.d to 5.f are built with factors restricted to their most significant measure. Both series reveal that the highest returns (green bars) are obtained with almost equal weights between the three factors. Restricting the factors to one single measure preserves the general impression, but it clearly results in a reduction of subsequent performance for almost all combinations.

[Insert Table X approximately here]

Table X compares yearly performance of the quintile portfolios for the selected measures. We test quintile portfolios built using a linear combination of the factors, keeping only equally weighted portfolios (three times 33%), portfolios for which one factor is underweighted (20%, 40% and 40%) and portfolios for which one factor is overweighed (40%, 30% and 30%). Hence, we control for the power of the three retained factors. In panel A, we consider the factor in its full version. The best quintile portfolios provide much higher returns, almost always higher than 6%, with a peak of 6.24% when factor 5 is overweighed.

If we consider factors reduced to their most significant component (Panel B), returns of the best quintile portfolios are slightly reduced, in the neighborhood of 5.90%. The optimum is now attained for portfolios when factor 2 is slightly overweighed. As volatility of portfolios built using factors is barely 1% higher, Sharpe ratios are also much better for Q1 than Q5 portfolios. We get a peak value of 0.282 for corrected Sharpe ratio when considering portfolios built with a combination of the three factors, overweighting factor 5.

Panel C considers classical measures for the construction of quintile portfolios. When no penalty is applied for a closed fund (left part of the Table), the best quintile portfolios built with Morningstar ratings provide a yearly return lower than 4.5%, while using the Sharpe ratio we get returns near 5%. The best result is obtained by merely using total returns, which allows us to obtain 5.18%. In comparison, randomly built portfolios provide returns near 3.75%. All portfolios display

very similar levels of volatility, and the ranking according to Sharpe ratio is the same. However, correcting for autocorrelations as proposed by Lo (2002), the quintile portfolio built using Sharpe ratio takes the first rank with 0.224, still largely below the performance achieved with combinations of factors.

If we integrate a penalty of 5% when a fund closes, the right side of Table X reveals that results are simply rescaled: the best quintile portfolio, obtained by combination of factors that slightly overweighs factor 5, provides a return of 6.09%, while total return gives 5.04%, Morningstar or Sharpe remains below 4.80%, and randomly built portfolios never exceed 4%.

We also report in Table X the returns, volatilities and Sharpe ratios for the worst quintile portfolio, indicating that the factors are discriminating well, as the performance for the worst quintile is always lower than for quintile portfolios built through Morningstar, Sharpe ratio and total return. As expected, for randomly based portfolios, there is clearly no relation between the level of the quintile and the average return.

For the sake of completeness, we perform some tests to check how significant the difference between the best quintile portfolio and each of the four other quintile portfolios is. For the average return, we compute the p-values; for the Sharpe ratio, we perform Jobson-Korkie (1981) and Opdyke (2007) tests. For all measures reported in Table X, except random, these four tests are always significant at a threshold of 99%[xxviii].

5.2. FUNDS OF SMARTLY SELECTED FUNDS

In the previous subsection, we have obtained the results of Figure 5 and Table X by building rebalanced portfolios that use all 1,625 funds available in our dataset, meaning that all quintile portfolios are a mix of more than two or three hundred funds. However, a manager who intends to manage a fund of funds will probably have fewer funds in its portfolio. To take this constraint into

account, we build quintile portfolios by selecting 20 funds from the set of funds constituting the performance quintile on the last three years. This selection of 20 funds is done randomly but respecting the proportions by which each benchmark is represented in the quintile (e.g. if 40% of the funds are benchmarked against the S&P 500 index, the same proportion of funds has to be represented in the portfolio). To control for the randomization, we reproduce this exercise of rebalancing 100 times on the whole sample of 1,625 funds. Finally, we repeat the procedure twice, splitting the whole sample into two subsamples, built according to the last digit of the ISIN code (odd or even).

[Insert Table XI approximately here]

Table XI reports that average returns are always higher for quintile portfolios built using a combination of the factors: the increase in return presents similar orders of magnitude as observations in Table X. The best choice also appears to be a combination where factor 5 is slightly overweighed, whatever the level of applied penalty for a closed fund. When considering factors reduced to their most significant measure, there is no optimal combination of the three sets of funds (full sample, subsample 1, and subsample 2). However, similar to evidence presented in Table X, the combination overweighting factor 2 is always in the top two returns. Table XI also highlights that our results are robust to the sampling in the funds.

[Insert Figures 6.a to 6.h approximately here]

To complete evidence presented in Table XI on average returns, we report in Figure 6 the whole distributions of returns for the best quintile portfolios built by taking 20 funds in the whole sample, according to the four classical measures or randomly, and we compare these with the distributions when considering a combination of full factors with overweighting factor 5, and a combination of reduced factors when overweighting factor 2.

It clearly appears from Figure 6 that the distributions of returns when considering the factors, complete or reduced, are shifted to the right. This observation can be quantified through the computation of p-values. Table XII reports the descriptive statistics for portfolios built with 20 funds from the whole sample, according to the two optimal combinations of factors and classical measures. It also features the percentages of times that a portfolio selected with each criterion (in rows) provides yearly returns higher than the average of yearly returns of portfolios similarly built, but using another measure, i.e. $\frac{1}{100}\sum_{i=1}^{100} \bar{R}_i(X) \geq \left(\frac{1}{100}\sum_{j=1}^{100} \bar{R}_j(Y)\right)$, where X is the measure used for building the portfolio (row) and Y is the measure with which the portfolio is compared (column). Visually, considering Figure 6, this measure corresponds to the proportion of observations using one bar chart that are higher than the mean value of returns using another bar chart. These proportions, corresponding to the statistical p-values, provide a simultaneous indication of both the dominating and the dominated character of each portfolio.

[Insert Table XII approximately here]

Building portfolios of funds using full factors permits the gain of returns that are always higher than with randomly based portfolios, 96% of the time for Morningstar, 89% of the time for Sharpe ratio and 85% of the time for total return. Using simplified factors, reduced to their most significant component, these percentages are in the same range, but usually lower. Overall, the Table thus provides evidence that selectively choosing funds based on full or reduced factors results in outperformance of portfolios based on classical measures more than 4 times out of 5 in all cases. On the other hand, portfolios constructed with classical measures or that are randomly selected never dominate factor-based portfolios more than 10% of the time.

5.3. IMPACT OF AGE OR SIZE OF THE FUNDS

To control for the robustness of our results, we run the same procedure on subsamples of the funds based on the age and on the size of the funds. On one hand, we divide the 1,625 funds in

terciles according to their age; on the other, we divide the 1,108 funds for which we know the AUM in terciles too. Then, we run the same rebalancing process described in subsection 6.2 for each of those subsamples[xxix].

[Insert Table XIII approximately here]

Results are extremely similar for each subsample. The previously selected full factors as well as the simplified factors largely outperform the classical measures. Predictability remains strong, after controlling for age and size.

These results, in parallel with those reported in Table V, provide evidence that our decomposition is not sensitive to potentially important measurable fund characteristics such as age and size, that previous researchers proposed to explain a variation in performance and in persistence of performance.

In summary, exploiting the three factors that we have highlighted before permits investors in funds of funds to gain substantially higher returns than those provided by the observation of criteria usually considered by practitioners. Evidence presented in this section suggests a high level of reliability of these findings.

6. CONCLUSION

The financial literature has proposed more than a hundred performance measures that can be used to rank portfolios. It is expected that a certain redundancy exists in this list, and that only a few of these measures are to be considered. Furthermore, for the investor, it is of prime interest to make a choice and to select the most appropriate for him or her, but also the most efficient. Only a few papers have focused on this, but always with a reduced set of analyzed measures, most of the time considering only US mutual funds. Some researchers conclude that the ranking remains very stable across measures, and that is enough to consider only the classical Sharpe ratio. These papers do

not aim to go further in the comprehensive analysis of the properties of these measures, and they do not cover other topics like the persistence and the relevance of the different measures for investors.

In this paper, we extend the existing literature and attempt to treat an exhaustive list, by considering a large sample of 147 parametric variations of the measures, on weekly prices for a sample of widely diversified international mutual funds.

We build a subset of 36 measures, having eliminated the highest correlations, on which we process a principal component analysis. We show finally that 18 factors are enough to represent almost all information carried in these performance measures. Focusing on the most important of these, which we set to seven, we build portfolios with the quintiles of the factors and we follow these portfolios over 14 years. We show that a factor that reflects the performance sought by highly risk-averse investors presents the highest persistence. We also report two factors that are persistent on a part of the scale, mainly on the worst performers.

We control for the robustness of our results according to the chosen sample, the size and the age of the funds.

As classical measures are widely published in reports and taken as reference for many investors, we analyze the link between the factors and those classical measures: we bring out factors which are the most suitable to exploit in order to obtain better results, or at least to avoid the worst.

Finally, we construct new performance measures, based on the three factors highlighted. With well-chosen weights for the factors, we emphasize that portfolios built according to the ranking of this performance during the last three years provide substantially higher returns.

We have verified the robustness of our results through their high stability regarding the sample, the horizon of the performance and the currency of denomination. This enables us to affirm

that performance contains fundamental elements, and it is not only a perceptual consideration from the investors.

7. REFERENCES

Adcock, C., Areal, N., Armada, M., Ceu Cortez, M., Oliveira, B. and Silva, F. (2010). Does the Use of Downside Risk-Adjusted Measures Impact the Performance of UK Investment Trusts? *Unpublished working paper, University of Minho, Braga*.

Berk, J. and Green, R. (2004). Mutual Fund Flows and Performance in Rational Markets. *Journal of Political Economy, 112*, 1269–1295.

Bollen, N. P. B. and Busse, J.A. (2001). On the Timing Ability of Mutual Fund Managers. *Journal of Finance, 56*, 1075-1094.

Bollerslev, T, Gibson, M. and Zhou, H. (2011). Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities. *Journal of Econometrics, 160*, 235-245.

Brown, S. J. and Goetzmann, W.N. (1995). Performance Persistence. *Journal of Finance, 50*, 679-698.

Caporin, M. and Lisi, F. (2011). Comparing and Selecting Performance Measures Using Rank Correlations. *Economics E-Journal*.

Caporin, M., Jannin, G., Lisi, F. and Maillet, B. (2014). A Survey on the Four Families of Performance Measures. *Journal of Economics Surveys, 28*, 917-942.

Carhart, M.M. (1997). On Persistence in Mutual Fund Performance. *Journal of Finance, 52*, 57-82.

Chen, N.-F., Roll, R. and Ross, S.A. (1986). Economic Forces and the Stock Market. *Journal of Business, 59*, 383-403.

Chen, J., Hong, H., Huang, M. and Kubik, J. (2004). Does Fund Size Erode Mutual Fund Performance? The Role of Liquidity and Organization. *American Economic Review, 94*, 1276–1302.

Chevalier, J. and Ellison, G. (1999). Are Some Mutual Fund Managers Better than Others? Cross-Sectional Patterns in Behavior and Performance. *Journal of Finance, 54*, 875-899.

Christopherson, J.A., Ferson, W.E. and Glassman, D.A. (1998). Conditioning Manager Alphas on Economic Information: Another Look at the Persistence of Performance. *Review of Financial Studies, 11*, 111-142.

Christopherson, J.A., Ferson, W.E. and Turner, A.L. (1999). Performance Evaluation Using Conditional Alphas and Betas. *Journal of Portfolio Management, 26*, 59-72.

Cogneau, P. and Hübner, G. (2009). The (more than) 100 Ways to Measure Portfolio Performance - Part 1: Standardized Risk-Adjusted Measures. *Journal of Performance Measurement, 13*, 56-71.

Cogneau, P. and Hübner, G. (2009a). The (more than) 100 Ways to Measure Portfolio Performance - Part 2: Special Measures and Comparison. *Journal of Performance Measurement, 14*, 56-69.

Cogneau, P. and Hübner, G. (2015). The Prediction of Fund Failure through Performance Diagnostics. *Journal of Banking and Finance, 50*, 224-241.

Coles, J. L., Daniel, N.D. and Nardari, F. (2006). "Does the Choice of Model or Benchmark Affect Inference in Measuring Mutual Fund Performance?" *Unpublished working paper, W.P. Carey School of Business, Tempe*.

Edelen R., Evans and R., Kadlec, G. (2013). Shedding Light on "Invisible" Costs: Trading Costs and Mutual Fund Performance. *Financial Analysts Journal, 69*, 33–44.

Eling, M. (2008). Does the Measure Matter in the Mutual Fund Industry? *Financial Analysts Journal, 64*, 54-66.

Eling, M. and Schuhmacher, F. (2005). Performance Measurement of Hedge Fund Indices - Does the Measure Matter? *Unpublished working paper, Operations Research Proceedings - Berlin*.

Eling, M. and Schuhmacher, F. (2007). Does the Choice of Performance Measure Influence the Evaluation of Hedge Funds? *Journal of Banking and Finance, 31*, 2632-2647.

Eling, M., Farinelli, S., Rossello, D. and Tibiletti, L. (2011). One-Size of Tailor-Made Performance Ratios for Ranking Hedge Funds. *Journal of Derivatives & Hedge Funds, 16*, 267-277.

Elton, E. J., Gruber, M.J. and Blake, C.R. (2012). Does Mutual Fund Size Matter? The Relationship between Size and Performance. *Review of Asset Pricing Studies, 2*, 31–55.

Farinelli, S., Ferreira, M., Rossello, D., Thoeny, M. and Tibiletti, L. (2009). Optimal Asset Allocation Aid System: From "One-Size" vs "Tailor-Made" Performance Ratio. *European Journal of Operational Research, 192*, 209-215.

Ferreira, M.A., Keswani, A., Miguel, A.F. and Ramos, S. (2013). The Determinants of Mutual Fund Performance: A Cross-Country Study. *Review of Finance, 17*, 483-525.

Ferson, W.E. and Harvey, C.R. (1995). Predictability and Time-Varying Risk in World Equity Markets. *Research in Finance, 13*, 25-88.

Filbeck, G. and Tompkins, D. (2004). Management Tenure and Risk-Adjusted Performance on Mutual Funds. *Journal of Investing, 13*, 72-80.

Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika, 10*, 507–521.

Frohlich, C., Schnusenberg, O. and Pennathur, A.K. (2006). Are Mutual Fund Performance Measures Created Equal? An Analysis of Mutual Fund Performance and Ranking. *Unpublished working paper, FMA Annual Meeting – Salt Lake City*.

Gemmill, G., Hwang, S. and Salmon, M. (2005). Performance Measurement with Loss Aversion. *Unpublished working paper, Warwick Business School, Coventry*.

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

Hübner, G. (2011). The Alpha of a Market Timer. *Unpublished working paper, University of Liège*.

Hwang, S. and Salmon, M. (2002). An Analysis of Performance Measures Using Copulae. In: John Knight and Stephen Satchell (Eds.), Performance Measurement in Finance - Firms, Funds and Managers, Butterworth-Heinemann – Elsevier, Oxford.

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

Jobson, J.D. and Korkie, B.M. (1981). Performance Hypothesis Testing with the Sharpe and Treynor Measures. *Journal of Finance, 36*, 889-908.

Khorana, A. (2001). Performance Changes Following Top Management Turnover: Evidence from Open End Mutual Funds. *Journal of Financial and Quantitative Analysis, 36*, 371-393.

Kostovetsky, L. and Warner, J.B. (2015). You're Fired! New Evidence on Portfolio Manager Turnover and Performance. *Journal of Financial and Quantitative Analysis, 50*, 729-755.

Lo, A. (2002). The Statistics of Sharpe Ratios. *Financial Analysts Journal, 58*, 36-52.

Lobosco, A. and Dibartolomeo, D. (1997). Approximating the Confidence Intervals for Sharpe Style Weights. *Financial Analysts Journal, 53*, 80-85.

Markowitz, H. (1952). Portfolio Selection. *Journal of Finance, 7*, 77-91.

Opdyke, J.D. (2007). Comparing Sharpe Ratios: So Where Are the P-Values? *Journal of Asset Management, 8*, 308-336.

Ornelas, J.R.H., Silva Junior, A.F. and Fernandes, J.L.B. (2012). Yes, the Choice of Performance Measure Does Matter for Ranking of US Mutual Funds. *International Journal of Finance and Economics, 17*, 61-72.

Pastor, L., Stambaugh, R.F. and Taylor, L.A. (2015). Scale and Skill in Active Management. *Journal of Financial Economics, 116*, 23-45.

Pedersen, C.S. and Rudholm-Alfvin, T. (2003). Selecting a Risk-Adjusted Shareholder for Performance Measure. *Journal of Asset Management, 4*, 152-172.

Peterson, J., Pietranico, P., Riepe, M. and Xu, F. (2001). Explaining the Performance of Domestic Equity Mutual Funds. *Journal of Investing, 10*, 81-92.

Pfingsten, A., Wagner, P. and Wolferink, C. (2004). An Empirical Investigation of the Rank Correlation between Different Risk Measures. *Journal of Risk, 6*, 55-74.

Plantinga, A. and De Groot, S. (2001). Risk-Adjusted Performance Measures and Implied Risk-Attitudes. *Journal of Performance Measurement, 6*, 9-19.

Razafitombo, H. (2010). A Statistical Analysis of Mutual Fund Performance Measures: The Relevance of IR, Betas, and Sharpe Ratios. *Journal of Index Investing, 1*, 89-106.

Razafitombo, H. (2012). Fund Analysis and Selection: An Approach Based on Distances and Similarities between Performance Measures. *Unpublished working paper, CEREFIGE - Université de Lorraine, Nancy*.

Schuhmacher, F. and Eling, M. (2011). Sufficient Conditions for Expected Utility to Imply Drawdown-Based Performance Rankings. *Journal of Banking and Finance, 35*, 2311-2318.

Schuhmacher, F. and Eling, M. (2012). A Decision-Theoretic Foundation for Reward-to-Risk Performance Measures. *Journal of Banking and Finance, 36*, 2077-2082.

