The Fund Synthetic Index: An Alternative Benchmark for Mutual Funds

VIRGINIE TERRAZA*
Associate professor in finance
CREA, University of Luxembourg
Associate Researcher LAMETA, University of Montpellier I

HERY RAZAFI-TOMBO**
Associate professor in finance
CEREFIGE, IAE University Paul Verlaine – Metz

Evaluation of the performance of investment funds is a topic of considerable interest to practitioners and academic researchers. Performance indicators of financial places have for long posed interesting challenges with regards to funds investors, but also to legal regulators authorities. The two major issues that need to be addressed in any performance study are how to choose an appropriate benchmark for comparison and how to adjust a fund's return for risk. Indeed, investors desire information about representative market indexes as a norm to evaluate the performance of their portfolios. MSCI Indexes are frequently used by institutional investors around the world as benchmarks to decide allocation of funds across asset classes and regions. Despite this wide acceptance, MSCI country index has in its original application a number of drawbacks and limitations. The main problems can be traced to the presence of usual biases, such as sampling, survivorship and instant history biases (Fung and Shieh, 2002), involving problems into the aggregation procedure. Thus, one can explain why certain financial places are less representative, specifically for funds distribution places. Some results also indicate problems related to misclassification in mutual funds (Sharpe, 1992). Each of these phenomena can have a significant impact on international diversification for fund managers. Based on these empirical findings, Ferreira, Miguel and Ramos (2007) examine cross-country mutual fund performance using several alternative benchmark models including a domestic and an international version of the Carhart (1997) four-factor model. Using multiple regressions, they obtain significant determinants explaining funds performance, like the funds size, the fees, the management style... .

In this article, we contribute to the existing discussion on alternative benchmarks to compare financial places, by conducting an analysis on funds time variation structure. Contrary to previous literature, we propose to use directly the information contained in the NAV to extract performance characteristics of funds. Then, each domiciliation place is compared by constructing a fund synthetic index that will capture the time structure of mutual fund performance.

Usual statistical approach consists to estimate financial returns of each fund in each country, which involves dealing with huge data sets that may cause the calculation processes to become slow and cumbersome and the results difficult to be interpreted and used in further applications. To reduce data sets, and give conclusions for each financial place, indicators of the mean of fund's returns can be used. Then one can identify classes of domicile funds that are subject to common properties. But this classical approach, gives only approximate results because it is based on an aggregation of average performance and risk and a boxplot statistical format. The construction of fund synthetic portfolio avoids this issue. First, it avoids the logic of representativity through market capitalization that is often difficult to apply to the mutual fund universe. Second, it is based on factor analysis techniques to generate indexes that are able to capture a very large fraction of the information. More precisely, it permits to take into account the common properties of fund returns relative to their domicile while keeping the maximum of information given by the original data. Indeed, it may be very useful to use a transformation to form a simplified data set retaining the characteristics of the original data set. Principal component analysis, abbreviated as PCA, is a method of statistical analysis useful in data reduction and interpretation of multivariate data sets by identifying factors of common behavior such that not much of the contained information is lost. In our context, we use this method to derive portfolio weights in order to construct a synthetic portfolio of each financial place. For that, we propose to replace the matrix of returns and to derive an index which keeps the global representation of each financial place.

I. Methodology and data

The identification of representative index represents a major challenge in analyzing the performance (Agarwal,

* virginie.terraza@uni.lu
** razafitombo@univ-metz.fr

2001). The challenge is even more assertive when it comes from investment funds performance analysis. The main challenge is to take into account all the common properties of funds.

Generally, the construction of indexes poses two problems: the representativeness and the purity. Martellini, Vaissié and Goltz (2004) indicate that these two problems explain number of index available and the heterogeneity of methodology used. Besides the two methods: indexes segmented by style[1] and indexes explained by factors[2], we use the Zurich Capital Market "pure style" approach to construct the synthetic fund index to represent groups of funds by domicile. On one hand, index providers do not offer a benchmark based on funds domicile. On the other hand, the indexes based on factor exposures require complex modeling strategy that is not relevant to funds domiciliation purpose. However, the pure style index approach permits to meet the needs of practitioners in terms of transparency and reliability. It uses all available information in the series of returns based on non-observable factor, fully representative and unbiased.

I.1. FUND SYNTHETIC INDEX CONSTRUCTION

The main goal of our analysis consists in investigating the construction of a fund index by country in order to better compare them. We propose a method to design fund benchmarks satisfying all defining properties for a good index. More specifically, our methodology is based on the concept of factor portfolios analysis. Black and Litterman (1992), or Chan, Karceski and Lakonishok (1998), have used principal component and factor analysis to examine the existence of common movements in asset returns. These analyses represent alternatives to fundamental approaches which relate the factors influencing financial asset returns to macroeconomic measures. In this context, historical returns are used to estimate orthogonal statistical factors and their relationship with the original variables. Huberman, Kandel and Stambaugh (1987) formalize the construction of replicating portfolios for the statistical factors. More recently, Martellini, Vaissié and Goltz (2004) apply this methodology to the benchmarking of hedge fund style returns. In the same order of idea, we extend the previous approach in order to define an alternative benchmark for the mutual funds. We assume that the index is a linear combination of adjusted prices or returns:

$$S = \sum_{i=1}^{n} w_i R_i = w^T R \qquad (1)$$

The main difficulty is the choice of an index weighting scheme. In this context, the index is provided in order to replicate the return pattern of funds of each country. We assume that capturing the structure of fund returns may improve the results.

