# THE INFORMATION CONTENT OF HEDGE FUND INVESTMENT STYLES – A RETURN-BASED ANALYSIS WITH SELF-ORGANIZING MAPS

# Short Title: THE INFORMATION CONTENT OF HEDGE FUND INVESTMENT STYLES

RAMIN BAGHAI-WADJI,[a] RAMI EL-BERRY,[b] STEFAN KLOCKER[a] AND MARKUS SCHWAIGER[a,c]

[a] *Department of Corporate Finance, Institute of Finance and Financial Markets, Vienna University of Economics and Business Administration, Vienna, Austria*

[b] *E-Commerce Competence Center, Vienna, Austria*

[c] *Financial Markets Analysis and Surveillance Division, Oesterreichische Nationalbank, Vienna, Austria*

JEL-Classifications: C45, G11, G23

Keywords: Hedge Funds, Neural Networks, Self organizing map

**ABSTRACT**

Notwithstanding their common features, hedge funds remain an extremely diverse asset class. A consistent classification system is however important for numerous purposes such as portfolio construction, performance attribution as well as risk management. As fund self-declaration is prone to strategic misclassification, return based taxonomies grouping funds along similarities in realized returns can be used to avoid this pitfall. In this paper we use Self-Organizing Maps (SOM) to find homogeneous groups of hedge funds based on similar (return) characteristics. We can identify nine hedge fund classes – whereas managed futures, sector financial and short sell-hedge funds are largely consistent in their self declared strategies, we detect a number of declared hedge fund styles displaying no or very limited return similarities. Especially the so called "equity hedge"-style does not seem to be a useful self classification, or, put otherwise, encompasses too many different substyles with different return characteristics. The SOM furthermore detects similarities in a number of declared strategies, such as merger arbitrage funds and distressed securities funds. Another important aspect that our paper addresses is the tendency of fund managers to perform undisclosed changes of their trading style or to strategically misdeclare their funds. Our results show that so called "style creep" is an issue in the hedge fund business with funds which misclassified themselves once being very likely to change their trading style again, although our results do not support the hypothesis of style creep being driven by strategic style gaming.

---

The authors thank Stefan Bogner and Alois Geyer for helpful comments on earlier versions of the paper.

## 1. INTRODUCTION

The hedge fund universe consists of a great variety of completely different investment and trading strategies. Despite having some common features (e.g. an unregulated organizational structure, flexible investment strategies, sophisticated investors, etc.), hedge funds remain an extremely diverse asset class (see Ackermann et al., 1999). As a consequence, both practitioners and academicians are far from agreeing on a common hedge fund-classification system (see e.g. Brittain, 2001). While hedge fund index and database providers rely on their proprietary classification systems, academic research has just begun to adapt mutual fund based classification methodologies to the idiosyncrasies of the hedge fund business. A consistent classification system is however important for numerous reasons – it will help improve investment-choices of market participants, and funds of funds will refer to it in the construction of their portfolio to avoid undiversified exposures. A grouping of funds based on return characteristics can furthermore help evaluate the discriminatory power of different styles. In this context, a consistent classification system contributes to an improved performance attribution by peer group analysis (see e.g. Kandel et al.'s 2004 five factor model in this respect). It can also be useful in conceiving risk management models for hedge fund investments.

Several methods of fund-classification can be mentioned. The most evident one is fund self-declaration. One problem with this classification method is the so called "style gaming", i.e. the strategic misclassification of funds used to polish the fund's own performance with respect to its peers (see e.g. Brown and Goetzmann, 1997). Return based taxonomies avoid this pitfall by grouping funds along similarities in realized returns. Sharpe (1992) was the first to show that a regression of mutual fund returns on a limited number of indices can be used to specify different fund styles. Both Brown and Goetzmann (2003) and Fung and Hsieh (1997, 1998) adapted these models to the hedge fund universe. Whereas this methodology is well fit for traditional buy-and-hold long only investments, it is problematic in the case of hedge funds, as is well documented by Fung and Hsieh (1997), due to the unique features of hedge funds, namely dynamic trading strategies including alternating long and short positions that lead to an averaging error in a standard regression.[2] While the static Sharpe (1992) model implies time-invariant factor loadings, this is in clear conflict with earlier research documenting a relatively high degree of variability in hedge fund factor exposure over time, which could either be an indication of style gaming, or of an adaptation of strategies to changing market conditions (see e.g. Ennis and Sebastian, 2003, or McGuire et al, 2005). Alternatively, traditional statistical clustering approaches have been used to classify hedge funds to eschew some of these problems (see e.g. Miceli and Susinno, 2003, and Barès et al., 2001).