Sharpe, W.F. (1966). Mutual Fund Performance. *Journal of Business, 39*, 119-138.

Sharpe, W.F. (1992). Asset Allocation: Management Style and Performance Measurement. *Journal of Portfolio Management, 18*, 7-19.

Yan, X. (2008). Liquidity, Investment Style, and the Relation between Fund Size and Fund Performance. *Journal of Financial and Quantitative Analysis, 43*, 741–767.

Zakamouline, V. (2011). The Performance Measure You Choose Influences the Evaluation of Hedge Funds. *Journal of Performance Measurement, 15*, 48-64.

Table I. Summary statistics of the funds returns

Panel A – Statistics by currency of denomination

	ALL	CHF	EUR	GBP	JPY	USD
Nb. of funds	1625	26	663	706	30	200
mean Nb. of obs.	555	556	558	560	597	520
Mean weekly return	0.0903%	0.0499%	0.0544%	0.1295%	-0.0151%	0.0922%
Mean yearly return	4.70%	2.60%	2.83%	6.73%	-0.79%	4.79%
Mean yearly std.dev.	20.6%	20.0%	20.9%	19.2%	22.0%	23.9%
Mean skewness	-0.422	-0.482	-0.404	-0.446	-0.334	-0.397
Mean Exc. Kurtosis	3.923	5.557	4.323	3.450	2.963	4.198
% Jarque-Bera at 5%	95.69%	88.46%	97.13%	95.33%	86.67%	94.50%

Panel B – Statistics by country of incorporation

	AT	BE	FR	GB	IT	LU	Others
Nb. of funds	58	89	178	696	114	405	85
mean Nb. of obs.	526	652	555	560	680	511	481
Mean weekly return	0.0262%	0.0834%	0.0793%	0.1303%	0.0831%	0.0468%	0.0539%
Mean yearly return	1.36%	4.34%	4.12%	6.78%	4.32%	2.43%	2.80%
Mean yearly std.dev.	20.3%	22.3%	21.3%	19.3%	17.8%	22.6%	21.9%
Mean skewness	-0.464	-0.588	-0.374	-0.447	-0.414	-0.377	-0.335
Mean Exc. Kurtosis	3.341	5.992	4.439	3.450	4.497	4.082	3.418
% Jarque-Bera at 5%	91.38%	100.00%	98.88%	95.26%	98.25%	94.32%	94.12%

Table I reports descriptive statistics for the linear returns of the 1,625 open-ended accumulation mutual funds with major or full allocation in equities. Prices are extracted for the period starting on Friday December 30th 1994 and ending on Friday January 8th 2010, from Thomson Reuters Datastream. Funds are grouped by currency of denomination (panel A) and by country of incorporation (panel B). The first rows report the numbers of funds, then the mean number of weekly observations. The following rows report the averages of the first four moments of the distributions, then the percentages of observations of funds for which a Jarque-Bera test permits to reject the normality at a threshold of 95%.

Table II. Spearman correlation matrix for the reduced dataset of measures

	Aftal_Ponc_2	Alpha_TM_cond_beta	Alpha_mkt_tim_TM	Stdzd_infor_ratio_4	eSDAR	Fama_French_alpha	Far_Tib_ifl_2x4	Fouse_rf_1	Gamma_TM_cond_beta	Gen_Bla_Trn_alpha	Hnr_Mrt_3_f_alpha	Hnr_Mrt_3_f_gam_hml	Hnr_Mrt_3_f_gam_rm	Hnr_Mrt_3_f_gam_smb	Hnr_Mrt_alpha	Hwang_Satchell	Isr_Infor_Ratio	Jensen_alpha	Mod_Treynor	Moses_Cheney_Veit	Rv_avg_dup_ddwn_rf	Rv_ifl_v01_v05_v05	Rv_max_dup_ddwn_rf	Rv_rf_v01_v2_v05	Rv_rf_v01_v5_v2	SRAP	Sortino_SK_ifl	Sortino_Sat_ifl_2	Stutzer_ifl	Stutzer_rf	Total_risk_alpha	Trn_Maz_alpha	Trn_Maz_cub_delta	Trn_Maz_cub_gamma	Trn_Maz_gamma	Upsd_pot_ratio_ifl
Aftal_Ponc_2		0	6	12	8	-22	39	1	6	1	0	3	-3	3	5	15	2	7	2	17	5	15	9	6	14	11	15	6	3	12	5	-7	2	2	-14	8
Alpha_TM_cond_beta	0		17	7	8	2	7	-52	7	9	-3	-17	-5	21	2	3	74	10	16	8	1	8	3	4	3	5	9	6	3	13	20	2	-16	-17	4	4
Alpha_mkt_tim_TM	6	17		52	80	35	45	-7	59	48	1	-23	-2	72	25	63	16	66	7	65	21	67	27	30	58	53	75	44	37	81	89	1	-25	-28	54	54
Stdzd_infor_ratio_4	12	7	52		45	43	53	-2	34	26	2	-3	2	38	28	57	5	67	5	72	29	68	31	28	51	62	76	42	35	61	49	3	-4	-5	60	69
eSDAR	8	8	80	45		45	44	2	74	43	2	0	4	50	27	68	14	63	7	65	31	64	25	20	63	62	72	45	40	82	70	5	-4	-2	56	48
Fama_French_alpha	-22	2	35	43	45		-1	1	33	10	6	15	-3	15	25	36	5	45	7	40	56	45	35	14	39	65	48	32	33	41	28	4	14	15	89	45
Far_Tib_ifl_2x4	39	7	45	53	44	-1		-1	35	24	1	0	4	33	28	50	7	48	6	65	36	64	32	22	41	62	68	43	33	52	43	-2	-4	-3	14	55
Fouse_rf_1	1	-52	-7	-2	2	1	-1		3	-4	7	22	5	-16	3	2	-5	2	-3	0	1	0	-1	-2	2	0	-1	2	5	-2	-13	-4	19	23	1	2
Gamma_TM_cond_beta	6	7	59	34	74	33	35	3		33	2	-1	3	37	20	48	12	46	7	50	25	49	20	15	48	48	55	38	34	62	52	5	-4	-2	42	36
Gen_Bla_Trn_alpha	1	9	48	26	43	10	24	-4	33		32	-45	38	62	3	30	7	33	7	31	2	30	9	14	26	21	36	25	22	36	56	9	-43	-43	22	27
Hnr_Mrt_3_f_alpha	0	-3	1	2	2	6	1	7	2	32		12	-25	-4	1	4	4	4	1	2	5	2	-1	-1	4	6	3	3	3	3	0	2	4	4	3	4
Hnr_Mrt_3_f_gam_hml	3	-17	-23	-3	0	15	0	22	-1	-45	12		9	-70	25	2	-2	-1	-3	3	18	3	3	-7	3	13	0	-1	2	-3	-46	-9	79	85	7	4
Hnr_Mrt_3_f_gam_rm	-3	-5	-2	2	4	-3	4	5	3	38	-25	9		-6	5	0	-3	3	2	4	-3	1	-5	-4	-2	0	3	5	6	-1	-6	4	5	9	-1	3
Hnr_Mrt_3_f_gam_smb	3	21	72	38	50	15	33	-16	37	62	-4	-70	-6		1	40	11	44	6	42	5	43	17	24	35	29	50	29	24	54	89	5	-67	-71	33	37
Hnr_Mrt_alpha	5	2	25	28	27	25	28	3	20	3	1	25	5	1		27	7	32	2	31	26	32	21	12	24	35	35	21	15	32	13	-1	27	28	27	27
Hwang_Satchell	15	3	63	57	68	36	50	2	48	30	4	2	0	40	27		6	61	4	67	29	65	24	21	86	60	71	40	35	68	57	3	-2	0	50	53
Isr_Infor_Ratio	2	74	16	5	14	5	7	-5	12	7	4	-2	-3	11	7	6		16	19	10	3	11	4	3	6	8	12	10	9	17	14	2	-4	-3	8	3
Jensen_alpha	7	10	66	67	63	45	48	2	46	33	4	-1	3	44	32	61	16		5	68	29	69	29	27	52	61	75	43	38	64	59	3	-3	-4	60	65
Mod_Treynor	2	16	7	5	7	7	6	-3	7	7	1	-3	2	6	2	4	19	5		7	3	8	4	1	2	8	8	15	17	7	7	3	-2	-2	8	5
Moses_Cheney_Veit	17	8	65	72	65	40	65	0	50	31	2	3	4	42	31	67	10	68	7		34	88	30	25	62	74	89	53	45	74	59	2	-1	1	57	73
Rv_avg_dup_ddwn_rf	5	1	21	29	31	56	36	1	25	2	5	18	-3	5	26	29	3	29	3	34		42	64	33	29	68	41	23	21	31	16	1	17	18	45	36
Rv_ifl_v01_v05_v05	15	8	67	68	64	45	64	0	49	30	2	3	1	43	32	65	11	69	8	88	42		38	29	62	77	88	52	44	73	60	-1	-2	0	61	73
Rv_max_dup_ddwn_rf	9	3	27	31	25	35	32	-1	20	9	-1	3	-5	17	21	24	4	29	4	30	64	38		82	21	39	37	22	27	26		-12	4	-2	42	34
Rv_rf_v01_v2_v05	6	4	30	28	20	14	22	-2	15	14	-1	-7	-4	24	12	21	3	27	1	25	33	29	82		16	15	30	15	13	23	30	-9	-8	-13	29	29
Rv_rf_v01_v5_v2	14	3	58	51	63	39	41	2	48	26	4	3	-2	35	24	86	6	52	2	62	29	62	21	16		59	68	38	33	68	51	3	0	1	50	48
SRAP	11	5	53	62	62	65	62	0	48	21	6	13	0	29	35	60	8	61	8	74	68	77	39	15	59		83	51	44	68	46	7	10	14	66	69
Sortino_SK_ifl	15	9	75	76	72	48	68	-1	55	36	3	0	3	50	35	71	12	75	8	89	41	88	37	30	68	83		58	49	82	68	4	-5	-3	67	82

Measure																																			
Sortino_Sat_ifl_2	6	6	44	42	45	32	43	2	38	25	3	-1	5	29	21	40	10	43	15	53	23	52	22	15	38	51	58		79	48	38	3	-4	-2	43	47
Stutzer_ifl	3	3	37	35	40	33	33	5	34	22	3	2	6	24	15	35	9	38	17	45	21	44	20	13	33	44	49	79		41	33	3	0	-1	41	41
Stutzer_rf	12	13	81	61	82	41	52	-2	62	36	3	-3	-1	54	32	68	17	64	7	74	31	73	27	23	68	68	82	48	41		73	5	-8	-6	58	50
Total_risk_alpha	5	20	89	49	70	28	43	-13	52	56	0	-46	-6	89	13	57	14	59	7	59	16	60	26	30	51	46	68	38	33	73		5	-46	-55	47	49
Trn_Maz_alpha	-7	2	1	3	5	4	-2	-4	5	9	2	-9	4	5	-1	3	2	3	3	2	1	-1	-12	-9	3	7	4	3	3	5	5		-9	-4	-1	-1
Trn_Maz_cub_delta	2	-16	-25	-4	-4	14	-4	19	-4	-43	4	79	5	-67	27	-2	-4	-3	-2	-1	17	-2	4	-8	0	10	-5	-4	0	-8	-46	-9		79	6	-3
Trn_Maz_cub_gamma	2	-17	-28	-5	-2	15	-3	23	-2	-43	4	85	9	-71	28	0	-3	-4	-2	1	18	0	-2	-13	1	14	-3	-2	-1	-6	-55	-4	79		4	-1
Trn_Maz_gamma	-14	4	54	60	56	89	14	1	42	22	3	7	-1	33	27	50	8	60	8	57	45	61	42	29	50	66	67	43	41	58	47	-1	6	4		61
Upsd_pot_ratio_ifl	8	4	54	69	48	45	55	2	36	27	4	4	3	37	27	53	3	65	5	73	36	73	34	29	48	69	82	47	41	50	49	-1	-3	-1		61

Table II reports Spearman's correlations (in %) between the 36 measures from the set of 147 whose two-by-two correlations do not exceed 90%. Correlations higher than 75% are highlighted in green, while negative correlations are highlighted in red.

Computed values are the averages of the 15 yearly correlations for non-overlapping durations of 1 year. Each measure is designed by its acronym: the full names are in Appendix 2.

Table III. Reduced factors after Varimax rotation

	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7	Factor8	Factor9	Factor10	Factor11	Factor12	Factor13	Factor14	Factor15	Factor16	Factor17	Factor18
Rv_avg_dup_ddwn_rf	0.88 (70%)																	
Rv_max_dup_ddwn_rf	0.85 (71%)																	
Stutzer_ifl	0.95 **(82%)**																	
Stutzer_rf	0.92 (77%)																	
Alpha_mkt_tim_TM		0.79																
Total_risk_alpha		0.89 (95%)																
Hnr_Mrt_3_f_gam_r			0.95															
Trn_Maz_cub_gamma			0.83 (90%)															
Trn_Maz_gamma			0.93 (93%)															
eSDAR				0.86 (81%)														
Fouse_rf_1				0.77 (75%)														
Stdzd_infor_ratio_4				0.84 **(87%)**														
Far_Tib_ifl_2x4					0.87													
Alpha_TM_cond_beta						0.87 **(92%)**												
Jensen_alpha						0.93 (92%)												
Rv_rf_v01_v2_v05							0.84 **(95%)**											
Rv_rf_v01_v5_v2							0.98 (95%)											
Hnr_Mrt_3_f_gam_sm								0.99										
Hnr_Mrt_3_f_gam_h									0.99									
SRAP										0.90								
Gamma_TM_cond_be											0.96							
Aftal_Ponc_2												0.97						
Trn_Maz_cub_delta													0.98					
Mod_Treynor														0.98				
Gen_Bla_Trn_alpha															0.99			
Hwang_Satchell																1.00		
Moses_Cheney_Veit																	1.00	
Rv_ifl_v01_v05_v05																		0.82

Table III reports the decomposition of the first eighteen factors issued from the principal component analysis, after Varimax rotation. Values of the loadings that are lower than 75% are floored to 0 and are not reported in the Table. When a factor is compounded with more than 1 measure, we report in brackets the Spearman correlation between each measure and the factor. The highest correlation is highlighted in bold.

Table IV. Comparison of factors for performance durations of 1, 3 and 5 years

	1 year		3 years		5 years
PC	Measure	PC	Measure	PC	Measure
1	Rv_avg_dup_ddwn_rf Rv_max_dup_ddwn_rf Stutzer_ifl Stutzer_rf	1	Rv_avg_dup_ddwn_rf Rv_max_dup_ddwn_rf Stutzer_ifl Stutzer_rf	1	Stdzd_Infor_Ratio_4 eSDAR Rv_avg_dup_ddwn_rf Rv_max_dup_ddwn_rf Sortino_SK_ifl Sortino_Sat_ifl_2
2	Alpha_mkt_tim_TM Total_risk_alpha	2	Fama_French_alpha Alpha_mkt_tim_TM Trn_Maz_alpha Total_risk_alpha	2	Fama_French_alpha Alpha_mkt_tim_TM Trn_Maz_alpha Total_risk_alpha
3	Trn_Maz_cub_gamma Hnr_Mrt_3_f_gam_rm Trn_Maz_gamma	3	Trn_Maz_cub_gamma Hnr_Mrt_3_f_gam_rm Trn_Maz_gamma	3	Trn_Maz_cub_gamma Hnr_Mrt_3_f_gam_rm Trn_Maz_gamma
4	Stdzd_Infor_Ratio_4 eSDAR Fouse_rf_1	4	Stdzd_Infor_Ratio_4 eSDAR	9	Stutzer_rf
5	Far_Tib_ifl_2x4	5	Upsd_pot_ratio_ifl Far_Tib_ifl_2x4	4	Upsd_pot_ratio_ifl Far_Tib_ifl_2x4
6	Jensen_alpha Alpha_TM_cond_beta	7	Jensen_alpha Alpha_TM_cond_beta	6	Jensen_alpha Alpha_TM_cond_beta
7	Rv_rf_v01_v2_v05 Rv_rf_v01_v5_v2	6	Rv_rf_v01_v2_v05 Rv_rf_v01_v5_v2	5	Rv_rf_v01_v2_v05 Rv_rf_v01_v5_v2
8	Hnr_Mrt_3_f_gam_smb	9	Hnr_Mrt_3_f_gam_smb	8	Hnr_Mrt_3_f_gam_smb
9	Hnr_Mrt_3_f_gam_hml	8	Hnr_Mrt_3_f_gam_hml	7	Hnr_Mrt_3_f_gam_hml
10	SRAP	17	SRAP	17	SRAP
11	Gamma_TM_cond_beta	10	Gamma_TM_cond_beta	10	Gamma_TM_cond_beta
12	Aftal_Ponc_2	12	Aftal_Ponc_2	15	Aftal_Ponc_2
13	Trn_Maz_cub_delta	13	Trn_Maz_cub_delta	11	Trn_Maz_cub_delta
14	Mod_Treynor	16	Mod_Treynor	18	Mod_Treynor
15	Gen_Bla_Trn_alpha	11	Gen_Bla_Trn_alpha	12	Gen_Bla_Trn_alpha
16	Hwang_Satchell	14	Hwang_Satchell	14	Hwang_Satchell
17	Moses_Cheney_Veit	15	Moses_Cheney_Veit	13	Moses_Cheney_Veit
18	Rv_ifl_v01_v05_v05	18	Rv_ifl_v01_v05_v05	16	Rv_ifl_v01_v05_v05

Table IV reports the factors issued from the principal component analysis, after Varimax rotation, for three different horizons of performance: 1 year, 3 years and 5 years. In the column corresponding to 1 year, factors are sorted according to decreasing explained variance. In the columns corresponding to 3 and 5 years, factors are reordered (with

mention of their real rank) according to the rank of the corresponding factor for 1 year. The transfers of measures from a factor to another when changing the horizon are highlighted by an arrow.