The index will be more informative if weights are assigned in a way which captures the maximum variance of the set of reference fund returns over time. To obtain in index with its property, is the extraction from the covariance matrix of stock prices of the largest characteristic root.

At the same time, large data set is a constraint to obtain a suitable index. We use a Principal Component Analysis approach, a mathematical technique used in risk management applications to reduce a complex data set to a lower dimension. Indeed, PCA obtains a new set of uncorrelated variables that are a linear combination of the returns. The dimensionality of the returns matrix can be reduced by selecting only that subset of PCs that contribute the most to the variance of the returns, Consider the set of n fund returns $R_1, \ldots, R_n$ as a random vector R with zero empirical mean and non-singular covariance matrix $\sum$:

$$E(R) = 0 \; ; \; Cov(R) = \sum \qquad (2)$$

The objective is to find a linear combination of random variables $R_1, \ldots, R_n$ that contains as much of the variability of the random variables as possible[3], Let F a re-representation of that data set. We can rewrite the equation 1 by replacing the original data set by the matrix of principal component:

$$S = \sum_{i=1}^{n} w_i F_i \qquad (3)$$

With

$$F_i = b_i H = b_{1i} H_1 + \ldots + b_{ni} H_n \qquad (4)$$

are the i[th] principal component i = 1,...,n

Vector b is a weight vector that tells us by what weight does each of the variables H_j affect the variance of the linear combination.

Specifically the goal of the methodology is to derive weights which contribute most to the significant improvement on benchmark performance. According the PCA methodology, our alternative benchmark is a linear combination of a large set of funds returns that weighting scheme takes into account the underlying variance-covariance structure of funds. Hence equation 1 can be rewritten to provide the fund synthetic index S, mapped into its exposures on the first K principal components as:

$$S = \sum_{j=1}^{k} \sum_{i=1}^{n} R_i b_{i,j}^2 \qquad (5)$$

or in the matrix form:

$$\underset{(N,1)}{S} = \underset{(N,n)}{R} \times \underset{(n,n)}{B}^2 \times \underset{(n,1)}{U} \qquad (6)$$

where $b_{i,j}^2$ represent absolute contributions of the original variables in principal components F and $\sum_{i=1}^{n} b_i^2 = 1$.

The fund synthetic index presents several advantages. Firstly, it permits to take into account the common properties of funds return relative to their domicile while keeping the maximum of information given by the original data. Indeed, it may be very useful to simplify the data or the data structure by identifying factors of common behaviour such that not much of the contained information is lost. Secondly, using the principal Component

Approach, it avoids biases in linear weighting scheme of portfolios, reducing the dimensionality of the data and keeping the representatively of financial markets. Synthetic fund indexes permit to better compare fund markets when structural information of returns is used, means-based measures face a bias if managers can trade between observation dates. The new measures avoid this interim trading bias.

Next, the variance of the synthetic fund index can be computed easily:

$$\begin{aligned} S &= \sum_{j=1}^{k}\left(\sum_{i=1}^{n} R_i b_{i,j}^2\right) \\ &= (R_1 b_{1,1}^2 + \dots + R_n b_{n,1}^2) + \dots \\ &\quad + (R_1 b_{1,k}^2 + \dots + R_n b_{n,k}^2) \\ &= \alpha_1 R_1 + \dots + \alpha_n R_n \end{aligned} \tag{7}$$

$$\sigma^2(S) = \alpha^T \sum \alpha$$

where $\alpha_i = b_i^2$

However, in general, the estimation of the variance becomes difficult, when introducing correlations effects. In fact, it is well known that the number of correlations increases geometrically with the number of assets. Then, it is more likely that some correlations will be measured inaccurately or incorrectly. Furthermore the computation time of covariance matrix can increase dramatically, which is not feasible for making quick decisions on trading portfolio positions in fast-changing markets.

Then in order to simplify the estimation of the variance – covariance matrix $\sum$, we can extend our approach. Following equation 2, and taking the first K principal components, the decomposition of S is given by:

$$\begin{aligned} S &= \sum_{i=1}^{n} w_i R_i \approx w_1(b_{11}F_1 + \dots + b_{1k}F_K) + \dots \\ &\quad + w_n(b_{n1}F_1 + \dots + b_{nk}F_K) \\ &= (w_1 b_{11} + \dots + w_n b_{n1})F_1 + \dots \\ &\quad + (w_1 b_{1k} + \dots + w_n b_{nk})F_k \\ &= \beta_1 F_1 + \dots + \beta_k F_k \end{aligned} \tag{8}$$

Where $\beta_i = w^T b_i$ represent the weighted exposure to the ith principal component.

Let Ω the diagonal $K \times K$ variance-covariance matrix of the $K \times 1$ vector of principal component:

$$\Omega = \begin{pmatrix} \lambda_1 & & 0 \\ & \dots & \\ 0 & & \lambda_K \end{pmatrix}$$

The variance of fund synthetic indexes becomes:

$$\sigma^2(S) = \beta^T \Omega \beta = \beta_1^2 \sigma^2(F_1) + \dots + \beta_k^2 \sigma^2(F_k) \tag{9}$$

The variance of the fund synthetic index can be defined through a few linear combinations of these variables. Since PCA can also be used in the field of portfolio risk management to reduce the dimensionality of the risk factor space, PCA may be used in significant simplification of Value at Risk (VaR) models.