In contrast to findings for mutual funds where the self declared strategies are reasonably characteristic for underlying investment styles (such as. Brown and Goetzmann, 1997, or diBartolomeo and Witkowski, 1997), the evidence for style consistency in the hedge fund universe is rather mixed. For example, Barès et al. (2001) and Miceli and Susinno (2003), using traditional statistical clustering procedures, document that self declared hedge fund strategies are a reliable characterization of the underlying hedge fund

---

[2] Recently, contingent claims methodology has been shown to be of value for the classification and/or performance attribution of hedge fund strategies. The work of Fung and Hsieh (2001) and Mitchell and Pulvino (2001) show for trend following strategies and merger arbitrage strategies respectively, that option-like features in the strand of Glosten and Jagannathan (1994) capture the underlying risk return profile of hedge funds much better. See also Agarwal and Naik (2000, 2004) for a multi-factor approach to evaluate hedge fund performance which is based on option-strategies. Note however that, as already pointed out by Glosten and Jagannathan (1994), each strategy requires the use of different (compound) options, making this technique rather hard to handle for classification purposes.

styles.[3] Amenc and Martellini (2003), on the other hand, also perform cluster-based peer grouping on hedge funds and find that there is rather limited correlation between self-declared styles and their cluster-based classifications.[4]

In this paper, we employ a novel methodology to deal with the specifics of the hedge fund universe. We use Self-Organizing Maps (SOMs) to find homogeneous groups of hedge funds based on similar (return) characteristics. The SOM is a single-layered unsupervised neural network which maps data points from a higher dimensional space into a lower dimensional space using non-linear mapping functions. By employing an unsupervised neural network approach which has proven to be reliable in a myriad of disciplines[5], we are able to avoid a number of problems associated with the regression-based factor approach. As is documented in the literature, the SOM also leads to superior results vis-à-vis traditional statistical clustering approaches such as single linkage, complete linkage, median linkage and K-Means.[6] In our paper we demonstrate that the SOM-approach is perfectly suited for dynamic trading strategies, which previous models have been unable to deal with efficiently. Furthermore, in contrast to other approaches used in the literature on hedge fund style analysis (see for example

---

[3] Note however that due to the extremely low number of funds analysed in Miceli and Susinno (2003) - their sample only includes 62 funds – their results may suffer a rather severe sample selection bias. As for Bares et al. (2001), a clear disadvantage of their clustering approach is that they have to decide on the number of possible style clusters a priori.

[4] Amenc and Martellini (2003) also use a rather limited sample of 581 hedge funds; therefore, their analysis is prone to be affected by selection bias. Furthermore, their clustering approach is based on grouping hedge funds by style weight vectors obtained from Sharpe's (1992) style analysis technique, an approach which is also exposed to the aforementioned critique on Sharpe's (1992) method in hedge fund applications - it appears that grouping hedge funds directly via their return characteristics would be a more promising approach.

[5] In the field of finance, for example, applications include determining similarities in market timing strategies of investment newsletters (Kumar and Pons, 2002), stock picking (Deboeck and Ultsch, 2000), term structure modelling (De Bondt and Cottrell, 1998) as well as the classification of mutual funds (see Deboeck, 1998 and Moreno et al., 2002).

Brown and Goetzmann, 2003, and Barès et al., 2001) our SOM-based classification procedure does not assume the number of style categories to be known a priori. The number of styles is determined after the completion of the training process and is therefore not based on any premature assumptions.[7]

As most studies on hedge fund styles to date are based on samples of return histories up to the year 2000 only, and hedge funds have undergone a spectacular growth since then (see e.g. ECB, 2004 and SEC, 2003), it seems natural to ask whether results based on the hedge fund market from several years ago are representative enough for today's market environment.