Table V. Reduced factors when PCA is processed on subsamples of funds

	all	sample_1	sample_2	age0	age1	age2	AUM0	AUM1	AUM2	Years 1-9	Years 10-14
factor 1											
Rv_avg_dup_ddwn_rf	0.883	0.880	0.886	0.912	0.865	0.858	0.872	0.856	0.879	0.904	0.841
Rv_max_dup_ddwn_rf	0.846	0.842	0.850	0.879	0.831	0.813	0.828	0.824	0.836	0.875	0.797
Sortino_Sat_ifl_2				*0.758*							
Stutzer_ifl	0.947	0.947	0.947	0.946	0.943	0.949	0.953	0.943	0.942	0.944	0.952
Stutzer_rf	0.924	0.939	0.907	0.888	0.938	0.947	0.948	0.939	0.935	0.901	0.948
factor 2											
Alpha_mkt_tim_TM	0.790	0.778	0.800	0.821	0.771	0.766	*0.723*	0.769	0.845	0.766	0.832
Fama_French_alpha				*0.750*							
Total_risk_alpha	0.893	0.897	0.889	0.890	0.901	0.888	0.915	0.894	0.814	0.900	0.849
Trn_Maz_alpha				*0.762*					*0,764*		

factor 3										
Hnr_Mrt_3_f_gam_rm	0.946	0.948	e.942	0.945	0.938	0.949	0.947	0.943	e.919	0.952	0.932
Trn_Maz_cub_gamma	0.830	0.829	0.833	0.838	0.848	0.812	0.825	0.828	0.867	0.824	0.834
Trn_Maz_gamma	0.926	0.925	0.927	0.934	0.930	0.905	0.923	0.923	0.931	0.926	0.925
factor 4										
Stz_Infor_Ratio_4	0.835	0.840	0.830	0.822	0.840	0.850	0.844	0.838	0.844	0.808	0.868
Fouse_rf_1	0.770	0.774	0.765	*0.742*	0.771	0.799	0.770	0.781	0.817	*0.684*	0.829
eSDAR	0.857	0.857	0.857	0.866	0.857	0.848	0.844	0.854	0.864	0.839	0.877
factor 5										
Far_Tib_ifl_2x4	0.872	0.875	0.869	0.868	0.861	0.887	0.886	0.875	0.856	0.831	0.918
Upsd_pot_ratio_ifl						*0.776*	*0.768*	*0.754*			*0.796*
factor 6										
Alpha_TM_cond_beta	0.872	0.877	0.868	0.882	0.872	0.852	0.872	0.865	0.857	0.880	0.825
Jensen_alpha	0.932	0.933	0.931	0.941	0.934	0.911	0.933	0.928	0.924	0.935	0.926
factor 7										
Rv_rf_v01_v2_v05	0.839	0.821	0.855	0.835	0.831	0.854	0.868	0.854	*0.666*	0.847	0.813

Rv_rf_v01_v5_v2	0.981	0.983	0.980	0.983	0.982	0.979	0.982	0.980	0.997	0.981	0.98

Table V reports the weights of the factors issued from the principal component analysis, after Varimax rotation, when they are higher than 75%, for different subsamples of the funds. Column "all" reports results for the whole sample, while sample_1 and sample_2 report results for two subsamples randomly chosen. In the columns age0, age1 and age2, the subsamples are respectively the older, the intermediaries and the younger funds. Column AUM0 reports results for the largest funds, AUM2 for the smallest funds and AUM1 for the remaining funds. Finally, column Years 1_9 reports results for the first 9 years and column Years 10_14 reports results for the last 5 years. We highlight in italics the measure and the weights that are not selected for the whole sample, or the measures selected for the whole sample which will be rejected for a subsample. The performance horizon for the computed measures is one year.

Table VI. Persistence of the quintile portfolios according to the set of factors

Factor	Quintile	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
	Q5	68.5	68.7	43.0	74.5	69.2	49.9	60.3
	Q4	50.9	65.7	48.0	71.8	63.5	67.1	64.1
	Q3	39.7	59.4	55.5	61.7	50.1	61.5	50.1
F1	Q2	42.7	40.6	49.4	40.3	43.3	46.3	45.3
	Q1	44.4	15.9	49.5	5.5	24.8	26.3	31.2
	Fano	2.7	9.7	0.4	16.2	6.1	5.0	3.4
	Hit rate	0.293	0.467	0.207	0.460	0.407	0.327	0.307
	Q5	67.6	77.9	45.9	72.5	80.1	60.0	44.8
	Q4	67.0	74.9	46.0	67.7	60.9	77.7	35.2
	Q3	47.5	55.8	55.5	44.9	40.4	52.4	44.7
F2	Q2	43.3	39.6	48.2	40.0	35.6	40.2	46.9
	Q1	14.7	6.8	52.5	18.7	36.9	19.4	72.0
	Fano	9.8	16.7	0.4	9.8	7.3	9.5	3.9
	Hit rate	0.420	0.467	0.227	0.407	0.453	0.400	0.153
	Q5	32.5	31.1	84.5	60.1	68.7	21.7	58.0
	Q4	18.7	28.0	50.0	41.1	31.5	21.5	40.0
	Q3	25.5	43.6	31.7	39.5	31.6	44.5	35.9
F3	Q2	55.7	64.5	48.4	54.0	56.4	67.9	62.4
	Q1	81.5	65.3	22.7	36.5	57.8	77.2	50.5
	Fano	15.5	6.8	11.8	2.3	5.7	14.2	2.6
	Hit rate	0.080	0.120	0.367	0.240	0.140	0.120	0.173
	Q5	53.9	74.6	73.5	80.5	84.5	49.2	63.8
	Q4	50.0	70.1	48.1	60.7	63.9	67.9	49.1
	Q3	59.6	57.8	52.0	57.2	43.8	65.5	45.5
F4	Q2	69.2	40.2	42.3	20.5	24.8	46.7	25.6
	Q1	25.4	9.7	30.4	17.4	40.1	19.6	73.9
	Fano	5.2	13.8	5.1	15.9	10.4	7.5	6.6
	Hit rate	0.247	0.473	0.333	0.520	0.413	0.380	0.200
	Q5	61.6	74.1	80.5	88.3	93.8	51.6	60.7
	Q4	43.3	49.0	65.2	75.6	78.7	47.3	50.4
	Q3	40.3	39.5	60.1	46.9	44.3	42.5	49.3
F5	Q2	33.7	39.5	29.8	26.5	21.6	48.8	40.7
	Q1	70.3	32.2	6.9	6.7	4.9	49.1	33.9
	Fano	4.8	5.7	18.2	23.4	28.8	0.2	2.2
	Hit rate	0.240	0.340	0.493	0.633	0.633	0.213	0.273
	Q5	63.2	55.5	38.2	58.5	53.9	61.3	38.5
	Q4	50.5	50.9	58.1	53.5	44.2	51.8	49.3
	Q3	48.5	51.3	56.5	50.0	43.5	47.4	47.5
F6	Q2	44.6	41.3	48.6	48.6	40.5	40.7	56.2
	Q1	49.1	46.6	42.2	47.5	40.5	45.0	46.6
	Fano	1.0	0.6	1.6	0.4	0.7	1.3	0.9
	Hit rate	0.293	0.280	0.207	0.240	0.287	0.287	0.220
	Q5	77.9	75.7	51.5	49.8	62.9	66.6	59.9
F7	Q4	55.9	56.4	76.3	63.8	62.3	53.1	60.9
	Q3	30.8	37.3	62.7	53.9	54.0	37.9	36.7

Factor	Quintile	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
	Q2	28.0	28.1	48.8	44.0	36.3	35.5	37.4
	Q1	55.9	43.9	9.3	33.0	22.1	50.4	44.2
	Fano	8.6	7.1	12.6	2.7	6.7	3.2	2.9
	Hit rate	0.273	0.367	0.360	0.240	0.373	0.307	0.273

Table VI reports the average percentiles of the performance for the quintile portfolios built according to the seven most important factors issued from the principal component analysis, when performance is computed on a duration of one year. Returns of mutual funds are in the currency of denomination when building the quintile portfolios. Values are reported for a return penalty of 0% when a fund ceases to report. The diagonal (highlighted) elements represent the average percentile of each factor achieved on its own value. The Fano factor is a dispersion measure computed as the ratio of the variance to the average. The hit rate is the proportion of quintile portfolios that are in the same quintile the following year.

Table VII. Performance of the quintile portfolios according to classical measures

Factor	Quintile	Sharpe	Treynor	Total ret.	MRAR2	Jensen's α	F&F α
F1	Q5	80.5	68.3	58.8	66.0	50.5	57.2
	Q4	68.7	63.3	69.5	55.6	65.3	62.3
	Q3	55.8	51.2	65.1	66.9	67.6	55.4
	Q2	33.9	40.1	43.1	44.6	46.5	42.2
	Q1	13.7	27.3	12.1	11.3	16.0	21.2
	Fano	14.3	5.6	10.9	10.7	8.7	5.7
	Hit rate	0.480	0.360	0.400	0.467	0.427	0.347
F2	Q5	75.0	75.9	66.4	59.3	55.0	70.7
	Q4	80.7	67.6	78.7	77.6	81.3	76.9
	Q3	57.9	50.3	59.5	62.6	66.5	66.5
	Q2	36.4	41.7	45.7	51.4	48.7	50.4
	Q1	6.9	17.6	7.1	7.3	12.4	4.8
	Fano	17.8	10.3	14.7	13.6	12.6	15.7
	Hit rate	0.500	0.520	0.473	0.393	0.400	0.473
F3	Q5	30.9	43.9	44.5	40.6	41.8	60.8
	Q4	30.1	25.2	33.5	26.2	30.4	39.1
	Q3	38.8	34.7	39.6	40.3	42.2	37.4
	Q2	59.9	52.7	61.1	67.8	61.8	54.0
	Q1	67.8	79.6	52.1	56.0	54.2	41.9
	Fano	6.6	9.2	2.5	5.6	3.2	2.3
	Hit rate	0.1	0.1	0.1	0.1	0.1	0.3
F4	Q5	73.5	75.9	83.4	84.4	65.8	84.0
	Q4	66.1	69.2	76.9	69.8	74.9	67.8
	Q3	66.5	57.8	65.0	62.8	66.2	62.8
	Q2	38.8	50.0	32.2	40.7	46.0	41.2
	Q1	9.8	18.1	5.7	5.5	9.5	6.8
	Fano	13.9	9.4	20.5	17.9	13.1	16.9
	Hit rate	0.460	0.493	0.547	0.573	0.507	0.613
F5	Q5	64.0	78.0	71.1	59.4	68.1	78.3
	Q4	47.8	47.7	58.5	54.6	63.1	58.4
	Q3	44.3	34.9	45.0	46.0	47.5	42.5
	Q2	36.5	31.7	35.5	38.0	35.2	37.1
	Q1	42.8	39.7	22.7	31.5	21.1	18.7
	Fano	2.3	7.5	7.7	2.9	8.0	10.8
	Hit rate	0.247	0.347	0.447	0.373	0.400	0.480
F6	Q5	57.2	56.3	56.5	62.1	53.0	55.1
	Q4	52.3	51.2	53.4	46.2	51.9	49.9
	Q3	50.2	45.8	52.2	56.4	52.4	52.5
	Q2	43.3	40.7	39.4	37.7	41.1	38.7
	Q1	41.6	40.8	42.0	46.3	41.2	42.8
	Fano	0.9	1.0	1.2	1.8	0.8	1.0
	Hit rate	0.260	0.293	0.220	0.300	0.247	0.260
F7	Q5	72.1	76.1	63.4	59.7	63.1	71.5
	Q4	54.5	55.7	60.5	62.4	63.9	67.7
	Q3	40.3	38.5	47.4	46.1	47.2	44.2
	Q2	31.5	29.2	32.9	36.5	31.9	28.1
	Q1	41.5	34.6	31.2	32.4	28.2	23.0
	Fano	5.2	7.8	4.8	0.4	6.0	10.5
	Hit rate	0.280	0.300	0.327	0.340	0.367	0.393

Table VII reports for each quintile portfolio of the seven most important factors issued from the principal component analysis, the average percentile of the performance according to six classical measures (Sharpe ratio, Treynor ratio, total return, Jensen's alpha, Fama & French's alpha and Morningstar's MRAR). Results are displayed for a time horizon of one year. The Fano factor is a dispersion measure computed as the ratio of the variance to the average. The hit rate is the proportion of quintile portfolios that are in the same quintile the following year.

Table VIII. Persistence of the quintile portfolios for longer durations.

Panel A. Duration of performance = 3 years

Factor	Type	Factor' 1	Factor' 2	Factor' 3	Factor' 4	Factor' 5	Factor' 6	Factor' 7	Sharpe ratio	Treynor ratio	Total return	MorningStar 2	Jensen's alpha	Fama&French alpha
F'1	Fano	17.7	3.8	5.0	2.9	1.8	5.5	11.4	3.97	6	7.48	7.64	6.14	2.74
	Hit rate	0.087	0.173	0.12	0.153	0.153	0.113	0.067	0.093	0.107	0.067	0.133	0.06	0.187
F'2	Fano	7.5	16.5	5.4	14.7	6.9	11.4	5.9	20.64	9.01	14.62	14.62	9.1	10.35
	Hit rate	0.213	0.433	0.173	0.48	0.333	0.387	0.193	0.567	0.28	0.407	0.407	0.333	0.427
F'3	Fano	22.3	12.4	18.5	3.2	7.1	11.0	1.6	20.19	19.96	9	6.19	3.3	1.87
	Hit rate	0.06	0.133	0.58	0.267	0.193	0.113	0.187	0.06	0.067	0.087	0.107	0.173	0.3
F'4	Fano	2.0	7.2	5.0	7.3	11.9	2.6	3.7	10.74	3.56	7.31	2.68	12.91	22.68
	Hit rate	0.173	0.393	0.34	0.327	0.34	0.273	0.227	0.453	0.3	0.3	0.267	0.407	0.533
F'5	Fano	8.4	7.7	2.9	23.8	27.8	3.7	8.9	3.71	6.16	11.31	8.17	6.19	7.02
	Hit rate	0.12	0.313	0.24	0.54	0.593	0.24	0.38	0.227	0.233	0.28	0.24	0.347	0.407
F'6	Fano	2.5	4.8	5.0	3.8	0.9	3.5	2.4	1.49	1.87	0.85	1.47	0.69	3.94
	Hit rate	0.147	0.387	0.193	0.12	0.273	0.32	0.28	0.193	0.287	0.24	0.267	0.24	0.307
F'7	Fano	0.6	3.1	2.4	0.5	0.6	3.4	3.0	0.29	1.13	0.56	1.86	1.36	3.2
	Hit rate	0.14	0.32	0.2	0.24	0.26	0.327	0.293	0.18	0.207	0.253	0.267	0.28	0.36

Panel B. Duration of performance = 5 years

Factor	Type	Factor" 1	Factor" 2	Factor" 3	Factor" 4	Factor" 5	Factor" 6	Factor" 7	Sharpe ratio	Treynor ratio	Total return	MorningStar 2	Jensen's alpha	Fama&French alpha
F"1	Fano	12.3	16.4	2.0	6.9	10.5	11.3	2.8	16.6	12.2	14.0	9.5	15.8	18.4
	Hit rate	0.467	0.473	0.153	0.447	0.4	0.427	0.173	0.56	0.487	0.52	0.433	0.527	0.573
F"2	Fano	12.4	20.2	5.2	8.0	19.6	11.8	3.6	25.6	19.1	24.8	17.8	16.6	21.6
	Hit rate	0.413	0.58	0.267	0.34	0.507	0.46	0.253	0.627	0.647	0.593	0.48	0.513	0.613
F"3	Fano	10.9	6.9	19.2	5.7	3.5	7.5	3.4	8.3	12.9	3.1	3.0	1.2	3.2
	Hit rate	0.133	0.107	0.493	0.3	0.28	0.167	0.167	0.087	0.107	0.127	0.147	0.233	0.28
F"4	Fano	1.8	4.4	3.1	15.2	2.5	1.4	6.0	1.9	5.3	2.8	2.3	3.8	3.7
	Hit rate	0.18	0.293	0.313	0.54	0.18	0.253	0.327	0.167	0.28	0.32	0.227	0.327	0.327
F"5	Fano	7.5	14.3	2.9	26.0	10.8	6.5	6.5	5.4	9.3	8.8	3.8	12.9	17.8
	Hit rate	0.32	0.42	0.24	0.687	0.48	0.393	0.287	0.287	0.353	0.4	0.26	0.427	0.42
F"6	Fano	2.5	1.8	0.7	0.3	2.9	3.0	1.3	1.6	2.1	1.4	1.8	0.7	1.5
	Hit rate	0.24	0.293	0.147	0.273	0.133	0.32	0.213	0.233	0.273	0.3	0.28	0.267	0.333
F"7	Fano	1.3	0.4	9.3	3.7	3.0	0.3	2.4	0.4	2.0	4.2	7.5	4.1	2.1
	Hit rate	0.167	0.187	0.433	0.327	0.147	0.213	0.24	0.14	0.26	0.32	0.367	0.273	0.247

Table VIII reports the persistence of the seven main factors issued of the principal component analysis composed as in Table III, and their predictability to six classical measures, when performance is computed on a time horizon of three years (Panel A) and five years (Panel B). The left part of this Table reports the Fano coefficient and the hit rates for the quintile portfolios built according to the seven most important factors issued from the principal component analysis. The right part reports, for each quintile portfolio of the seven most important factors issued from the principal component analysis, the Fano coefficient and the hit rate according to six classical measures (Sharpe ratio, Treynor ratio, total return, Morningstar's MRAR Jensen's alpha and Fama & French's alpha). Returns of mutual funds are in the currency of denomination when building the quintile portfolios, and values are reported for a null penalty when a fund delists.