I.2. DATA

In this study, we use daily frequency observations for 2147 equity global funds from the Lipper funds database. The time period covered for the study is from January 2000 to December 2007. Different ranking periods are used in order to isolate a crisis period (2000-2003), a relative period of stability (2004-2007). The global period (2000-2007) will capture the long-run past performance[4]. Our sample includes both live and defunct funds to mitigate the impact of survivorship bias. The number of funds include in our sample are listed in table 1[5].

For these countries, the MSCI World index is chosen as a benchmark and the Euribor (6- month) is used to estimate the risk premium and one performance indicator widely used the Sharpe ratio.

In this study, each fund synthetic index represents the evolution of the mutual fund industry in its respective country. Portfolios have been composed by filtering representative funds using Principal Component Analysis.

By definition, The PCA aims to explain the behavior of observed variables – funds returns -using a smaller set of unobserved implied variables. The number of principal components to be retained for further analysis is determined by the correlation structure of the data. If the data are all highly mutually correlated, one or two principal components will suffice to explain a large fraction of total data variation. In our experiment, only a few principal components are required for achieving the highest accuracy for each country (cf. infra correlation issues). More precisely, two principal components are sufficient to explain most of the variability present in the data.

Then, the principal component can be expressed as:

$$F_j = b_j H = b_{1i} H_1 + b_{2i} H_2 \tag{10}$$

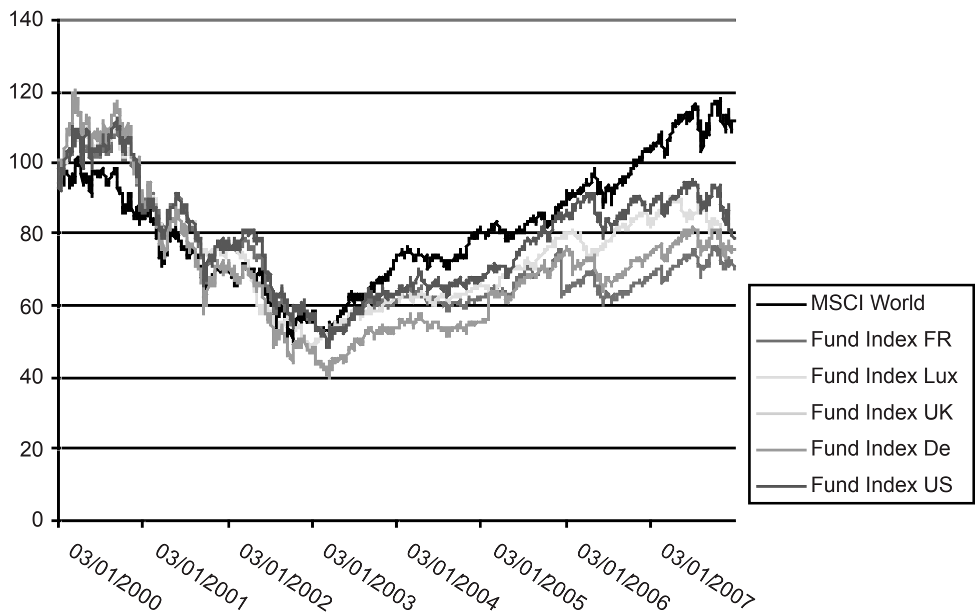
and the fund synthetic index by:

$$S = \sum_{j=1}^{2}\sum_{i=1}^{n} R_i b_{i,j}^2 \tag{11}$$

Table 1: Database

	DE	FR	UK	LU	US	Total
Initial database	187	254	287	552	867	2,147
2000-2007	74	62	49	148	218	551
2000-2003	74	66	78	147	218	583
2004-2007	152	131	72	232	251	838

Figure 1: The evolution of benchmarks



From this figure, we provide a brief overview of the evolution of our benchmarks during the whole period. We can observe an interesting phenomenon: the relative underperformance of the MSCI index compared to funds synthetic indexes during the crisis period. One possible explanation for this could be the information provided by the MSCI index, estimated using a weighted arithmetic average together with the concept of chain-linking. Then, during extreme events periods, information on volatility contained in data structure is ignored by the MSCI. In contrast, during the relative period of stability, the MSCI index outperforms fund synthetic indexes. Furthermore, it seems that a high correlation between indexes is found. However in order to obtain more coherent conclusions, in the next section we look at the statistical quality of our results in order to assess and quantify the performance of the benchmarks.

II. Results

This section presents the survey results. Especially, we analyse the key indicators of performance and risk measures used in fund industry. In a first subsection, a statistical comparison between benchmarks provides insights about performance in a sample of equity investment funds of 5 countries. The second subsection turns to the key indicators of Value at Risk estimated using the parametric approach. The third subsection examines correlation issues, in order to analyse the common influences between indexes.

II.1. A STATISTICAL COMPARISON

Table 2 reports the descriptive statistics of our indexes sample. We observe that the average mean of returns over the entire sample period is slightly negative (-6.52% per year) for all indexes. During the first sample sub-period all average index mean returns are negative (-22.58%), whereas they all become positive (+17.95%) during the second sample sub-period. We may interpret this as an improvement of market conditions from periods 2000–2003 to 2004-2007.