To conclude, our method enables us to derive and visualize a consistent taxonomy for today's hedge fund market. This will provide us with answers to the following questions:

- Are self declared hedge fund styles a useful or misleading "label" ?
- Is there a connection between mislabeling and hedge fund survival?
- Do hedge funds change their styles over time, i.e. display the so called "style creep" and if so, is there any evidence for strategic "style gaming"?
- Are certain groups of funds particularly prone to misclassification and/or style creep?

---

[6] See for instance Mangiameli et al. (1996) for the superiority of Self-Organizing Maps as a clustering method for "messy data" sets where the number of clusters is assumed to be known and Ultsch and Vetter (1994) for the case when the number of clusters (homogeneous groups) in the data are assumed to be unknown a priori.

[7] In a related article, Maillet and Rousset (2003) had a first try at the use of SOM to classify hedge funds. Their results are however based on a very narrow sample of funds (294) and are thus likely to display a severe sample selection bias as is also acknowledged by the authors themselves. This may be one reason behind their failure to come up with a well trained map for hedge fund styles.

In our answers to these questions we can see that especially in the recent past, hedge funds have become less proficient at assigning themselves to a particular style as previous research suggests. Our results will help improve the choices of investors in terms of the construction of their portfolio, as well as contribute to an improved performance evaluation. Due to the opaque nature of the hedge fund business, which is based on proprietary (and secretive) trading strategies, getting the most out of available data seems all the more important for an informed investment decision.

The paper is organized as follows. In the next section we present a brief outline of the most important characteristics of the Self-Organizing Map. In section 3, we give an overview of the data and provide some summary statistics. Section 4 contains the main empirical results of our research. In section 5 the main conclusions are drawn.

## 2. METHODOLOGY

The Self-Organizing Map (SOM)[8] is an ideal tool for grouping and visualizing high-dimensional data; it is a single-layered unsupervised neural network which does not require any human intervention during the training process.[9] In the course of the training process, the SOM identifies the key features of the input space via a given set of input vectors. The SOM maps high-dimensional input data into a lower dimensional (usually two-dimensional, hence the term "map") output space while preserving the inherent structure of the original data input, thus allowing the visualization of complex data sets. Therefore, if two vectors are similar in terms of the distance measure employed, their images will end up in the vicinity of each other on the map. In the present paper, each hedge fund represents an input vector, the dimension of which is given by the number of monthly return observations. After the completion of the training process, hedge funds exhibiting similar return characteristics will be represented as homogeneous clusters on a two-dimensional surface.

The Self-Organizing Map consists of a single array of neural processing elements called nodes. Each node *i* has an associated reference vector $m_i \in \mathbb{R}^n$. In our case, the initial values of the reference vectors are chosen randomly. In each training pass *t*, an input vector $x(t)$ is drawn randomly from the input set and is compared with all reference vectors. The location of response is defined to be the node where the distance[10] between the input vector $x(t)$ and the reference vector $m_i(t)$ associated with that node achieves

---

[8] The Self-Organizing Map was originally developed by Teuvo Kohonen's research group and enhanced by many others since the initial publication of the material more than a quarter of a century ago (see Kohonen, 1997, for an exhaustive treatise on the subject).

[9] This characteristic distinguishes the SOM from the *supervised* neural network techniques where both input and output data are fed into the system; a network of that type is useful when a given input-output relationship has to be learned, but it is unsuitable for our research problem.

a minimum: $m_c(t) = \min_i \left\| x(t) - m_i(t) \right\|$. After $m_c(t)$, the reference vector corresponding to the so-called "winner node", has been determined, the value of its reference vector as well as that of its neighboring nodes is adjusted toward the value of the input vector $x$ - this is in fact what constitutes the learning process. Following the completion of the pre-specified number of training passes, each input vector is finally assigned to the trained node most similar in terms of the distance measure employed.