Table X. Risk and return statistics for rebalanced quintile portfolios

Panel A: Portfolios based on full factors

Name	Quintile	Penalty = 0%				Penalty = 5%			
		Mean	S.D.	SR	SR-corr	Mean	S.D.	SR	SR-corr
$F2_{33}\_F4_{33}\_F5_{33}$	Q5	3.17%	15.9%	0.075	0.075	2.85%	15.9%	0.055	0.055
	Q1	6.01%	18.4%	0.236	0.269	5.86%	18.4%	0.229	0.261
$F2_{20}\_F4_{40}\_F5_{40}$	Q5	3.37%	15.4%	0.085	0.085	3.05%	15.4%	0.065	0.065
	Q1	6.13%	18.6%	0.242	0.273	5.98%	18.6%	0.234	0.265
$F2_{40}\_F4_{20}\_F5_{40}$	Q5	3.08%	15.8%	0.070	0.070	2.77%	15.8%	0.050	0.050
	Q1	6.11%	18.3%	0.242	0.276	5.96%	18.3%	0.234	0.267
$F2_{40}\_F4_{40}\_F5_{20}$	Q5	3.01%	16.3%	0.068	0.068	2.68%	16.3%	0.048	0.048
	Q1	5.84%	18.1%	0.228	0.261	5.69%	18.1%	0.220	0.252
$F2_{40}\_F4_{30}\_F5_{30}$	Q5	3.08%	16.1%	0.071	0.071	2.76%	16.0%	0.051	0.051
	Q1	6.00%	18.2%	0.236	0.270	5.84%	18.2%	0.228	0.261
$F2_{30}\_F4_{40}\_F5_{30}$	Q5	3.12%	15.8%	0.072	0.072	2.80%	15.8%	0.052	0.052
	Q1	5.98%	18.3%	0.235	0.267	5.83%	18.3%	0.227	0.259
$F2_{30}\_F4_{30}\_F5_{40}$	Q5	3.20%	15.6%	0.075	0.075	2.88%	15.6%	0.056	0.056
	Q1	6.24%	18.5%	0.248	0.282	6.09%	18.5%	0.240	0.273

Panel B: Portfolios based on reduced factors

Name	Quintile	Penalty = 0%				Penalty = 5%			
		Mean	S.D.	SR	SR-corr	Mean	S.D.	SR	SR-corr
$F2_{33}\_F4_{33}\_F5_{33}$	Q5	3.69%	15.6%	0.106	0.106	3.38%	15.6%	0.086	0.086
	Q1	5.93%	18.6%	0.232	0.260	5.77%	18.6%	0.224	0.251
$F2_{20}\_F4_{40}\_F5_{40}$	Q5	3.63%	15.2%	0.101	0.101	3.31%	15.2%	0.080	0.080
	Q1	5.67%	18.8%	0.218	0.242	5.51%	18.8%	0.210	0.233
$F2_{40}\_F4_{20}\_F5_{40}$	Q5	3.43%	15.4%	0.089	0.089	3.12%	15.4%	0.069	0.069
	Q1	5.93%	18.6%	0.232	0.262	5.77%	18.6%	0.223	0.252
$F2_{40}\_F4_{40}\_F5_{20}$	Q5	3.49%	16.2%	0.096	0.096	3.16%	16.2%	0.076	0.076
	Q1	5.90%	18.4%	0.231	0.259	5.73%	18.4%	0.222	0.249
$F2_{40}\_F4_{30}\_F5_{30}$	Q5	3.66%	15.8%	0.105	0.105	3.33%	15.8%	0.084	0.084
	Q1	5.96%	18.5%	0.234	0.263	5.80%	18.5%	0.226	0.254

Name	Quintile	Mean	S.D.	SR	SR-corr	Mean	S.D.	SR	SR-corr
$F2_{30}$_$F4_{40}$_$F5_{30}$	Q5	3.75%	15.7%	0.110	0.110	3.42%	15.7%	0.090	0.090
	Q1	5.86%	18.5%	0.229	0.255	5.71%	18.5%	0.221	0.246
$F2_{30}$_$F4_{30}$_$F5_{40}$	Q5	3.54%	15.3%	0.096	0.096	3.23%	15.3%	0.076	0.076
	Q1	5.93%	18.7%	0.231	0.259	5.77%	18.7%	0.224	0.251

Panel C: Portfolios based on other measures

Name	Quintile	Penalty = 0%				Penalty = 5%			
		Mean	S.D.	SR	SR-corr	Mean	S.D.	SR	SR-corr
MRAR2	Q5	3.70%	17.6%	0.113	0.124	3.34%	17.5%	0.093	0.093
	Q1	4.48%	17.4%	0.156	0.194	4.35%	17.4%	0.149	0.186
MRAR3	Q5	3.80%	17.7%	0.119	0.119	3.45%	17.7%	0.100	0.100
	Q1	4.31%	17.1%	0.146	0.183	4.18%	17.1%	0.139	0.174
Sharpe	Q5	3.09%	16.2%	0.072	0.072	2.72%	16.2%	0.050	0.050
	Q1	4.92%	17.3%	0.180	0.224	4.79%	17.3%	0.173	0.215
Total return	Q5	3.40%	16.8%	0.093	0.093	3.07%	16.8%	0.074	0.074
	Q1	5.18%	17.9%	0.194	0.218	5.04%	17.9%	0.187	0.210
Random	Q5	3.76%	16.1%	0.112	0.124	3.53%	16.1%	0.098	0.108
	Q1	3.75%	16.4%	0.112	0.137	3.52%	16.4%	0.099	0.120

Table X reports yearly returns, standard deviations and Sharpe ratios (standard version and version corrected for autocorrelation as described in Lo (2002)) for the worst (Q5) and the best (Q1) quintile portfolios. Panel A uses the portfolios based on linear combinations of the factors. Panel B uses the portfolios based on the linear combinations of the reduced factors, as represented by the most correlated performance measure with each factor. Panel C presents results for quintile portfolios based on classical measures or randomly constructed. Results are reported for no penalty when a fund delists, and for a penalty of 5%.

Table XI. Mean returns for the best rebalanced quintile portfolios of 20 funds

Panel A: Portfolios based on full factors

Name	Penalty = 0%			Penalty = 5%		
	Full sample	Subsample 1	Subsample 2	Full sample	Subsample 1	Subsample 2
$F2_{33}\_F4_{33}\_F5_{33}$	6.02%	6.32%	5.25%	5.87%	6.19%	5.09%
$F2_{20}\_F4_{40}\_F5_{40}$	6.14%	6.77%	5.46%	5.99%	6.62%	5.31%
$F2_{40}\_F4_{20}\_F5_{40}$	6.11%	6.72%	5.30%	5.96%	6.57%	5.15%
$F2_{40}\_F4_{40}\_F5_{20}$	5.70%	6.13%	5.21%	5.56%	6.00%	5.04%
$F2_{40}\_F4_{30}\_F5_{30}$	6.07%	6.17%	5.18%	5.91%	6.03%	5.02%
$F2_{30}\_F4_{40}\_F5_{30}$	5.78%	6.33%	5.33%	5.64%	6.18%	5.19%
$F2_{30}\_F4_{30}\_F5_{40}$	6.22%	6.80%	5.41%	6.07%	6.66%	5.26%

Panel B: Portfolios based on reduced factors

Name	Penalty = 0%			Penalty = 5%		
	Full sample	Subsample 1	Subsample 2	Full sample	Subsample 1	Subsample 2
$F2_{33}\_F4_{33}\_F5_{33}$	5.76%	6.50%	5.42%	5.60%	6.34%	5.25%
$F2_{20}\_F4_{40}\_F5_{40}$	5.56%	6.12%	5.25%	5.41%	6.03%	5.08%
$F2_{40}\_F4_{20}\_F5_{40}$	6.06%	6.24%	5.59%	5.89%	6.09%	5.41%
$F2_{40}\_F4_{40}\_F5_{20}$	5.87%	6.22%	5.30%	5.71%	6.06%	5.13%
$F2_{40}\_F4_{30}\_F5_{30}$	5.83%	6.46%	5.55%	5.67%	6.30%	5.39%
$F2_{30}\_F4_{40}\_F5_{30}$	5.69%	6.40%	5.40%	5.53%	6.24%	5.24%
$F2_{30}\_F4_{30}\_F5_{40}$	6.01%	6.30%	5.42%	5.84%	6.14%	5.25%

Panel C: Portfolios based on other measures

Name	Penalty = 0%			Penalty = 5%		
	Full sample	Subsample 1	Subsample 2	Full sample	Subsample 1	Subsample 2
MRAR2	4.48%	4.88%	4.04%	4.03%	4.76%	3.92%
MRAR3	4.48%	4.77%	3.66%	4.03%	4.65%	3.53%
Sharpe	4.94%	5.23%	4.51%	4.49%	5.11%	4.35%
Total return	5.10%	5.43%	4.77%	4.66%	5.28%	4.64%
Random	3.82%	4.23%	3.75%	3.38%	4.01%	3.51%

Table XI reports yearly returns for the best quintile portfolios. Panel A uses the portfolios based on linear combinations of the factors. Panel B uses the portfolios based on the linear combinations of the reduced factors, as represented by the most correlated performance measure with each factor. Panel C presents results for quintile portfolios based on classical measures or randomly constructed. Results are reported for no penalty when a fund delists, and for a penalty of 5%. Subsamples are built according to the last digit of the ISIN code (odd or even).

Table XII. Descriptive statistics and p-values for the best rebalanced quintile portfolios of 20 funds

	Descriptive statistics				p-values							
	Mean	S.D.	Skew	Kurt	p-full	p-red$_a$	p-red$_b$	p-M2	p-M3	p-SR	p-TR	p-rand
$F2_{30}\_F4_{30}\_F5_{40}$(full)	6.22%	0.96%	-0.14	-0.64	49%	61%	60%	96%	96%	89%	85%	100%
$F2_{30}\_F4_{30}\_F5_{40}$(red$_a$)	6.01%	0.90%	0.04	-0.42	43%	45%	46%	97%	97%	86%	85%	100%
$F2_{40}\_F4_{20}\_F5_{40}$(red$_b$)	6.06%	0.88%	-0.41	-0.29	49%	58%	56%	95%	95%	88%	83%	98%
MRAR2	4.48%	0.86%	0.13	-0.23	3%	4%	4%	53%	53%	28%	21%	73%
MRAR3	4.48%	0.70%	0.12	0.23	2%	3%	3%	55%	56%	25%	15%	81%
Sharpe	4.94%	0.75%	0.19	0.20	5%	8%	6%	72%	72%	50%	42%	94%
Total return	5.10%	0.78%	-0.13	2.09	5%	7%	7%	83%	83%	62%	53%	94%
Random	3.82%	0.89%	0.07	-0.24	0%	1%	1%	28%	28%	12%	8%	50%

Table XII reports the mean, standard deviation, skewness and kurtosis of yearly returns for the best quintile portfolios of selected measures (first column). The right part of the Table reports, for each row, the percentages of quintiles portfolios Q1 built with 20 funds according to the selected measure whose return is higher than the mean of the Q1 portfolios built with 20 funds according to the measure in the corresponding column. Results are reported for no penalty when a fund delists.

Table XIII. Subsamples based on size and age: P-values for the best rebalanced quintile portfolios of 20 funds

		MRAR2	MRAR3	Random	Sharpe	Total Return
Oldest	$F2_{30}\_F4_{30}\_F5_{40}$(full)	0.99	0.99	1	0.99	0.81
	$F2_{30}\_F4_{30}\_F5_{40}$(red$_a$)	0.97	0.99	1	0.97	0.72
	$F2_{40}\_F4_{20}\_F5_{40}$(red$_b$)	0.9	0.94	0.99	0.9	e0.65
Intermediate age	$F2_{30}\_F4_{30}\_F5_{40}$(full)	0.98	0.97	1	0.88	0.83
	$F2_{30}\_F4_{30}\_F5_{40}$(red$_a$)	0.99	0.99	1	0.87	0.8
	$F2_{40}\_F4_{20}\_F5_{40}$(red$_b$)	1	0.98	1	0.86	0.78
Largest	$F2_{30}\_F4_{30}\_F5_{40}$(full)	1	1	1	0.91	0.75
	$F2_{30}\_F4_{30}\_F5_{40}$(red$_a$)	0.98	0.99	0.99	0.86	0.71
	$F2_{40}\_F4_{20}\_F5_{40}$(red$_b$)	0.99	0.99	0.99	0.86	0.74
Midsize	$F2_{30}\_F4_{30}\_F5_{40}$(full)	0.98	0.93	0.97	0.87	0.9
	$F2_{30}\_F4_{30}\_F5_{40}$(red$_a$)	0.99	0.96	0.99	0.88	0.9
	$F2_{40}\_F4_{20}\_F5_{40}$(red$_b$)	0.94	0.91	0.93	0.87	0.89
Smallest	$F2_{30}\_F4_{30}\_F5_{40}$(full)	1	1	1	1	0.91
	$F2_{30}\_F4_{30}\_F5_{40}$(red$_a$)	0.97	1	1	0.97	0.82
	$F2_{40}\_F4_{20}\_F5_{40}$(red$_b$)	0.99	1	1	0.98	0.84

Table XIII reports, for each row, the percentages of quintiles portfolios Q1 built with 20 funds according to the selected measure whose return is higher than the mean of the Q1 portfolios built with 20 funds according to the measure in the corresponding column. Results are reported for no penalty when a fund delists, for five subsamples of the universe of funds: oldest, intermediary age, largest, intermediary size, smallest.

Figure 1. Explained variance due to factors
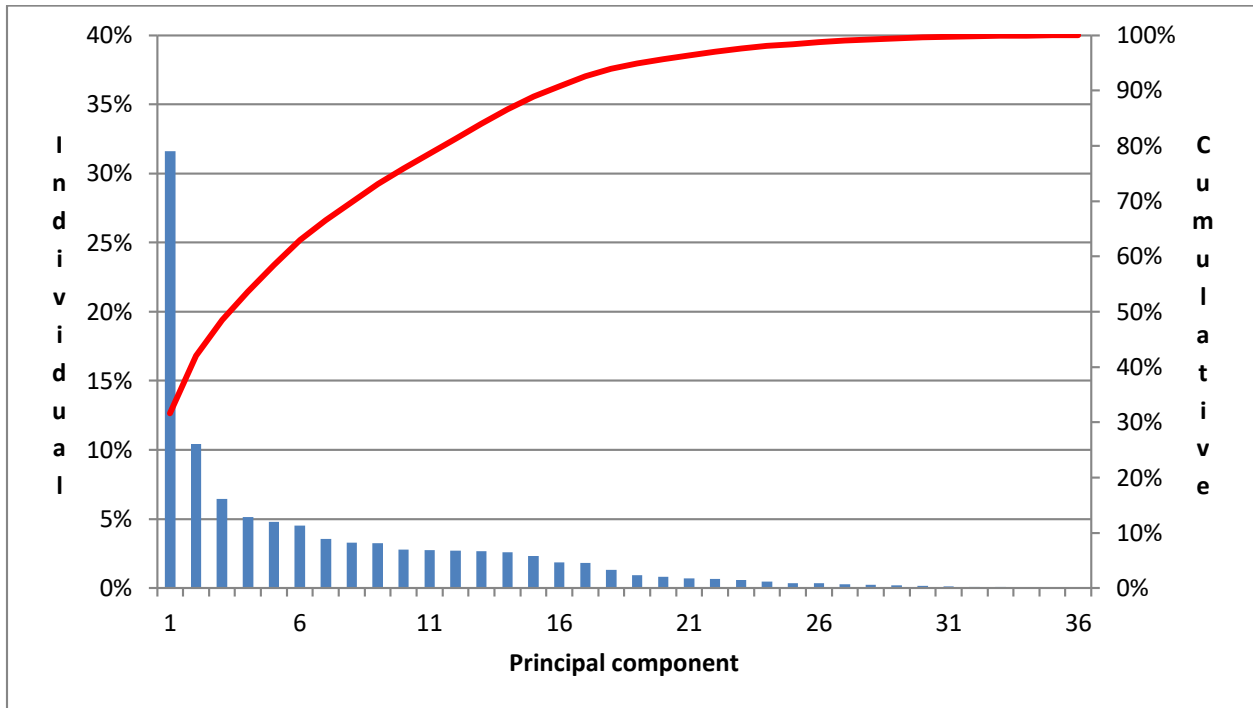
Figure 1 shows the proportion of explained variance by each of the factor issued of the principal component analysis. The bar chart with values scaled on the left represents the values per component, while the curve scaled on the right displays the cumulative.