As statistics for individual group of domicile, the gaps between average performance for all synthetic fund indexes and the MSCI are high. This finding was behind the construction of fund synthetic indexes. It is important to note that these differences are merely the reflection of the original data. It reflects the differences in performance and risk between each group of investment funds (cf. Appendix 2). Indeed, using a classical statistical comparison - boxplot format-, we observe that there is a little evidence of homogeneity of funds within a group and their difference from funds in other domicile groups. However, this statistical approach does not allow for a more accurate analysis of the differences between the groups of funds, especially because it is based on an aggregation of average performance or risk for each fund within a group (Razafitombo and Terraza, 2008).

All groups of domicile present a lower performance than MSCI index (2.51%). No group exceeds the performance MSCI over 2000-2007 and 2000-2003. Over these two periods, we observe that UK fund synthetic indexes display the least bad performance, -4.98% and -21.09% respectively, whereas DE fund synthetic index has the worst performance, -8.89% and -24.81% during the same periods. However, fund synthetic indexes show a better performance (17,95%) than MSCI benchmark (12.23%) over the bullish period (2004-2007), except US funds synthetic index. We can state that Fund synthetic indexes are on average more volatile than MSCI. More precisely, during the bullish period, LU Synthetic fund index is the least risky while US fund synthetic index presents the higher performance. As Regards the coefficient of variation and Sharpe ratios, no conclusion can be given for 2000-2007 and 2000-2003 periods because of their negative value. Over the 2004-2007period, the performance indicator confirm the previous results as we find that the US synthetic fund display a poor prime to variability with a Sharpe ratio equal to 0,169, whereas UK funds present a higher Sharpe ratio.

Our sample of benchmark properties reveals that average sample skewness is consistency negative for all three sample periods which indicate significant non-normality. The first sub-period is less negative (-0.229) than the second sub-period (-0.579), implying a generally bearish market trend with contracting volatility, relative to the previous sub-period. We find also that most funds synthetic indexes display statistically significant clustering in the extremes of the left tail.

As we can see, FR funds significantly dominate the negative skew factor and the kurtosis measure in periods 2000-2003 and 2000-2007. This implies a high concentration of events around the mean, which is negative. In period 2004-2007, excess Kurtosis is at the same

Table 2: Data properties of indexes

Data properties of indexes							
Period	**2000-2007**						
	DE	**FR**	**UK**	**LU**	**US**	**MSCI**	**Average**
Nbr of observations	2,089						
Minimum	−0.138	−0.138	−0.068	−0.059	−0.111	−0.039	−0.103
Maximum	0.095	0.083	0.075	0.064	0.094	0.047	0.082
Mean (annualized)	−8.89%	−5.85%	−4.98%	−5.49%	−7.40%	2.51%	−6.52%
Std. dev. (annualized)	31.85%	20.51%	22.37%	23.04%	31.72%	14.07%	25.90%
Coefficient of variation	−3.582	−3.507	−4.496	−4.193	−4.284	5.613	−4.013
Skewness	−0.347	−0.923	−0.365	−0.298	−0.170	−0.009	−0.421
Excess Kurtosis	3.293	10.664	2.062	1.600	2.048	2.297	3.934
Sharpe ratio	−0.428	−0.517	−0.435	−0.444	−0.383	−0.159	−0.441
Period	**2000-2003**						
	DE	**FR**	**UK**	**LU**	**US**	**MSCI**	**Average**
Nbr of observations	1,042						
Minimum	−0.087	−0.122	−0.067	−0.061	−0.109	−0.039	−0.089
Maximum	0.070	0.070	0.059	0.064	0.092	0.047	0.071
Mean (annualized)	−24.81%	−23.50%	−21.09%	−23.08%	−20.42%	−6.28%	−22.58%
Std. dev. (annualized)	30.46%	24.75%	26.73%	28.17%	35.79%	17.03%	29.18%
Coefficient of variation	−1.23	−1.05	−1.27	−1.22	−1.75	−2.71	−1.304
Skewness	−0.230	−0.606	−0.148	−0.126	−0.037	0.123	−0.229
Excess Kurtosis	0.876	5.165	0.661	0.577	1.536	1.444	1.763
Sharpe ratio	−0.890	−1.043	−0.875	−0.901	−0.635	−0.504	−0.869
Period	**2004– 2007**						
	DE	**FR**	**UK**	**LU**	**US**	**MSCI**	**Average**
Nbr of observations	1,043						
Minimum	−0.039	−0.041	−0.044	−0.038	−0.078	−0.025	−0.048
Maximum	0.032	0.032	0.031	0.028	0.064	0.021	0.037
Mean (annualized)	19.89%	18.61%	23.01%	19.30%	8.92%	12.23%	17.95%
Std. dev. (annualized)	15.21%	14.96%	17.49%	14.54%	24.67%	10.30%	17.37%
Coefficient of variation	0.76	0.80	0.76	0.75	2.76	0.84	1.169
Skewness	−0.669	−0.638	−0.577	−0.629	−0.382	−0.325	−0.579
Excess Kurtosis	1.571	1.656	1.190	1.489	1.847	0.880	1.550
Sharpe ratio	0.996	0.927	1.044	1.001	0.169	0.727	0.827

level for all our indexes. Therefore, FR funds have a high concentration of events around its negative mean. From these results, we can state that FR synthetic funds index can be considered as less risky than both the US and DE. We know about the high likelihood of negative returns, with comparatively moderate deviation from the mean. Thus, rational investors can profit from this bear cycle by adjusting strategies such as short positions, stop-loss strategies and others. On average in the whole period, excess indexes' kurtosis is significant at +3.934, which implies fat tails for the left tail of their distributions. The agglomeration of extreme events at the left tails of our sample distributions further confirms this hypothesis.