The aforementioned adjustments of the winner node $m_c$ and its neighbor nodes can be expressed in the following fashion: $m_i(t+1) = m_i(t) + \alpha(t)\left[x(t) - m_i(t)\right]$. This learning process is only applied to those nodes $m_i$ lying within a pre-specified distance from the winner $m_c$; the other nodes remain unchanged, i.e. $m_i(t+1) = m_i(t)$. The learning rate factor, $\alpha(t)$ with $0 < \alpha(t) < 1$, which establishes the magnitude of the adjustments, as well as the function defining the topological neighborhood of the winner node are both chosen to be monotonically decreasing in time (i.e. the number of completed training passes). [11]

It should be noted that the mapping process is not influenced by elements, i.e. return realizations at a given time, which exhibit similar values across all input vectors.[12]

---

[10] Euclidean distance is used in most practical applications as well as in the present case.

[11] For specifics regarding the SOM methodology, please refer to Kohonen (1997), Deboeck and Kohonen (1998) or the SOM_PAK documentation.

[12] If we consider for example the case that all input vectors (i.e. individual hedge funds in our case) feature a return close to 0.1 at element 15 (i.e. the $15^{th}$ observation within a fund's return history), then all trained reference vectors will have a value close to or equal to 0.1 at position 15. Therefore, the absolute distance between each input vector and all properly trained reference vectors with respect to element 15 will be very close to zero and hence does not contribute to the determination of the winner node.

From a more practical side, it should be mentioned that we use the original SOM_PAK library along with an adjusted version of the labeling algorithm of Merkl and Rauber (2001).[13]

[13] The SOM_PAK was downloaded from http://ftp.funet.fi/pub/sci/neural/cochlea/som_pak/.

## 3. DATA AND SUMMARY STATISTICS

Our paper is based on monthly return data from the CISDM (Center for International Securities and Derivatives Markets) hedge fund database, formerly known as the Managed Account Reports, Inc. (MAR) database. CISDM also provides a summary of the self-declared investment strategy and style for each fund. This proprietary classification will be compared to our neural network/return-based classification approach.

Our initial data set covers a ten year time period from May 1994 to April 2004 and comprises 4,231 hedge funds. In order to assure a sufficiently high degree of computational stability, we only include funds with a minimum of 24 monthly return observations in our sample, as recommended by Ackermann et al. (1999).[14] This eliminates 879 funds from our original data set. Furthermore, the fund of funds category is excluded from the analysis a priori in order to allow a focus on the "pure" trading strategies, which reduces our sample by another 853 funds. Following the same reasoning, we also exclude the 57 hedge fund indices of the original data set.

All of the above considered, this leaves us with a total sample of 2,442 funds[15]. It should be noted that our results are not subject to survivorship bias, as we include 844 non-surviving hedge funds in our analysis, i.e. funds which exhibit a minimum number of 24 observations but which have ceased to exist sometime within the period under observation. Table 1 summarizes our data sample.

[14] The requirement that a fund must have a sufficiently long return history for it to be included in the sample can give rise to a so-called "multi-period sampling bias." However, according to Fung and Hsieh (2000), the resulting upward performance bias is negligibly small. Ackermann et al. (1999) even find that for the data sample they use, the "multi-period sampling" requirement actually biases their statistics downwards. On the whole, the impact of a required return history 24 monthly observations appears to be of limited significance.

[15] 6 funds in the database whose style was labelled "unknown" were also excluded from our sample.

[Insert Table 1 about here]

## 4. EMPIRICAL ANALYSIS

### 4.1 Style consistency and misclassification

Our SOM-based classification procedure does not necessitate the specification of the number of style groups a priori. This allows us to determine the number of hedge fund style categories from the structure inherent in the data set. Training the SOM with the data sample described in the last section, we can identify nine hedge fund classes based on the number of clusters identifiable on the trained map's surface (see chart 1 for the resulting SOM and table 2 for a cross-tabulation of declared versus empirically confirmed hedge fund classes). Following Fung and Hsieh (1997) and Brown and Goetzmann (2003), the labeling is done according to the preponderance of managers of a given self-declared style in each group[16]: convertible arbitrage and fixed income (CA & FI), emerging markets (EM), managed futures (F), merger arbitrage and distressed securities (MA & DS), sector financial (SF), sector health care (SH), sector technology (ST), short selling (SS) and the class "other," which encompasses all funds that could not be included elsewhere.