Figure 2. Explained variance due to factors, after Varimax rotation
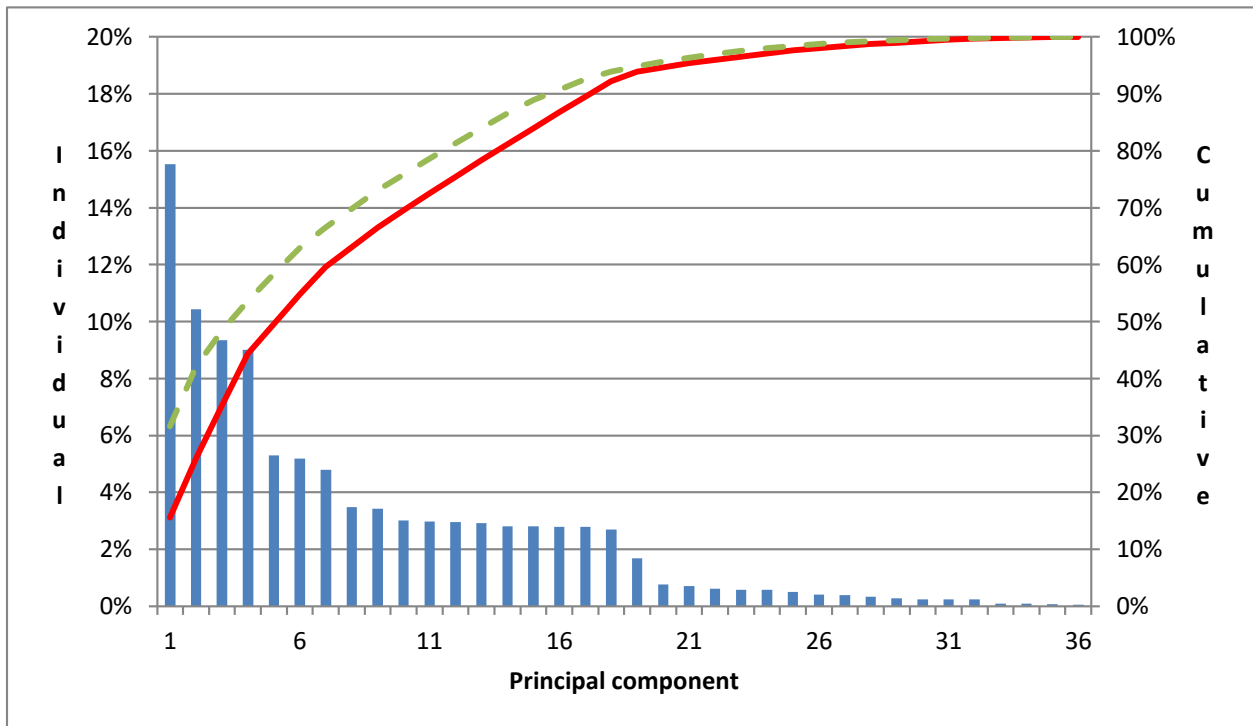
Figure 2 shows the proportion of explained variance by each of the factor issued of the principal component analysis after Varimax rotation. The bar chart with values scaled on the left represents the values per component, while the curve scaled on the right displays the cumulative. For comparison purpose, the dashed curve corresponds to the cumulative explained variance before Varimax rotation.

Figure 3. Dendrogram of the 147 computed measures

Clustering of   147 measures
Performance duration is 52 weeks
Selections: currency = / country =
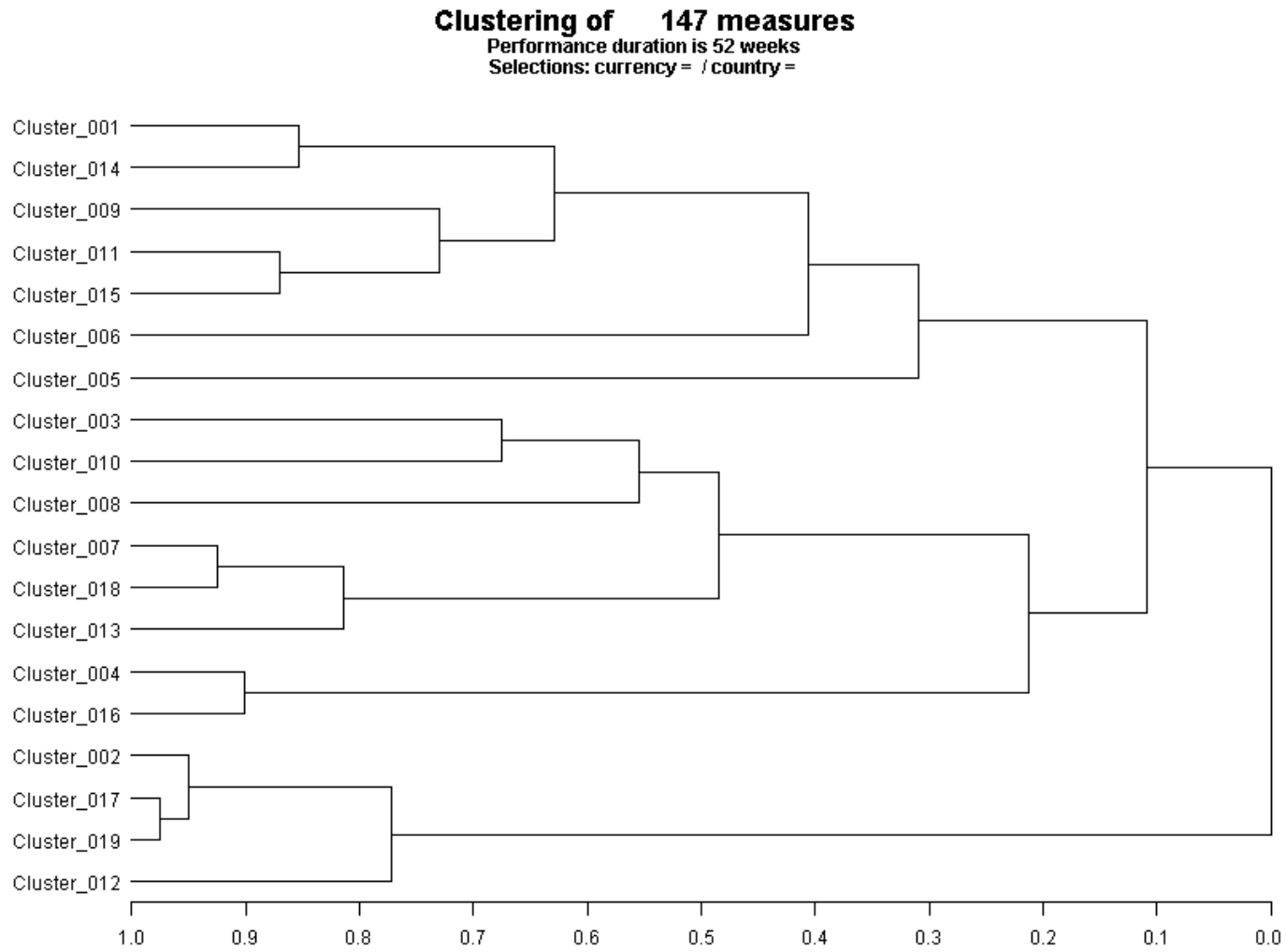
Note. This Figure shows the different steps in the clustering process for performance measures computed on a time horizon of one year. The considered sample contains all funds. The process starts with all measures in one cluster, and then it repeats the following steps:

1. A cluster is chosen for splitting; the selected cluster has the largest eigenvalue associated with the second principal component.

2. The chosen cluster is split into two clusters by finding the first two principal components, performing an orthoblique rotation, and assigning each variable to the rotated component with which it has the higher squared correlation.

3. Variables are iteratively reassigned to clusters, to try to maximize the variance accounted for by the cluster components.

The process stops splitting when every cluster has only one eigenvalue greater than one.

The composition of the 19 final buckets is as follows (the measures are labeled by their acronym - see Annex 2 for the full name):

Cluster_001 (54 measures): Adj_Skew_Sharpe_3, Bernardo_Ledoit_ifl, Bernardo_Ledoit_rf, Bernardo_Ledoit_zro, Burke_3, Calmar, Downsd_risk_Sharpe, Far_Tib_ifl_1_5x1, Far_Tib_rf_1_5x1, Gini, Isr_Roy_ifl, Isr_Roy_zro, Isr_Sharpe_ratio, MAD, Minimax, Prosp_rat_ifl_1, Prosp_rat_rf_1, Psp_S_K_rt_ifl_1, Psp_S_K_rt_ifl_2v25, Psp_S_K_rt_rf_1, Psp_S_K_rt_rf_2v25, RewVaR_rf_v05_v01, RewVaR_rf_v1_v01, Roy_ifl, Roy_zro, Rwd_to_half_variance, Sharpe_SK, Sharpe_ratio, Shp_VaRCF_ifl_v05, Shp_VaRCF_rf_v05, Shp_VaRCF_zro_v05, Shp_VaR_ifl_v1_v01, Shp_VaR_rf_v1_v01, Shp_VaR_zro_v1_v01, Sortino_SK_ifl, Sortino_SK_rf, Sortino_SK_zro, Sortino_Sat_ifl_1, Sortino_Sat_ifl_2, Sortino_Sat_ifl_3, Sortino_Sat_ifl_5, Sortino_Sat_rf_1, Sortino_Sat_rf_2, Sortino_Sat_rf_3, Sortino_Sat_rf_5, Sortino_Sat_zro_1, Sortino_Sat_zro_2, Sortino_Sat_zro_3, Sortino_Sat_zro_5, Sortino_ifl, Sortino_rf, Sortino_zro, Sterling_Calmar_3, eSDAR

Cluster_002 (6 measures): Far_Tib_ifl_2x2, Far_Tib_ifl_2x3, Far_Tib_rf_2x2, Far_Tib_rf_2x3, Upsd_pot_ratio_ifl, Upsd_pot_ratio_rf

Cluster_003 (12 measures): M2, M2_Sortino_rf, Martin, MorningStar_1, MorningStar_2, MorningStar_3, Sharpe_Alpha_1, Sharpe_Alpha_2, Sharpe_Alpha_3, Stdzd_Infor_Ratio_2, Stdzd_Infor_Ratio_4, Sterling

Cluster_004 (6 measures): Hnr_Mrt_3_f_alpha, Hnr_Mrt_3_f_gam_rm, Hnr_Mrt_alpha, Hnr_Mrt_gamma, Trn_Maz_cub_gamma, Trn_Maz_gamma

Cluster_005 (5 measures): Rv_ifl_v01_v5_v2, Rv_rf_v01_v2_v05, Rv_rf_v01_v5_v05, Rv_rf_v01_v5_v2, Rv_rf_v01_v5_v5

Cluster_006 (6 measures): Fouse_ifl_1, Fouse_ifl_2, Fouse_ifl_3, Fouse_rf_1, Fouse_rf_2, Fouse_rf_3

Cluster_007 (10 measures): Alpha_mkt_tim_HM, Alpha_mkt_tim_TM, Carhart_alpha, Fama_French_alpha, Gen_Bla_Trn_alpha, Stdzd_Infor_Ratio_1, Stdzd_Infor_Ratio_3, Total_risk_alpha, Trn_Maz_alpha, Trn_Maz_cub_alpha

Cluster_008 (4 measures): MRAP, Mod_Treynor, Modified_Jensen, Treynor

Cluster_009 (11 measures): Rv_ifl_v01_v05_v05, Rv_rf_v01_v05_v05, Shp_CVaR_ifl_v05_v01, Shp_CVaR_ifl_v1_v01, Shp_CVaR_rf_v05_v01, Shp_CVaR_rf_v1_v01, Shp_CVaR_zro_v05_v01, Shp_CVaR_zro_v1_v01, Shp_VaR_ifl_v05_v01, Shp_VaR_rf_v05_v01, Shp_VaR_zro_v05_v01

Cluster_010 (3 measures): Alpha_TM_cond_beta, Alpha_cond_beta, Jensen_alpha

Cluster_011 (8 measures): Rv_avg_dup_ddwn_ifl, Rv_avg_dup_ddwn_rf, Rv_avg_dup_ddwn_zro, Rv_max_dup_ddwn_ifl, Rv_max_dup_ddwn_rf, Rv_max_dup_ddwn_zro, Stutzer_ifl, Stutzer_rf

Cluster_012 (3 measures): Aftal_Ponc_2, Aftal_Ponc_3, Hnr_Mrt_3_f_gam_hml

Cluster_013 (4 measures): Information_Ratio, Isr_Infor_Ratio, SRAP, Semi_Var_Infor_Ratio

Cluster_014 (6 measures): Prosp_rat_ifl_2v25, Prosp_rat_ifl_5, Prosp_rat_rf_2v25, Prosp_rat_rf_5, Psp_S_K_rt_ifl_5, Psp_S_K_rt_rf_5

Cluster_015 (3 measures): Shp_VaRCF_ifl_v2, Shp_VaRCF_rf_v2, Shp_VaRCF_zro_v2

Cluster_016 (2 measures): Gamma_TM_cond_beta, Trn_Maz_cub_delta

Cluster_017 (2 measures): Alpha_cond_alpha, Hwang_Satchell

Cluster_018 (1 measures): Hnr_Mrt_3_f_gam_smb

Cluster_019 (1 measures): Moses_Cheney_Veit

Figure 4. Proportions of the ex-ante quintile portfolios in the 50 best and 50 worst ex-post portfolios, for six classical measures
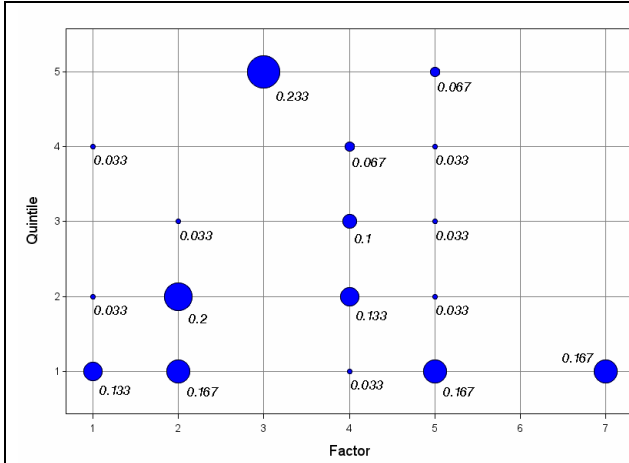

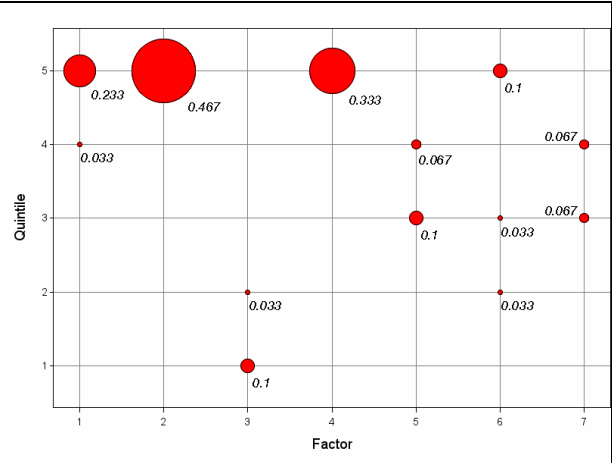Figure 4.a. Sharpe ratio - Top 50 portfolios

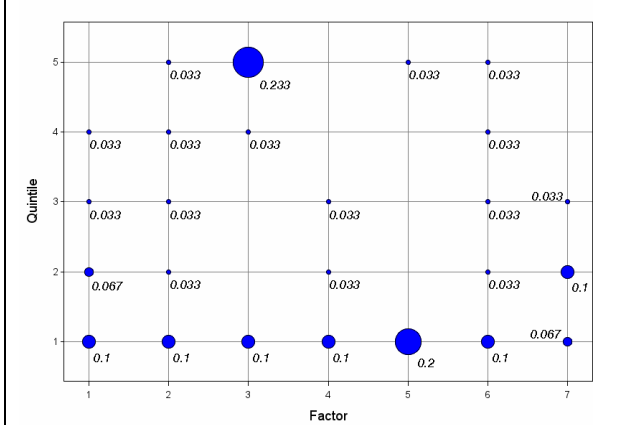Figure 4.b. Sharpe ratio - Bottom 50 portfolios

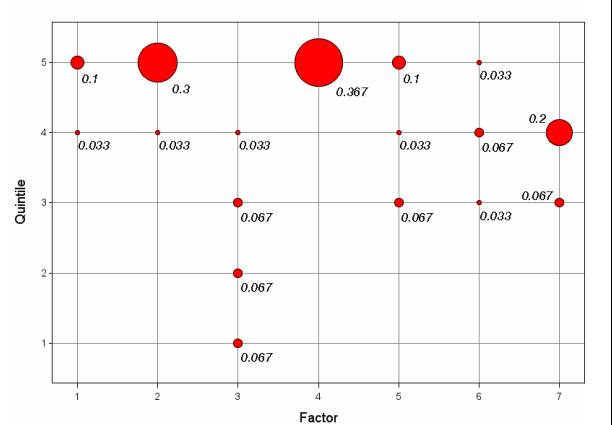Figure 4.c. Treynor ratio - Top 50 portfolios
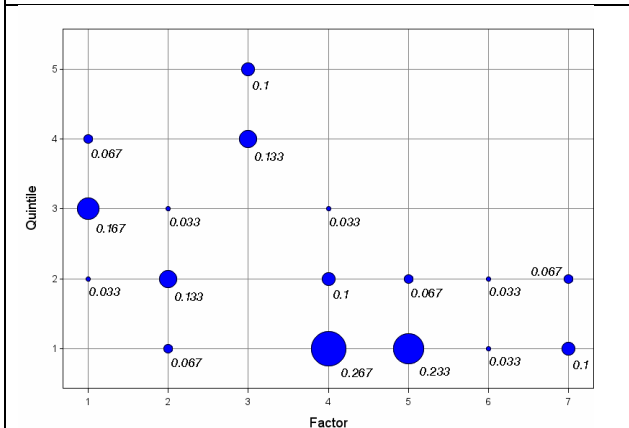
Figure 4.d. Treynor ratio - Bottom 50 portfolios