II.2. VALUE AT RISK ANALYSIS

To complete our analysis on fund synthetic fund's risk, another statistical measure widely used is Value at Risk. The Value at Risk (VaR) has rapidly become the standard quantitative benchmark for measuring the risk exposures of financial portfolios and has become a standard concept in risk management (Pichler and Selitsch, 1999 and Jorion, 2001). For the simplest model, the VaR depends on the mean and standard deviation of the normal density and the critical value α corresponds to a confidence level. However, it is well known that returns do scarcely follow a distribution that is approximately normal. Therefore the use of the classical VaR will result in a systematic undervaluation of the VaR measure. An adapted version has been developed by Cornish and Fisher (1937), who expanded the original equation to adjust for skewness and kurtosis of the actual returns distribution:

$$VaR_{CF}(\alpha) = \mu + \Omega(\alpha)\cdot\sigma$$

$$VaR_{CF}(\alpha) = \mu + \sigma\Big[z(\alpha) + \frac{1}{6}\left(z(\alpha)^2 - 1\right)S + \frac{1}{24}\left(z(\alpha)^3 - 3z(\alpha)\right)K - \frac{1}{36}\left(2z(\alpha)^3 - 5z(\alpha)\right)S^2\Big] \quad (12)$$

where μ is the mean, σ is the standard deviation of the sample of returns and $\Omega(\alpha)$ is the Cornish-Fisher value based on loss probability, skewness and kurtosis of the returns distribution. Hence this formulation allows calculating VaR for distributions that exhibit skewness and excess kurtosis. The results of Cornish Fisher VaR can be seen in table 3 using 99% as confidence level.

As we can see from table 3, globally, the MSCI seems to be the least risky of all indexes followed by FR and UK is the riskier during the two sub-periods. In all our results, risky indexes' tail events have significantly exceeded those of the MSCI. Hence using the MSCI carelessly as a risk measure for much riskier market environments may lead the investor to severely underestimate downside risks and thus VaR. At present, this traditional measure is still the most widely used risk measure in financial institutions. Unfortunately, this measure doesn't capture the dynamics of large losses. Many authors have shown that the efficiency of the VaR risk measures may be improved significantly by

Table 3. Cornish-Fisher Value at Risk

	DE	FR	UK	LU	US	MSCI
2000-2007	−0.036	−0.008	−0.026	−0.028	−0.041	−0.019
2000-2003	−0.040	−0.023	−0.050	−0.039	−0.037	−0.025
2004-2007	−0.014	−0.014	−0.028	−0.014	−0.018	−0.012
α = 99%						

implementing custom dynamic return distributions, such as the GARCH distributions. Then, conditional variances can be modelled much more precisely since effects such as volatility clustering leverage effects and leptokurtosis are captured. We therefore extend our Cornish Fisher VaR model by estimating our volatility using a GARCH process, where the variance is estimated as:

$$\sigma^2_{t,GARCH} = \alpha_0 + \sum_{i=1}^{p}\alpha_i r^2_{t-i} + \sum_{j=1}^{q}\beta_j\sigma^2_{t-j} \quad (13)$$

Where r^2 and σ^2 respectively are the squared returns (ARCH term) and the conditional variance (GARCH term) of the returns series and α_0 the long term volatility. For each fund index, we fit a GARCH (1,1) model using Maximum Likelihood procedures. For constructing our GARCH-VaR vector, we first calculate the Cornish-Fisher factor and then proceed to determine daily volatilities adjusted for GARCH effects. Table 4 reports the results of Cornish Fisher GARCH VaR with α =99%. We can read that the specification of the conditional volatility processes makes a significant difference in our results. Comparisons among the two VaR models show that the simple Cornish-Fisher VaR gives undervaluation of the risk exposure of the underlying asset during extreme events. Here, the gaps between the MSCI and fund synthetic indexes are more pronounced than those obtained before for some indexes. Results still indicate that the MSCI displays the less VaR measures based on GARCH volatilities while US index is the riskier index especially during the crisis period.

In all cases, we notice the undervaluation of the MSCI compared to the other fund synthetic indexes. More precisely, our analysis indicates that during crisis period, the MSCI have shown performance discrepancies regardless of fund synthetic indexes. Hence, this raises the question whether the MSCI is actually a suitable benchmark for measuring risk in the US and other important financial markets? Again the contrast between the US and MSCI indexes represent a good example of the danger.

II.3. CORRELATION ISSUES

Here, we consider another reduced form dimension in order to measuring possible gains by combining countries. First, we investigate correlations between fund synthetic indexes estimated for different countries. Second, we analyze combined data for the five countries

Table 4. Cornish-Fisher GARCH Value at Risk

	DE	FR	UK	LU	US	MSCI
2000-2007	−0.111	−0.083	−0.057	−0.062	−0.115	−0.058
2000-2003	−0.060	−0.064	−0.055	−0.057	−0.108	−0.055
2004-2007	−0.010	−0.010	−0.120	−0.009	−0.024	−0.013
α = 99%						

Table 5. Correlation matrix

		DE	FR	UK	LU	US	MSCI
DE	2000-2007	1					
FR		0.711	1				
UK		0.600	0.634	1			
LU		0.855	0.778	0.792	1		
US		0.476	0.588	0.781	0.785	1	
MSCI		0.584	0.417	0.331	0.475	0.108	1

		DE	FR	UK	LU	US	MSCI
		2004-2007					
DE	2000-2003	1	*0.848*	*0.339*	*0.881*	*0.312*	*0.323*
FR		0.745	1	*0.563*	*0.965*	*0.668*	*0.621*
UK		0.597	0.693	1	*0.540*	*0.588*	*0.644*
LU		0.821	0.829	0.843	1	*0.651*	*0.531*
US		0.409	0.646	0.782	0.815	1	*0.660*
MSCI		0.740	0.532	0.476	0.608	0.173	1

described above. This last approach suggests that some dimension reduction is possible from combining data across countries - this reduction reflects correlations in returns across countries.

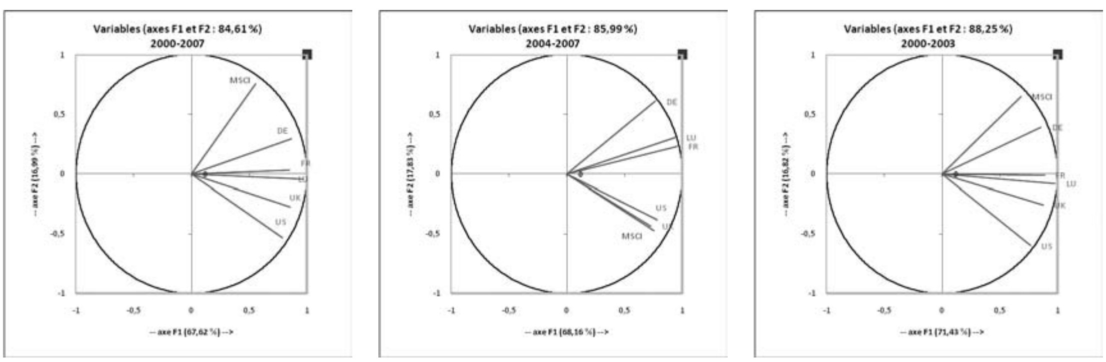
Table 5 compares correlations across countries over three time periods (2000-2003, 2004-2007 and 2000-2007). We find low correlation effects between the MSCI and each fund synthetic index, especially for US and UK during the whole period and the crisis period. The gap between indexes is less pronounced for 2004-2007 except for DE.

Next, we investigate correlations between principal components for our different countries and perform PCA for aggregate data.

Combining countries data gives a different perspective on correlations across country fund returns (see figure 2). These calculations suggest that there are minor additional reductions in dimension from combining data across countries. Evidence from correlations between country-level principal components suggests that there is some correlation between the first components extracted from different country fund return data. There is less evidence for correlation between second and third components for all indexes see results in the tables given in appendix 3. The major positive result is that relatively few principal components describe a large fraction of variance in the long ends of most country fund returns. Indeed, results show that the first two factorial axes explain more of 80% of the total variance whatever is the period.

In general, we see that the first principal component (labeled "PC1") is a roughly-equal linear combination for fund synthetic indexes except for the period 2004-2007 (Cf. Appendix 3). These indexes might reasonably be interpreted as a general fund return index. From tables, we observe that the second principal component (labeled "PC2") has different behavior depending the period. For 2000-2007, negative loadings for LU, US and UK, and positive loadings for DE, FR and MSCI. This loading appears to represent a specific index component. For 2000-2003, negative loadings for FR, LU, UK and US and positive loadings for DE and MSCI. For 2004-2007, negative loadings for MSCI, UK and US and positive loadings for FR, LU and DE. Furthermore, we can extract mainly two groups of variables with strong correlations, LU and FR for the first group and US, UK and the MSCI for the second group. The MSCI is an exception; movements in the index are

Figure 2: Correlation circles



described by components that are not highly correlated with other country fund returns. Furthermore, there is some evidence that correlations changed from the period 2000-2003 to the period 2004-2007 for the MSCI index. The index is stronger correlated with the LU, DE and FR group for the first period while it is stronger correlated with US and UK group in the second period.

These findings lead to various comments. Again, the fund synthetic indexes can highlight information that is simply not possible through traditional statistical comparison. The low correlation between all fund synthetic indexes and the MSCI confirms the existence of the problem of misclassification stated by Sharpe (1992). The cross-correlation and the positioning of each fund group on correlation circle suggest the existence of geographical proximity. There is a high and positive correlation between three European groups (DE, FR and LU). The UK and the US depart strongly from the other groups and from MSCI. This can be explained by the presence of home bias phenomenon. Following Hau and Rey (2008), the low correlation with the MSCI and with other fund synthetic index clearly suggests a high degree of home bias. For the UK and the LU groups, the comments would be more mitigated, insofar as those places present the more pure international funds. For the US, it seems that the fund sizes reinforce this evidence of home bias phenomenon. Finally and implicitly, the correlation analysis of synthetic fund indices allows highlighting the heterogeneity of investment strategies and institutional constraints that may characterize the different places.