Figure 4.e. Total return - Top 50 portfolios

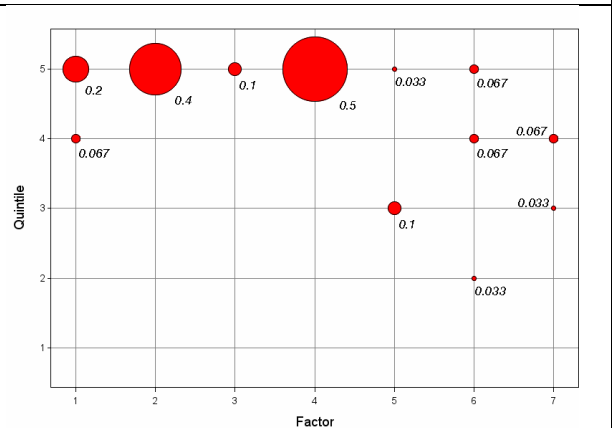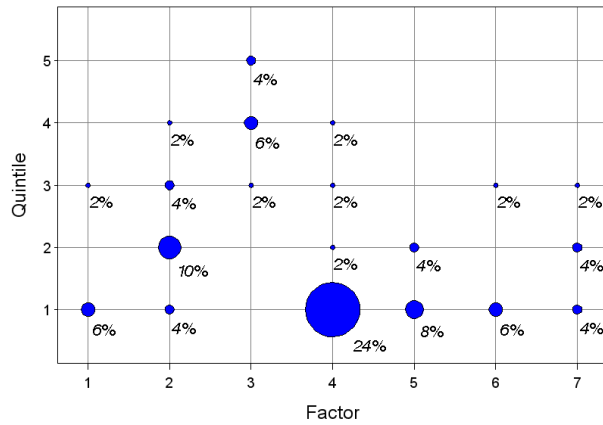Figure 4.f. Total return - Bottom 50 portfolios
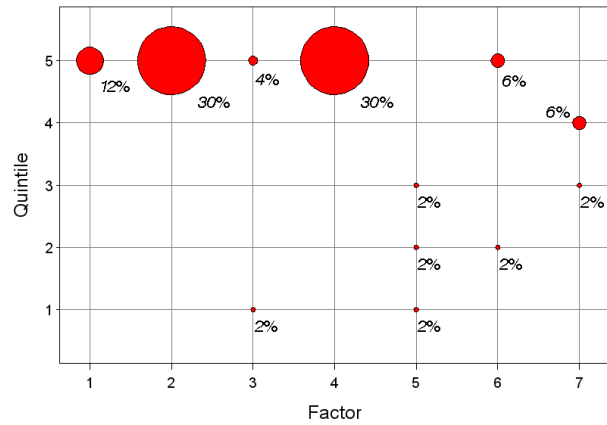
Figure 4.g. MRAR2 - Top 50 portfolios

Figure 4.h. MRAR2 - Bottom 50 portfolios
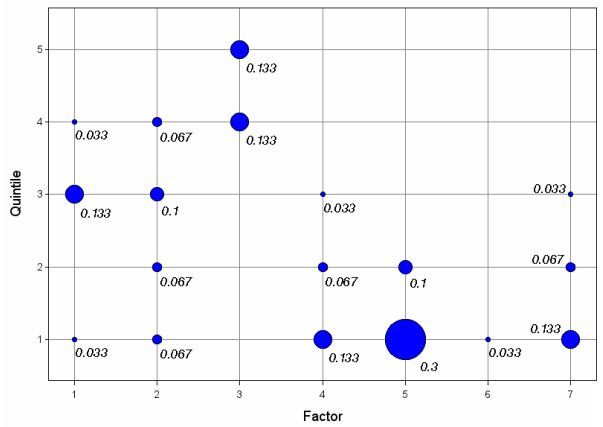
Figure 4.i. Jensen's alpha - Top 50 portfolios
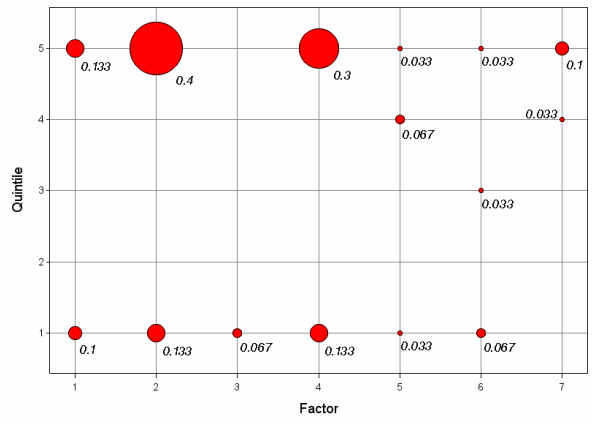
Figure 4.j. Jensen's alpha - Bottom 50 portfolios
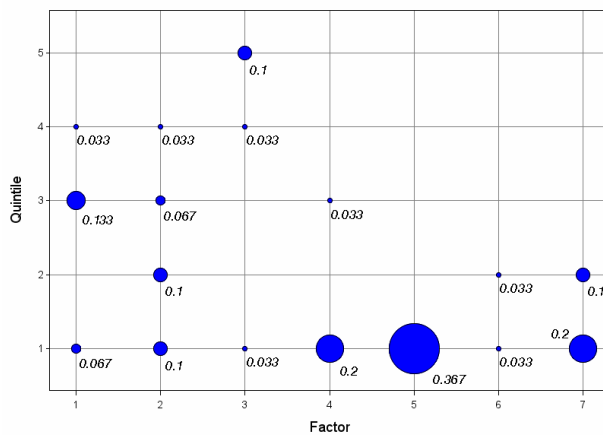
Figure 4.k. Fama & French alpha - Top 50 portfolios
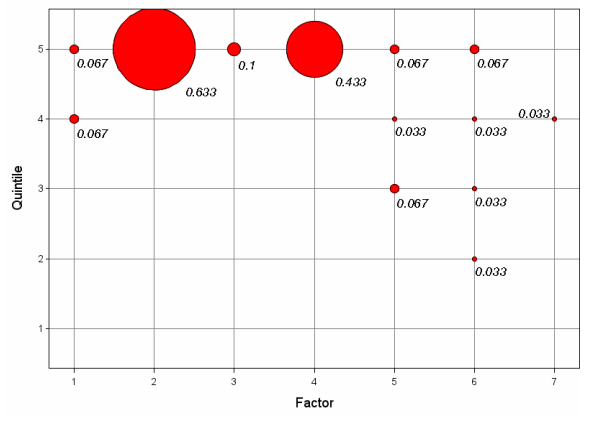
Figure 4.l. Fama & French alpha - Bottom 50 portfolios

Figures 4.a to 4.l report the proportions of occurrences of each ex-ante quintile portfolio according to the first 7 factors, in the Top-50, then in the Bottom-50 ex-post portfolios for the selected classical measures, when performance is computed over a duration of one year.

Figure 5. Yearly returns of best quintiles portfolios built with combinations of factors
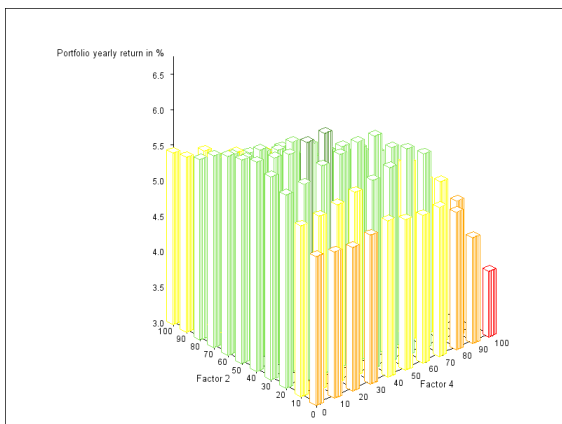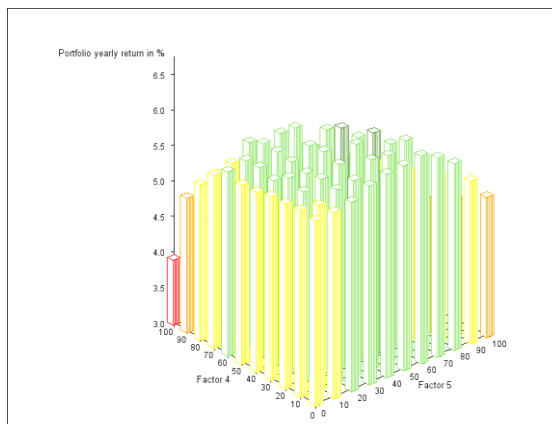

Figure 5.a. Combinations F2-F4

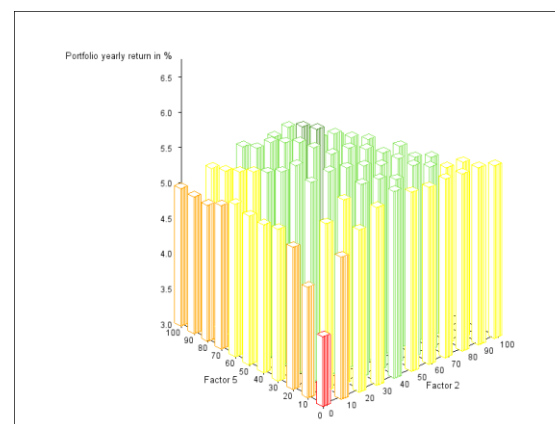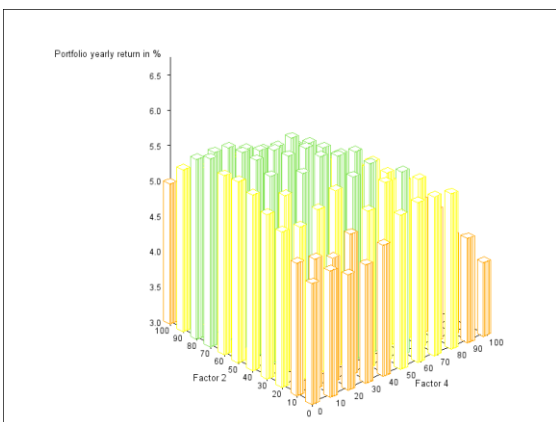Figure 5.b. Combinations F4-F5

Figure 5.c. Combinations F5-F2

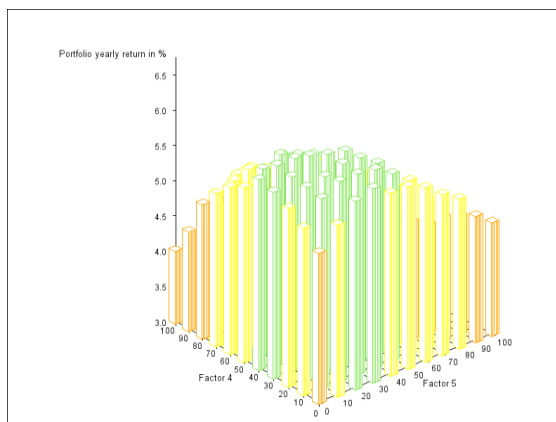Figure 5.d. Combinations restr. F2-F4
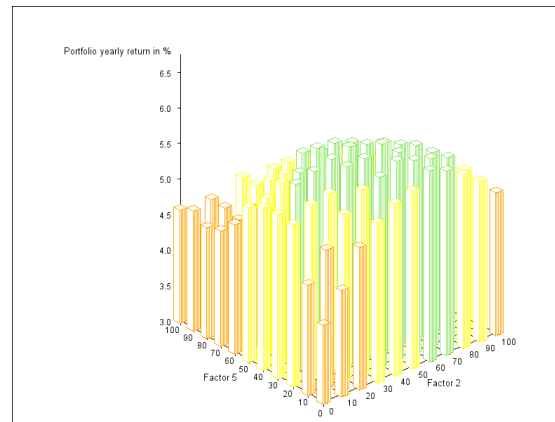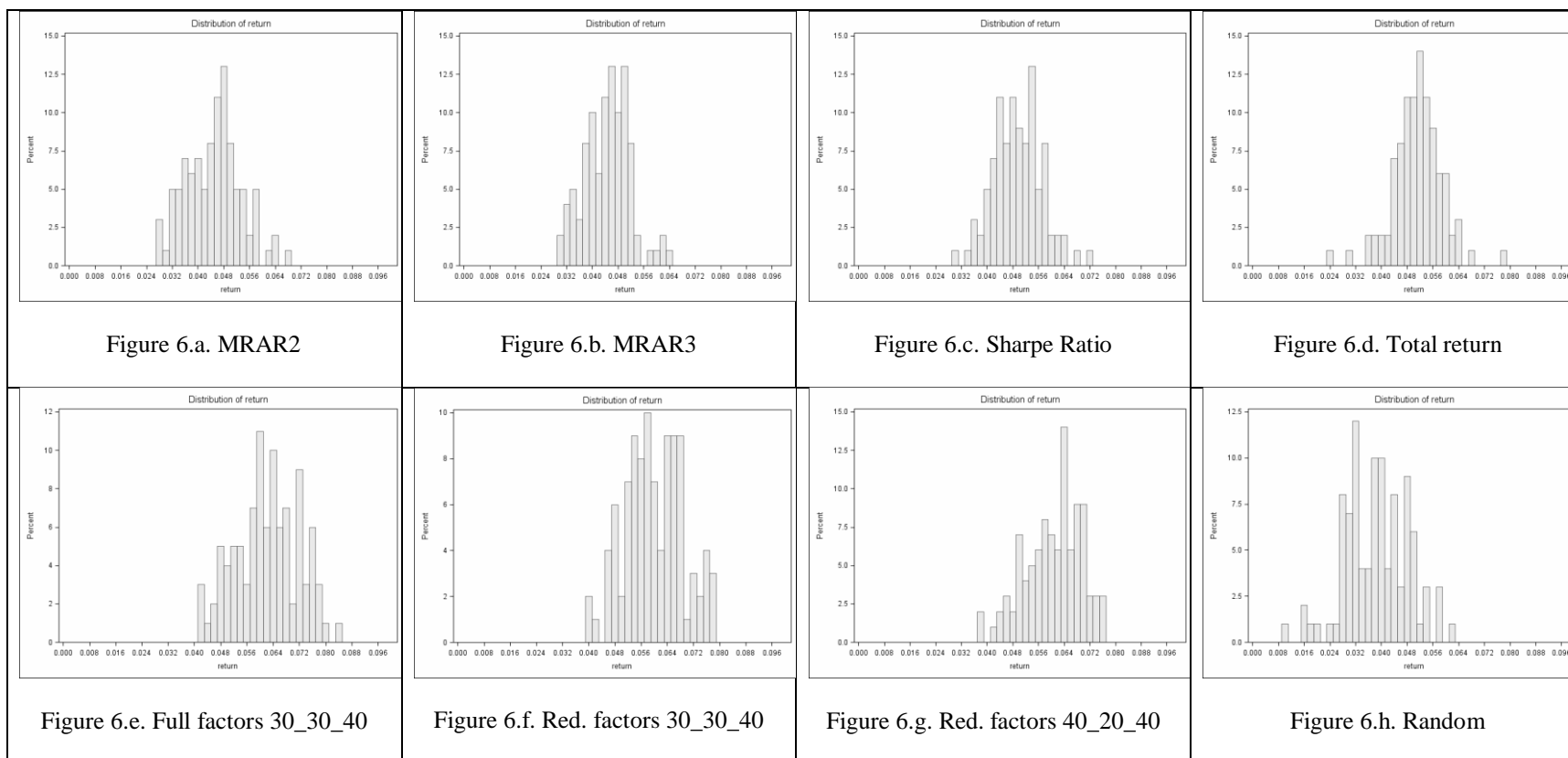

Figure 5.e. Combinations restr. F4-F5

Figure 5.f. Combinations restr. F5-F2

Figures 5.a to 5.f report the return of rebalanced portfolios created on basis of performance measures built using factors 2, 4 and 5 of the principal component analysis. Weights are multiple of 10% and the horizontal axes are the weights on two of the three factors (the weight of the third factor is 100 minus the weights of the reported factors). The reported return is for the portfolio corresponding to the best quintile. Figures 5.a to 5.c correspond to factor considered integrally. Figures 5.d to 5.f correspond to factors reduced to their most significant measure, as described in 5.4. The penalty when a fund closes is set to 5%.

Figure 6. Distributions of returns for the best quintile portfolios built with 20 funds.

Figure 6.a. MRAR2	Figure 6.b. MRAR3	Figure 6.c. Sharpe Ratio	Figure 6.d. Total return
Figure 6.e. Full factors 30_30_40	Figure 6.f. Red. factors 30_30_40	Figure 6.g. Red. factors 40_20_40	Figure 6.h. Random

Figures 6.a to 6.g report the distributions of returns for the best quintile portfolios according to the series of following measures: MRAR with γ equal to 2 and 3, Sharpe Ratio, total return, combination of full factors with overweighting factor 5, combination of reduced factors with overweighting factor 5, and a combination of reduced factors with overweighting factor 4. Figure 6.h reports the distribution when the selection is random. Quintile portfolios are built selecting 20 funds

from the set of funds constituting the performance quintile on the last three years. This selection of 20 funds is done randomly, but respecting the proportions of

funds per benchmark. To randomize the process, we repeat the procedure by rebalancing 100 times on the whole sample of funds.

Appendix 1: List of indexes used for determination of the benchmark

a) Eighteen indexes per country:
   - Eight European indexes: AEX for Netherlands, BEL-20 for Belgium, CAC40 for France, DAX for Germany, FTSE for United Kingdom, IBEX35 for Spain, MIBTEL for Italy and SMI for Helvetia.
   - Four American indexes: Dow Jones, Nasdaq, S&P 500 and MorningStar
   - Nikkei-225 for Japan
   - Five emerging markets indexes: BOVE for Brazil, FTSE Bursa Malaysia for Malaysia, Hang Seng for Hong Kong, Indonesia SEI for Indonesia, SGX for Singapore
b) Five indexes per region:
   - Euro Stoxx 50 for the Euro zone
   - MSCI and MorningStar indexes for Europe
   - MSCI for North America
   - MSCI for World
c) Nine indexes per investment style:
   - Small companies: RUSSEL 2000 for USA, MSCI index for worldwide
   - Small and mid-companies: MorningStar index for worldwide
   - Mid companies: MSCI index for Europe, MorningStar for USA and MSCI index for worldwide
   - Large companies: MSCI index for Europe, MorningStar for USA and MSCI index for worldwide
d) Six sectorial worldwide indexes:
   - Banks (ticker: CBNK)
   - Energy companies (ticker: MXWO0EN)
   - IT companies (ticker: MXWO0IT)
   - IT services companies (ticker: IXK)
   - Real Estate companies (ticker: BWREAL)
   - Retail companies (ticker: MXWD0RT)
e) Two commodities indexes:
   - Market Crude Oil
   - Gold

Appendix 2: List of performance measures

Acronym	Full name	Parameters	Class	Initial reference	Section	Select
Adj_Skew_Sharpe_3	Adjusted for Skewness Sharpe Ratio (ASSR)		Return-based Ratio	Zakamouline V. and Koekebakker S. (2009), *JBF*	3_1_1_4	N
Aftal_Ponc_2	Aftalion and Poncet's index	Price of risk = 2	Return-based Difference	Aftalion F. and Poncet P. (1991), *RB*	4_1_1_6	Y
Aftal_Ponc_3	Aftalion and Poncet's index	Price of risk = 3	Return-based Difference	id.	4_1_1_6	N
Alpha_cond_alpha	Conditional alpha		Return-based Difference	Christopherson J., Ferson W. and Turner A. (1999), *JPM*	4_2_2_6	N
Alpha_cond_beta	Alpha with conditional betas		Return-based Difference	Ferson W. and Schadt R. (1996), *JoF*	4_2_2_5	N
Alpha_mkt_tim_HM	Market timing alpha, according to Henriksson and Merton's model		Market timing	Hübner G. (2011), *WP*		N
Alpha_mkt_tim_TM	Market timing alpha, according to Treynor and Mazuy's model		Market timing	id.		Y
Alpha_TM_cond_beta	Alpha in Treynor and Mazuy's conditional market timing model		Market timing	Ferson W. and Schadt R. (1996), *JoF*	6_2_3_1	Y
Bernardo_Ledoit_ifl	Bernardo-Ledoit gain loss ratio, or Omega	Reserve return = inflation rate	Return-based Ratio	Bernardo A. and Ledoit O. (2000), *JPE*; Keating C. and Shadwick W. (2002), *JPMr*	3_1_3_1_1	N
Bernardo_Ledoit_rf	Bernardo-Ledoit gain loss ratio, or Omega	Reserve return = risk-free rate	Return-based Ratio	id.	3_1_3_1_1	N
Bernardo_Ledoit_zro	Bernardo-Ledoit gain loss ratio, or Omega	Reserve return = zero percent	Return-based Ratio	id.	3_1_3_1_1	N
Burke_3	Burke ratio		Return-based Ratio	Burke G. (1994), *Fu*	3_1_3_3_4_3	N
Calmar	Calmar ratio		Return-based Ratio	Young T. (1991), *Fu*	3_1_3_3_1	N
Carhart_alpha	Alpha based on Carhart's four factors model		Return-based Difference	Carhart M. (1997), *JoF*	4_2_2_2	N
Downsd_risk_Sharpe	Downside-risk Sharpe ratio		Return-based Ratio	Ziemba W. (2005), *JPM*	3_1_2_1_2	N
eSDAR	Excess standard deviation adjusted return (eSDAR)		Return-based Difference	Statman M. (1987), *JFQA*	4_1_1_5	Y

Acronym	Full name	Parameters	Class	Initial reference	Section	Select
Fama_French_alpha	Alpha based on Fama and French's three factors model		Return-based Difference	Fama E. and French K. (1992), *JoF*, Fama E. and French K. (1993), *JFE*	4_2_2_1	Y
Far_Tib_ifl_1_5x1	Farinelli-Tibiletti ratio	Reserve return = inflation rate, order of upper partial moment = 1.5, order of lower partial moment = 1	Return-based Ratio	Farinelli S. and Tibiletti L. (2008), *EJOR*	3_1_3_1_4	N
Far_Tib_ifl_2x4	Farinelli-Tibiletti ratio	Reserve return = inflation rate, order of upper partial moment = 2, order of lower partial moment = 2	Return-based Ratio	id.	3_1_3_1_4	Y
Far_Tib_ifl_2x9	Farinelli-Tibiletti ratio	Reserve return = inflation rate, order of upper partial moment = 2, order of lower partial moment = 3	Return-based Ratio	id.	3_1_3_1_4	N
Far_Tib_rf_1_5x1	Farinelli-Tibiletti ratio	Reserve return = risk-free rate, order of upper partial moment = 1.5, order of lower partial moment = 1	Return-based Ratio	id.	3_1_3_1_4	N
Far_Tib_rf_2x4	Farinelli-Tibiletti ratio	Reserve return = risk-free rate, order of upper partial moment = 2, order of lower partial moment = 2	Return-based Ratio	id.	3_1_3_1_4	N
Far_Tib_rf_2x9	Farinelli-Tibiletti ratio	Reserve return = risk-free rate, order of upper partial moment = 2, order of lower partial moment = 3	Return-based Ratio	id.	3_1_3_1_4	N
Fouse_ifl_1	Fouse's index	Reserve return = inflation rate, coefficient of aversion to risk = 1	Gain-based Difference	Sortino F. and Price L. (1994), *JoI*	4_3_3_1	N
Fouse_ifl_2	Fouse's index	Reserve return = inflation rate, coefficient of aversion to risk = 2	Gain-based Difference	id.	4_3_3_1	N
Fouse_ifl_3	Fouse's index	Reserve return = inflation rate, coefficient of aversion to risk = 3	Gain-based Difference	id.	4_3_3_1	N
Fouse_rf_1	Fouse's index	Reserve return = risk-free rate, coefficient of aversion to risk = 1	Gain-based Difference	id.	4_3_3_1	Y
Fouse_rf_2	Fouse's index	Reserve return = risk-free rate, coefficient of aversion to risk = 2	Gain-based Difference	id.	4_3_3_1	N
Fouse_rf_3	Fouse's index	Reserve return = risk-free rate, coefficient of aversion to risk = 3	Gain-based Difference	id.	4_3_3_1	N

Acronym	Full name	Parameters	Class	Initial reference	Section	Select
Gamma_TM_cond_beta	Conditional Treynor and Mazuy's coefficient		Market timing	Ferson W. and Schadt R. (1996), *JoF*	6_2_3_1	Y
Gen_Bla_Trn_alpha	Generalized Black-Treynor ratio		Return-based Ratio	Hübner G. (2005), *RoF*	3_2_2_2	Y
Gini	Gini ratio		Return-based Ratio	Yitzhaki S. (1982), *AER*	3_1_2_3_2	N
Hnr_Mrt_3_f_alpha	Alpha in Henriksson and Merton model with a three-factor context		Market timing	Chan L., Chen H.L. and Lakonishok J. (2002), *RFS*	6_2_2_4	Y
Hnr_Mrt_3_f_gam_hml	Gamma related to factor hml in Henriksson and Merton model with a three-factor context		Market timing	id.	6_2_2_4	Y
Hnr_Mrt_3_f_gam_rm	Gamma related to market in Henriksson and Merton model with a three-factor context		Market timing	id.	6_2_2_4	Y
Hnr_Mrt_3_f_gam_smb	Gamma related to factor smb in Henriksson and Merton model with a three-factor context		Market timing	id.	6_2_2_4	Y
Hnr_Mrt_alpha	Alpha in Henriksson and Merton's market timing model		Market timing	Henriksson R. and Merton R. (1981), *JB*	6_1_2_1	Y
Hnr_Mrt_gamma	Henriksson and Merton's coefficient		Market timing - gamma	id.	6_1_2_1	N
Hwang_Satchell	Higher moment measure of Hwang and Satchell		Return-based Difference	Hwang S. and Satchell S. (1999), *IJFE*	4_2_2_8	Y
Information_Ratio	Information ratio		Return-based Ratio	Treynor J. and Black F. (1973), *JB*	3_3_2_1	N
Isr_Infor_Ratio	Israelsen's modified information ratio		Return-based Ratio	Israelsen C. (2005), *JAM*	3_3_2_2	Y
Isr_Roy_ifl	Israelsen's modified Roy's measure	Reserve return = inflation rate	Return-based Ratio	id.	3_1_1_2	N
Isr_Roy_zro	Israelsen's modified Roy's measure	Reserve return = zero percent	Return-based Ratio	id.	3_1_1_2	N
Isr_Sharpe_ratio	Israelsen's modified Sharpe ratio		Return-based Ratio	id.	3_1_1_2	N
Jensen_alpha	Jensen's alpha		Return-based Difference	Jensen M. (1968), *JoF*	4_2_1_1_1	Y

Acronym	Full name	Parameters	Class	Initial reference	Section	Select
M2	M² index, or risk-adjusted performance (RAP)		Return-based Difference	Modigliani F. and Modigliani L. (1997), *JPM*	4_1_1_1	N
M2_Sortino_rf	M² for Sortino		Return-based Difference	Bacon C. (2008), *book*	4_1_1_2	N
MAD	Mean absolute deviation (MAD) ratio		Return-based Ratio	Konno H. and Yamazaki H. (1991), *MS*	3_1_2_3_1	N
Martin	Martin ratio or Ulcer performance index		Return-based Ratio	Martin P. and McCann B. (1989), *book*	3_1_2_3_4	N
Minimax	Minimax		Return-based Ratio	Young M. (1998), *MS*	3_1_2_3_3	N
Mod_Treynor	Modified Treynor ratio		Return-based Ratio	Bacon C. (2008), *book*	3_2_1_2	Y
Modified_Jensen	Modified Jensen		Return-based Ratio	Smith K. and Tito D. (1969), *JFQA*	3_2_2_1	N
MorningStar_1	Morningstar risk adjusted return (MRAR)	Relative aversion to risk coefficient = 1	Preference based	MorningStar (2007), *WP*	5_1_1_4	N
MorningStar_2	Morningstar risk adjusted return (MRAR)	Relative aversion to risk coefficient = 2	Preference based	id.	5_1_1_4	N
MorningStar_3	Morningstar risk adjusted return (MRAR)	Relative aversion to risk coefficient = 3	Preference based	id.	5_1_1_4	N
Moses_Cheney_Veit	Moses, Cheney and Veit's measure		Return-based Ratio	Moses E., Cheyney J. and Veit T. (1987), *JPM*	3_3_1_1	Y
MRAP	Market risk-adjusted performance (MRAP)		Return-based Difference	Scholz H. and Wilkens M. (2005), *JPMr*	4_1_1_3	N
Prosp_rat_ifl_1	Prospect ratio	Reserve return = inflation rate, extent of the loss aversion = 1	Preference based	Watanabe Y. (2006), *JPMr*	5_1_3_1	N
Prosp_rat_ifl_2v25	Prospect ratio	Reserve return = inflation rate, extent of the loss aversion = 2.25	Preference based	id.	5_1_3_1	N
Prosp_rat_ifl_5	Prospect ratio	Reserve return = inflation rate, extent of the loss aversion = 5	Preference based	id.	5_1_3_1	N
Prosp_rat_rf_1	Prospect ratio	Reserve return = risk-free rate, extent of the loss aversion = 1	Preference based	id.	5_1_3_1	N
Prosp_rat_rf_2v25	Prospect ratio	Reserve return = risk-free rate, extent of the loss aversion = 2.25	Preference based	id.	5_1_3_1	N

Acronym	Full name	Parameters	Class	Initial reference	Section	Select
Prosp_rat_rf_5	Prospect ratio	Reserve return = risk-free rate, extent of the loss aversion = 5	Preference based	id.	5_1_3_1	N
Psp_S_K_rt_ifl_1	Prospect + Skewness / Kurtosis	Reserve return = inflation rate, extent of the loss aversion = 1	Preference based	id.	5_1_3_2	N
Psp_S_K_rt_ifl_2v25	Prospect + Skewness / Kurtosis	Reserve return = inflation rate, extent of the loss aversion = 2.25	Preference based	id.	5_1_3_2	N
Psp_S_K_rt_ifl_5	Prospect + Skewness / Kurtosis	Reserve return = inflation rate, extent of the loss aversion = 5	Preference based	id.	5_1_3_2	N
Psp_S_K_rt_rf_1	Prospect + Skewness / Kurtosis	Reserve return = risk-free rate, extent of the loss aversion = 1	Preference based	id.	5_1_3_2	N
Psp_S_K_rt_rf_2v25	Prospect + Skewness / Kurtosis	Reserve return = risk-free rate, extent of the loss aversion = 2.25	Preference based	id.	5_1_3_2	N
Psp_S_K_rt_rf_5	Prospect + Skewness / Kurtosis	Reserve return = risk-free rate, extent of the loss aversion = 5	Preference based	id.	5_1_3_2	N
RewVaR_rf_v05_v01	Reward-to-VaR ratio	Reserve return = risk-free rate, threshold for VaR = 5%, computed according to the EVT with threshold = 1%	Return-based Ratio	Alexander G. and Baptista A. (2003), *JPM*	3_1_2_2_2	N
RewVaR_rf_v1_v01	Reward-to-VaR ratio	Reserve return = risk-free rate, threshold for VaR = 10%, computed according to the EVT with threshold = 1%	Return-based Ratio	id.	3_1_2_2_2	N
Roy_ifl	Roy's measure	Reserve return = inflation rate	Return-based Ratio	Roy A. (1952), Ec	3_1_1_9	N
Roy_zro	Roy's measure	Reserve return = zero percent	Return-based Ratio	id.	3_1_1_9	N
Rv_avg_dup_ddwn_ifl	Rachev average drawup/down ratio	Reserve return = inflation rate	Return-based Ratio	Ortobelli S., Biglova A., Rachev S. and Stoyanov S. (2009), *JAFA*	3_1_3_3_6	N
Rv_avg_dup_ddwn_rf	Rachev average drawup/down ratio	Reserve return = risk-free rate	Return-based Ratio	id.	3_1_3_3_6	Y
Rv_avg_dup_ddwn_zro	Rachev average drawup/down ratio	Reserve return = zero percent	Return-based Ratio	id.	3_1_3_3_6	N

Acronym	Full name	Parameters	Class	Initial reference	Section	Select
Rv_ifl_v01_v05_v05	Rachev ratio	Reserve return = inflation rate, threshold for CVaR are 5% for both numerator and denominator, computed according to the EVT with threshold = 1%	Return-based Ratio	Biglova A., Ortobelli S., Rachev S. and Stoyanov S. (2004), *JPM*	3_1_3_2_1	Y
Rv_ifl_v01_v5_v2	Rachev ratio	Reserve return = inflation rate, threshold for CVaR are 50% for numerator and 20% for denominator, computed according to the EVT with threshold = 1%	Return-based Ratio	id.	3_1_3_2_1	N
Rv_max_dup_ddwn_ifl	Rachev maximum drawup/down ratio	Reserve return = inflation rate	Return-based Ratio	Ortobelli S., Biglova A., Rachev S. and Stoyanov S. (2009), *JAFA*	3_1_3_3_5	N
Rv_max_dup_ddwn_rf	Rachev maximum drawup/down ratio	Reserve return = risk-free rate	Return-based Ratio	id.	3_1_3_3_5	Y
Rv_max_dup_ddwn_zro	Rachev maximum drawup/down ratio	Reserve return = zero percent	Return-based Ratio	id.	3_1_3_3_5	N
Rv_rf_v01_v05_v05	Rachev ratio	Reserve return = risk-free rate, threshold for CVaR are 5% for both numerator and denominator, computed according to the EVT with threshold = 1%	Return-based Ratio	Biglova A., Ortobelli S., Rachev S. and Stoyanov S. (2004), *JPM*	3_1_3_2_1	N
Rv_rf_v01_v2_v05	Rachev ratio	Reserve return = risk-free rate, threshold for CVaR are 20% for numerator and 5% for denominator, computed according to the EVT with threshold = 1%	Return-based Ratio	id.	3_1_3_2_1	Y
Rv_rf_v01_v5_v05	Rachev ratio	Reserve return = risk-free rate, threshold for CVaR are 50% for numerator and 5% for denominator, computed according to the EVT with threshold = 1%	Return-based Ratio	id.	3_1_3_2_1	N
Rv_rf_v01_v5_v2	Rachev ratio	Reserve return = risk-free rate, threshold for CVaR are 50% for numerator and 20% for denominator, computed according to the EVT with threshold = 1%	Return-based Ratio	id.	3_1_3_2_1	Y

Acronym	Full name	Parameters	Class	Initial reference	Section	Select
Rv_rf_v01_v5_v5	Rachev ratio	Reserve return = risk-free rate, threshold for CVaR are 50% for both the numerator and the denominator, computed according to the EVT with threshold = 1%	Return-based Ratio	id.	3_1_3_2_1	N
Rwd_to_half_variance	Reward to half-variance index		Return-based Ratio	Ang J. and Chua J. (1979), *JFQA*	3_1_2_1_1	N
Semi_Var_Infor_Ratio	Information ratio based on semi-variance		Return-based Ratio	Gillet Ph. And Moussavou J. (2000), *EIR*	3_3_2_3	N
Sharpe_Alpha_1	Sharpe's alpha	Coefficient of aversion to shortfall = 1	Gain-based Difference	Plantinga A. and De Groot S. (2001), *JPMr*	4_3_2_1	N
Sharpe_Alpha_2	Sharpe's alpha	Coefficient of aversion to shortfall = 2	Gain-based Difference	id.	4_3_2_1	N
Sharpe_Alpha_3	Sharpe's alpha	Coefficient of aversion to shortfall = 3	Gain-based Difference	id.	4_3_2_1	N
Sharpe_ratio	Sharpe ratio		Return-based Ratio	Sharpe W. (1966), *JB*	3_1_1_1	N
Sharpe_SK	Sharpe + Skewness / Kurtosis		Return-based Ratio	Watanabe Y. (2006), *JPMr*	3_1_1_6	N
Shp_CVaR_ifl_v05_v01	Sharpe ratio based on conditional VaR or STARR ratio	Reserve return = inflation rate, threshold for CVaR = 5%, computed according to the EVT with threshold = 1%	Return-based Ratio	Martin R., Rachev S., and Siboulet F. (2003), *Wi*	3_1_2_2_4	N
Shp_CVaR_ifl_v1_v01	Sharpe ratio based on conditional VaR or STARR ratio	Reserve return = inflation rate, threshold for CVaR = 10%, computed according to the EVT with threshold = 1%	Return-based Ratio	id.	3_1_2_2_4	N
Shp_CVaR_rf_v05_v01	Sharpe ratio based on conditional VaR or STARR ratio	Reserve return = risk-free rate, threshold for CVaR = 5%, computed according to the EVT with threshold = 1%	Return-based Ratio	id.	3_1_2_2_4	N
Shp_CVaR_rf_v1_v01	Sharpe ratio based on conditional VaR or STARR ratio	Reserve return = risk-free rate, threshold for CVaR = 10%, computed according to the EVT with threshold = 1%	Return-based Ratio	id.	3_1_2_2_4	N
Shp_CVaR_zro_v05_v01	Sharpe ratio based on conditional VaR or STARR ratio	Reserve return = zero percent, threshold for CVaR = 5%, computed according to the EVT with threshold = 1%	Return-based Ratio	id.	3_1_2_2_4	N