III. Conclusion

Performance benchmarks are important for several reasons. They help to measure and to compare investment performance of institutional fund managers. They provide client with a reference point for monitoring that performance. They can also have the effect of modifying the behaviour of fund managers. Usually, there are two manners to evaluate portfolio performance. First, in addition to the rate of return, we can use composite equity portfolio performance measures that combine risk and return into single value (e.g. Sharpe, Treynor, Information ratios, etc). Second, we can use single index benchmark or peer group benchmark comparison. For these two approaches, we have to collect the returns produced by a representative universe of investors over a specific period of time and display them in a simple boxplot format. To aid the comparison, the universe is divided into percentiles to indicate their relative ranking.

Despite their robustness, both methods present several potential problems, in particular for evaluation performance of groups of funds. First, and foremost, the boxplot statistics do not make any explicit adjustment for the risk level of the portfolio in the universe. In the fact, investment risk is only implicitly considered to the extent that all the portfolios in the universe have essentially the same level of volatility. As our statistical analysis shows, this is not the case for any group of funds. Our results indicate that the universe mixes probably portfolios with different investment styles. Second, it is almost impossible to form a truly comparable peer group that is large enough to make the percentile rankings valid and meaningful. By focusing on nothing more than relative returns, such a comparison loses sight of whether investor in question – or any universe, for that matter – has accomplished his individual objectives and satisfied his investment constraints. Finally, our statistical results indicate the difficulty for all groups of funds to beat the MSCI index. It seems that investors are not represented in all markets covered by the index or there is the so called "home country bias" that is impossible to capture with these classical methods.

As an alternative benchmark, the fund synthetic index methodology constitutes a median solution to these problems. Following recent development of hedge fund indexes construction methodology, they keep maximum of information given by the original data and take into account the common properties of funds return relative to their domicile. This empirical work could be pursued in various ways. We can use the synthetic indexes to investigate the main determinants of the performance of financial markets. In a practical point of view, we can define and construct a new range of indexes that is more accurate with the local fund management industry characteristics.

1 The indexes segmented by style classify the funds by strategies and/or style by using the approach based on multifactorial styles (Fung & Hsieh, 1997; Schneeweis & Spurgin, 1998), or the approach the classification algorithms (Liang, 1999).

2 Several factor exposures are used to explain the returns of funds (Schneeweis, Kazemi, and Martin, 2001).

3 see J. Edward (1991) for a tutorial on Principal Component Analysis.

4 In order to reduce the impact of survivorship bias, we apply the same methodology (not reported here) to fund data based on annual period of calculations. The results are quite similar.

5 See Appendix 1 for database summary statistics

References

- Agarwal and Naik (2001), "Characterizing hedge fund risks with option based and buy and hold strategies", Working paper, London School of Economics, March.
- Agarwal and Naik (2004), "Risks and portfolio decisions involving hedge funds", Review of Financial Studies, Vol. 17, pp. 63-68
- Black F. and R. Litterman (1992), "Global portfolio optimization", Financial Analysts' Journal, 48 (5), pp. 28-43.
- Carhart M. (1997), "On persistence in mutual fund performance", The Journal of Finance, 52, March, pp. 57-82.
- Chan L., J. Karceski and J. Lakonishok (1998) "The risk and return from factors", Journal of Financial and Quantitative Analysis, 33(2), pp. 159-188.
- Edward Jackson J. (1991), "A user's guide to principal components", New York, Wiley-Interscience.
- Ferreira Miguel and Ramos (2007), "The determinants of Mutual Fund Performance: A cross country Study", Working Papers, ISCTE Business School.
- Fung and Hsieh (2000), "Performance characteristics of hedge funds and CTA's: Natural vs Spurious biases", Journal of Financial and Quantitative Analysis, vol. 35, pp. 291-307.
- Fung and Hsieh (2001), "The risks in hedge fund strategies: information content and measurement biases", Review of Financial Studies, 14 (2), pp. 313-341.
- Fung and Hsieh (2002), "Benchmark of hedge fund performance, information content and measurement biases", Financial Analyst Journal, Vol. 58 (1)1 (Jan.-Feb), pp. 22-34.
- Hau H. and Rey H. (2008), "Home Bias at the Fund Level", American Economic Review, Volume 98, Issue 2, May, pp. 333-338.
- Hubermann G., S. Kandel and R. Stambaugh (1987), "Mimicking portfolios and exact arbitrage pricing", Journal of Finance, 42 (1), pp. 1-9.
- Jorion P. (2001), "Value at Risk: The New Benchmark for Managing Financial Risk", 2nd edition, McGraw-Hill, New York.
- Kazemi H., Schneeweis T. and Martin G. (2001), "Understanding hedge fund performance: Research issue revisited – Part I", Journal of Alternative investment, vol. 5, n°3, winter 2002, pp. 6-22.
- Khorana A., Servaes H. and Tufano P. (2009), "Mutual funds fees around the world." Review of Financial Studies 22 (3) pp. 1279-1310.
- Martellini L., Vaissié M. and Goltz F. (2004), "The brave new world of hedge fund indexes", Working Paper, Edhec Risk and Asset Management Research Centre.
- Mitchell M. and Pulvino T. (2001), "Characteristics of risk return in risk arbitrage" Journal of Finance, vol. 56, n°6, December, pp. 2135-2175.
- Pichler S., Selitsch K. (1999), "A Comparison of Analytical VaR: Methodologies for Portfolios That Include Options", Working Paper, Technische Universität Wien, Austria.
- Razafitombo H. and Terraza V. (2008), "Domiciliation and Performance Analysis of investment funds", Working papers, CREA University of Luxembourg.
- Razafitombo H. and Terraza V. (2011), "A Structural Analysis of Mutual Fund Performance: A Comparative Study for Domiciliation", The Journal of Index and Investing, Spring 2011, Vol. 1, n°4, pp. 81-91.
- Schneeweis and Spurgin (1998), "Multifactor analysis of hedge funds, managed futures and mutual fund return and risk characteristics", Journal of Alternative Investment, vol. 1, p. 1-24.
- Sharpe W. F. (1992), "Asset Allocation: Management Style and Performance Measurement", Journal of Portfolio Management, vol. 18, n° 2, pp. 7-19.

Appendix 1: Database statistics (december 2007)

	DE	FR	UK	LU	US
Currency	EUR	EUR	GBX	EUR	USD
Number of funds	187	254	287	552	867
Mean*	195.30	92.98	187.38	105.76	816.34
Median*	28.36	23.71	52.08	27.71	125.83
Standard deviation	653.68	214.54	407.84	306.79	3,475.04
Low*	0.19	0.02	0.08	0.01	0.07
High*	5,556.92	2,103.12	3,759.60	5,006.50	56,475.85

* market caps (millions)

Appendix 2: Average performance and risk by domicile

Absolute performance is measured by the difference between the net asset value at the end of the period and the net asset value at the beginning of the period for each fund. Absolute risk is measured by the standard deviation of the daily return for each fund.

Average Performance							
	MSCI World	Full sample	DE	FR	UK	LU	US
2000-2007	1.61%	−2.46%	−2.70%	−2.03%	−0.48%	−2.23%	−2.75%
2004 -2007	6.29%	4.77%	5.36%	4.91%	6.24%	5.33%	3.30%
2000 -2003	−4.41%	−6.67%	−7.65%	−6.92%	−6.36%	−7.15%	−6.07%
Average Risk							
	MSCI World	Full Sample	DE	FR	UK	LU	US
2000 -2007	14.09%	18.59%	18.91%	17.43%	17.20%	18.25%	19.54%
2004-2007	10.31%	13.24%	11.44%	11.89%	12.06%	12.28%	16.27%
2000 -2003	17.03%	21.87%	23.63%	21.36%	20.57%	22.02%	21.80%

Appendix 3: Average performance and risk by domicile

Table 2000-2007					
Eigenvalues					
				Cumulative	Cumulative
Number	Value	Difference	Proportion	Value	Proportion
1	4.057452	3.038042	0.6762	4.057452	0.6762
2	1.019410	0.616452	0.1699	5.076862	0.8461
3	0.402958	0.125802	0.0672	5.479820	0.9133
4	0.277156	0.082836	0.0462	5.756976	0.9595
5	0.194320	0.145616	0.0324	5.951296	0.9919
6	0.048704	---	0.0081	6.000000	1.0000

Eigenvectors (loadings),						
Variable	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6
LU	0.479988	−0.045328	−0.051985	−0.287789	0.200490	−0.801146
FR	0.422284	0.031658	−0.621491	0.656358	−0.038622	0.046095
US	0.389499	−0.527741	0.242328	0.011881	0.597252	0.392691
UK	0.425208	−0.272332	0.462134	0.134100	−0.716464	0.012705
DE	0.428706	0.291444	−0.319071	−0.633444	−0.172821	0.445362
MSCI	0.273836	0.747882	0.486775	0.258836	0.241690	0.057667

Appendix 3: Average performance and risk by domicile (suite)

Table 2004-2007

Eigenvalues

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	4.089842	3.020305	0.6816	4.089842	0.6816
2	1.069537	0.640043	0.1783	5.159379	0.8599
3	0.429494	0.096863	0.0716	5.588873	0.9315
4	0.332631	0.279591	0.0554	5.921504	0.9869
5	0.053039	0.027582	0.0088	5.974543	0.9958
6	0.025457	---	0.0042	6.000000	1.0000

Eigenvectors (loadings)

Variable	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6
FR	0.472878	−0.219805	−0.087940	0.051373	−0.640434	0.554566
DE	0.379489	−0.587173	0.185043	0.111768	0.648031	0.211042
UK	0.359967	0.419498	0.727930	−0.403678	0.019702	0.034906
LU	0.464018	−0.294160	−0.084212	−0.123394	−0.236527	− 0.787334
US	0.386205	0.369249	−0.648879	−0.415205	0.324015	0.126790
MSCI	0.371930	0.456136	−0.002667	0.796424	0.092629	−0.103582

Eigenvalues

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	4.285755	3.276433	0.7143	4.285755	0.7143
2	1.009323	0.675997	0.1682	5.295078	0.8825
3	0.333326	0.140665	0.0556	5.628404	0.9381
4	0.192661	0.036786	0.0321	5.821065	0.9702
5	0.155874	0.132814	0.0260	5.976939	0.9962
6	0.023061	---	0.0038	6.000000	1.0000

Eigenvectors (loadings)

Variable	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6
FR	0.427807	−0.009826	−0.643317	0.633528	−0.040072	0.007972
DE	0.412014	0.390083	−0.305758	−0.606109	−0.299778	0.356871
UK	0.422139	−0.259883	0.526697	0.203576	−0.658611	0.040468
LU	0.472106	−0.082745	−0.001358	−0.299201	0.192382	−0.802331
US	0.370362	−0.596694	0.076839	−0.154558	0.514800	0.460410
MSCI	0.330108	0.645995	0.457536	0.276446	0.415606	0.123408