Acronym	Full name	Parameters	Class	Initial reference	Section	Select
Shp_CVaR_zro_v1_v01	Sharpe ratio based on conditional VaR or STARR ratio	Reserve return = zero percent, threshold for CVaR = 10%, computed according to the EVT with threshold = 1%	Return-based Ratio	id.	3_1_2_2_4	N
Shp_VaR_ifl_v05_v01	Sharpe ratio based on the Value at Risk	Reserve return = inflation rate, threshold for VaR = 5%, computed according to the EVT with threshold = 1%	Return-based Ratio	Dowd K. (1999), *JPM*; Dowd K. (2000), *IREF*	3_1_2_2_1	N
Shp_VaR_ifl_v1_v01	Sharpe ratio based on the Value at Risk	Reserve return = inflation rate, threshold for VaR = 10%, computed according to the EVT with threshold = 1%	Return-based Ratio	id.	3_1_2_2_1	N
Shp_VaR_rf_v05_v01	Sharpe ratio based on the Value at Risk	Reserve return = risk-free rate, threshold for VaR = 5%, computed according to the EVT with threshold = 1%	Return-based Ratio	id.	3_1_2_2_1	N
Shp_VaR_rf_v1_v01	Sharpe ratio based on the Value at Risk	Reserve return = risk-free rate, threshold for VaR = 10%, computed according to the EVT with threshold = 1%	Return-based Ratio	id.	3_1_2_2_1	N
Shp_VaR_zro_v05_v01	Sharpe ratio based on the Value at Risk	Reserve return = zero percent, threshold for VaR = 5%, computed according to the EVT with threshold = 1%	Return-based Ratio	id.	3_1_2_2_1	N
Shp_VaR_zro_v1_v01	Sharpe ratio based on the Value at Risk	Reserve return = zero percent, threshold for VaR = 10%, computed according to the EVT with threshold = 1%	Return-based Ratio	id.	3_1_2_2_1	N
Shp_VaRCF_ifl_v05	Sharpe ratio based on Cornish-Fisher VaR	Reserve return = inflation rate, threshold for VaR is 5%	Return-based Ratio	Favre L. and Galeano J.A. (2002), *JAI*	3_1_2_2_3	N
Shp_VaRCF_ifl_v2	Sharpe ratio based on Cornish-Fisher VaR	Reserve return = inflation rate, threshold for VaR is 20%	Return-based Ratio	id.	3_1_2_2_3	N
Shp_VaRCF_rf_v05	Sharpe ratio based on Cornish-Fisher VaR	Reserve return = risk-free rate, threshold for VaR is 5%	Return-based Ratio	id.	3_1_2_2_3	N
Shp_VaRCF_rf_v2	Sharpe ratio based on Cornish-Fisher VaR	Reserve return = risk-free rate, threshold for VaR is 20%	Return-based Ratio	id.	3_1_2_2_3	N

Acronym	Full name	Parameters	Class	Initial reference	Section	Select
Shp_VaRCF_zro_v05	Sharpe ratio based on Cornish-Fisher VaR	Reserve return = zero percent, threshold for VaR is 5%	Return-based Ratio	id.	3_1_2_2_3	N
Shp_VaRCF_zro_v2	Sharpe ratio based on Cornish-Fisher VaR	Reserve return = zero percent, threshold for VaR is 20%	Return-based Ratio	id.	3_1_2_2_3	N
Sortino_ifl	Sortino ratio	Reserve return = inflation rate	Return-based Ratio	Bawa V. (1975), *JFE*; Ang J. and Chua J. (1979), *JFQA*; Sortino F. and Van der Meer R. (1991), *JPM*	3_1_2_1_3	N
Sortino_rf	Sortino ratio	Reserve return = risk-free rate	Return-based Ratio	id.	3_1_2_1_3	N
Sortino_Sat_ifl_1	Sortino-Satchell ratio or Kappa coefficient	Reserve return = inflation rate, power index = 1	Return-based Ratio	Sortino F. (2000), *PI*; Sortino F. and Satchell S. (2001), *book*; Kaplan P. and Knowles J. (2004), *JPMr*	3_1_2_1_5	N
Sortino_Sat_ifl_2	Sortino-Satchell ratio or Kappa coefficient	Reserve return = inflation rate, power index = 2	Return-based Ratio	id.	3_1_2_1_5	Y
Sortino_Sat_ifl_3	Sortino-Satchell ratio or Kappa coefficient	Reserve return = inflation rate, power index = 3	Return-based Ratio	id.	3_1_2_1_5	N
Sortino_Sat_ifl_5	Sortino-Satchell ratio or Kappa coefficient	Reserve return = inflation rate, power index = 5	Return-based Ratio	id.	3_1_2_1_5	N
Sortino_Sat_rf_1	Sortino-Satchell ratio or Kappa coefficient	Reserve return = risk-free rate, power index = 1	Return-based Ratio	id.	3_1_2_1_5	N
Sortino_Sat_rf_2	Sortino-Satchell ratio or Kappa coefficient	Reserve return = risk-free rate, power index = 2	Return-based Ratio	id.	3_1_2_1_5	N
Sortino_Sat_rf_3	Sortino-Satchell ratio or Kappa coefficient	Reserve return = risk-free rate, power index = 3	Return-based Ratio	id.	3_1_2_1_5	N
Sortino_Sat_rf_5	Sortino-Satchell ratio or Kappa coefficient	Reserve return = risk-free rate, power index = 5	Return-based Ratio	id.	3_1_2_1_5	N
Sortino_Sat_zro_1	Sortino-Satchell ratio or Kappa coefficient	Reserve return = zero percent, power index = 1	Return-based Ratio	id.	3_1_2_1_5	N
Sortino_Sat_zro_2	Sortino-Satchell ratio or Kappa coefficient	Reserve return = zero percent, power index = 2	Return-based Ratio	id.	3_1_2_1_5	N
Sortino_Sat_zro_3	Sortino-Satchell ratio or Kappa coefficient	Reserve return = zero percent, power index = 3	Return-based Ratio	id.	3_1_2_1_5	N
Sortino_Sat_zro_5	Sortino-Satchell ratio or Kappa coefficient	Reserve return = zero percent, power index = 5	Return-based Ratio	id.	3_1_2_1_5	N
Sortino_SK_ifl	Sortino + Skewness / Kurtosis ratio	Reserve return = inflation rate	Return-based Ratio	Watanabe Y. (2006), *JPMr*	3_1_2_1_4	Y
Sortino_SK_rf	Sortino + Skewness / Kurtosis ratio	Reserve return = risk-free rate	Return-based Ratio	id.	3_1_2_1_4	N

Acronym	Full name	Parameters	Class	Initial reference	Section	Select
Sortino_SK_zro	Sortino + Skewness / Kurtosis ratio	Reserve return = zero percent	Return-based Ratio	id.	3_1_2_1_4	N
Sortino_zro	Sortino ratio	Reserve return = zero percent	Return-based Ratio	Bawa V. (1975), *JFE*; Ang J. and Chua J. (1979), *JFQA*; Sortino F. and Van der Meer R. (1991), *JPM*	3_1_2_1_3	N
SRAP	Style risk-adjusted performance measure (SRAP)		Return-based Difference	Lobosco A. (1999), *JPM*	4_1_1_4	Y
Stdzd_Infor_Ratio_1	Standardized Information ratio n°1		Return-based Ratio	Bodson L., Cavenaile L., Hübner G. (2009), *WP*		N
Stdzd_Infor_Ratio_2	Standardized Information ratio n°2		Return-based Ratio	id.		N
Stdzd_Infor_Ratio_3	Standardized Information ratio n°3		Return-based Ratio	id.		N
Stdzd_Infor_Ratio_4	Standardized Information ratio n°4		Return-based Ratio	id.		Y
Sterling	Sterling ratio		Return-based Ratio	Kestner L. (1996), *Fu*	3_1_3_3_2	N
Sterling_Calmar_3	Sterling-Calmar ratio		Return-based Ratio		3_1_3_3_3_3	N
Stutzer_ifl	Stutzer index of convergence	Threshold = inflation rate	Preference based	Stutzer M. (2000), *FAJ*	5_1_1_2	Y
Stutzer_rf	Stutzer index of convergence	Threshold = risk-free rate	Preference based	id.	5_1_1_2	Y
Total_risk_alpha	Total risk alpha		Return-based Difference	Fama E. (1972), *JoF*	4_2_1_2_4	Y
Treynor	Treynor ratio		Return-based Ratio	Treynor J. (1965), *HBR*	3_2_1_1	N
Trn_Maz_alpha	Alpha in Treynor and Mazuy's market timing model		Market timing	Treynor J. and Mazuy K. (1966), *HBR*	6_1_1_1	Y
Trn_Maz_cub_alpha	Alpha in Treynor and Mazuy extended timing model		Market timing	Jagannathan R. and Korajczyk R. (1986), *JB*	6_2_1_1	N
Trn_Maz_cub_delta	Delta in Treynor and Mazuy extended timing model		Market timing	id.	6_2_1_1	Y
Trn_Maz_cub_gamma	Gamma in Treynor and Mazuy extended timing model		Market timing - gamma	id.	6_2_1_1	Y
Trn_Maz_gamma	Treynor and Mazuy's coefficient		Market timing - gamma	Treynor J. and Mazuy K. (1966), *HBR*	6_1_1_1	Y

Acronym	Full name	Parameters	Class	Initial reference	Section	Select
Upsd_pot_ratio_ifl	Upside potential ratio	Reserve return = inflation rate	Return-based Ratio	Sortino F., Van der Meer R. and Plantinga A. (1999), *JPM*	3_1_3_1_3	Y
Upsd_pot_ratio_rf	Upside potential ratio	Reserve return = risk-free rate	Return-based Ratio	id.	3_1_3_1_3	N

This Table reports the 147 measures used in the paper. For each measure, the columns c_i (from left to right) report the following information: (c_1) acronym of the measure as it will be used in other Tables and Figures; (c_2) full name; (c_3) parameters used in the computation, if applicable; (c_4) class of the measure in Cogneau and Hübner (2009 and 2009a); (c_5) name(s) of the author(s), year of publication, acronym of the Journal : AER = American Economic Review, Ec = Econometrica, EIR = European Investment Review, EJOR = European Journal of Operational Research, FAJ = Financial Analysts Journal, Fu = Futures, HBR = Harvard Business Review, IJFE = international journal of Finance and Economics, IREF = International Review of Economics and Finance, JAFA = Journal of Applied Functional Analysis, JAI = Journal of Alternative Investments, JAM = Journal of Asset Management, JB = Journal of Business, JBF = Journal of Banking and Finance, JFE = Journal of Financial Economics, JFQA = Journal of Financial and Quantitative Analysis, JoF = Journal of Finance, JoI = Journal of Investing, JPM = Journal of Performance Management, JPMr = Journal of Performance Measurement, MS = Management Science, PI = Pensions and Investments, RB = Revue Banque, RFS = Review of Financial Studies, RoF = Review of Finance, Wi = Wilmott, WP = Working Paper; (c_6) section where the measure is classified in Cogneau and Hübner (2009 and 2009a); (c_7) flag indicating if the measure is in the 36 measures whose two-by-two correlation is lower than 90%, and thus is included in the principal component analysis.

[i] To quote a few, distribution-based measures that involve a percentile, like Value-at-Risk, feature the choice of the confidence level; preference-based measures such as the Morningstar MRAR metric usually require the assessment of the investor's risk aversion; or loss-based measures like the Roy "safety-first" criterion have to specify a latent reservation rate that replaces the risk-free interest rate.

[ii] We are aware that performance measures are considered to be adequate for certain investor profiles. For instance, the Sharpe ratio is an adequate performance measure if the investor wishes to invest all of his wealth in one risky asset; in contrast, Jensen's Alpha is adequate if the risky asset under evaluation is simply a fraction of the investor's portfolio. Our goal is precisely to explore the whole sample of measures and to determine the dimensions they cover.

[iii] Here, we do not consider papers that focus on risk measures only, like Pfingsten *et al.* (2004).

[iv] Many commonly used distributions satisfy this condition, a.o. Beta, Extreme value, Gamma, Logistic, Normal, Student t, Uniform, Weibull, Normal inverse Gaussian.

[v] Following Schuhmacher and Eling (2012), an admissible measure is a strictly increasing function of the Sharpe ratio under similar distributional conditions.

[vi] The choice of weekly data in this context is defended by Bollen and Busse (2001), Coles *et al.* (2006) and Cogneau and Hübner (2015).

[vii] When a fund is segmented in different share classes, each segmented occurrence is considered as a different fund.

[viii] Fund type is checked through a manual search with Bloomberg. Restricting the sample to only open-ended accumulation funds enables us to avoid the issue of dividends, which may have a country-specific tax impact for investors. This condition leads us to discard a large number of funds denominated in USD, which explains their lower number, but the sample size remains sufficient to draw statistical inferences.

[ix] Because of the international character of the study, we prefer to rely on a single database. Nevertheless, we manually ran a number of probes to ensure the consistency of data retrieved from Thomson Reuters Datastream with the corresponding CRSP returns.

[x] We apply the following filters: (i) we exclude from the sample all funds having missing data in their price series, at least three times three consecutive identical prices, or at least eight times two consecutive identical prices, (ii) if the shares of a fund have once been divided or regrouped, we recalculate the whole series of prices starting from the day of the event; (iii) we perform a global check of the plausibility of the prices and manually fix anomalous prices in the series; (iv) we exclude all "cousin" funds (140 cases) displaying returns with more than 80% correlation with an included fund; and (v) we eliminate funds whose return-based style analysis reveals that they are invested in bonds or in money markets.

[xi] Previous research has highlighted that various values of the parameters permits to represent different investor profiles (see for instance Farinelli *et al.* (2009) for the various investor profiles mimicked by the Sortino-Satchell, Generalized Rachev and Farinelli-Tibiletti ratios). This inflation of potential measures will be counterbalanced by the further elimination process which is described in subsection 4.2.

[xii] The complete parameterization of these variations and the source documents are available upon request.

[xiii] Detailed results are available upon request.

[xiv] We have also processed the principal component analysis on the whole set of 147 measures. Due to the important level of correlation between measures, the process is less precise. In output, we obtain 9 main factors who include the 7 who are distinguished in the main analysis below. Detailed results are available upon request.

[xv] With a sample of 1,625 funds, a selected threshold of 90% has a confidence interval of [88.7124% 91.1476%] at a significance threshold of 99% (see Fisher, 1915).

[xvi] We have executed other orthogonal rotations: Quartimax, Biquartimax, Equamax, Parsimax, Orthomax… The rotated factors produced are always extremely similar: same decomposition in the basis measures, weights almost identical… Results are not reported here but are available upon request.

[xvii] Using correlation analysis, it is also possible to relate the highlighted factors to particular measures. For example, the Sharpe ratio is highly correlated to eSDAR and therefore to factor 4; the Fama & French alpha is correlated to the Market Timing alpha and to the Total risk alpha, thus to factor 2; MorningStar measures are correlated to the Modified Treynor, thus to factor 14; etc.

[xviii] Different methods can be used to select the number of factors: all eigenvectors whose eigenvalues is higher than 1, all eigenvectors having a percentage of variance higher than (100% / number of variables), all eigenvectors being to the left of an elbow on the graphic, all eigenvectors being to the left of the point when the convexity of

the curve changes, all eigenvectors being to the left of the point before a plateau: the latter is the method we selected. We also process the analysis presented below with more factors, but the main conclusions remain, in particular the highlighting of factors 2, 4 and 7.

[xix] We do not have an explanation for those (very limited) transfer of performance measures. We uncover that there might be a difference in the stability of performance over time that lead to slight differences in horizon-based dimensions: this is something that is left for future research.

[xx] The process is based on the FASTCLUS procedure from SAS, recommended for sets of more than 100 observations. It relies on a standard iterative algorithm for minimizing the sum of squared distances from the cluster means.

[xxi] We also process the same analysis considering the true factor as described in Table IV: results are similar and even more conclusive. Then, we also process the same analysis replacing each factor by the second measure in the factor. Once again, the reported persistence is very similar to what we have highlighted in Table VII and VIII – but slightly less conclusive. Both of those results are not reported here but are available upon request.

[xxii]The subsequent analysis is based on the same data set as before, but with a distinct objective, namely to detect predictability or persistence. Both analyses are in-sample, but operated distinctively. To assess the robustness of our results, each of these two analyses have been conducted on two distinct subsamples from the whole set of data, with extremely similar results leading to the same conclusions (results available upon request).

[xxiii] This penalty is supposed to represent the expenses needed for the fund closing: auditing and legal fees, communication expenses, liquidation expenses… Cogneau and Hübner (2015) analyze the influence of the penalty level when building portfolios.

[xxiv] We perform the same computations with penalties of 5% or 10% applied to each fund's disappearance, and we get extremely similar patterns for Tables VI and VII. We also verify the robustness of our findings by considering performance computed in EUR, then in GBP, and finally in USD. The results are all similar.

[xxv] These results are not presented in this paper but are available on request.

[xxvi] See for instance Jegadeesh and Titman (1993), Hendricks, Patel and Zeckhauser (1993), Christopherson, Ferson and Glassman (1998).

[xxvii] Detailed results are not provided in Table VIII but are available upon request.

[xxviii] Results are not reported here but are available upon request.

[xxix] Regarding the age subsample decomposition, this process was impossible to run for the youngest funds, as all of them were launched after the start date of the rebalancing process.