[Insert Chart 1 about here.]

These classes occupy sections of different sizes on the map. Whereas managed futures emerge as a large group in this respect, spanning an extensive section of the map, other styles, such as the sector exposed ones (financial sector funds, healthcare sector funds, technology sector funds and short selling funds) occupy relatively little space. The size

---

[16] In order for a node to be labelled according to a given style, we require funds of this style to be the largest individual group of all fund styles mapped onto this node and to represent at least 40 % of all funds assigned to that node. Note that for equity hedge funds this procedure resulted in the identification of multiple smaller scattered clusters, which were not connected to each other. For this reason it was not possible to locate a single homogenous equity hedge cluster on the map.

information can be used to evaluate the degree of dispersion within each of the nine style groups identified, as Euclidean distance is used to depict return similarities on the map.

In contrast to Brown and Goetzmann (2003) or Miceli and Susinno (2003), and in line with Amenc and Martellini (2003), our findings suggest that a differentiated picture of the consistency of self-declared fund styles has to be drawn (see table 2). We can see that some hedge fund styles do a fairly good job of self-classification: Particularly short-sell and sector-financial hedge funds, as well as the category comprising managed futures are largely consistent in their self-declared strategies. In all of these cases, more than 65% of the respective funds are clustered in a meaningful way: The fund's self-labeling therefore has economic content in terms of a certain return pattern. Futures and short-sell strategies are especially well grouped by our map, with the percentage of correct self-declaration exceeding 79% in both cases. For managed futures, this underpins the hypothesis that idiosyncratic trading strategies reflected in their returns distinguish them quite substantially from other hedge fund styles.

[Insert Table 2 about here.]

For several other strategies, we see that a proprietary trading style emerges, but a considerable number of funds misdeclare themselves. In the case of merger arbitrage, convertible arbitrage and fixed income hedge funds, only 50% to 60% of the funds can be meaningfully grouped with their peers. Furthermore, distressed securities, emerging markets and sector technology funds exhibit a considerable amount of misclassification. The map recognizes these styles, but well over half of the funds pertaining to one of these self-declared groups are spread over other classes on the map. As a caveat it should, however, be mentioned that all of these styles occupy a rather limited surface on the map and are still able to capture a reasonable percentage of peers within these boundaries. Nonetheless, these results dictate caution in the investment choice and performance evaluation when dealing with the above fund classes.

Furthermore, we detect a number of declared hedge fund styles displaying no or very limited return similarities in our analysis. Especially the so-called "equity hedge" style does not seem to be a useful self-classification. Put differently, this style encompasses too many different substyles that convert the style into a misleading label – "equity hedge" funds are basically spread all over the plane. A similar argument applies for global macro, multi-sector and long-only funds: Although these funds are more concentrated in several regions of the map, they do not cluster into a homogeneous group. One conclusion therefore is that these fund categories encompass a variety of different substyles, i.e. category names are of limited informational value for the actual investment strategies used by these funds. Once again, caution in the construction of fund of funds and in performance attribution has to be exercised with these fund classes.

In addition to these consistency results, the SOM also detects similarities in a number of declared hedge fund strategies, so that these styles could be interpreted as substitutes in the construction of fund of funds portfolios. Merger arbitrage funds and distressed securities funds, for instance, emerge as a single style. Due to the digital nature of the underlying business (deal closure or not, and bankruptcy or not) and the fact that companies that are being taken over are often in a state of financial "distress," the vicinity of merger arbitrage and distressed securities funds seems to be perfectly rational from an economic point of view. Convertible arbitrage hedge funds and fixed income hedge funds also appear as a single style. Their exposure to bonds can be quoted as a reason for this result. Furthermore, funds with sector exposure (technology, health care, financial) are located in relatively remote sections of the map. The distance of these groups to managed futures, for instance, is in line with the economic rationale that these funds are driven by equity markets to a much greater extent than managed futures are. The map could therefore also be split in terms of equity market exposure, which seems to be important in the case of the lower and left section of the plane (see chart 1).

### 4.2 Hedge fund styles and return moments

Besides this general analysis of the map, we examined the mean, standard deviation, skewness, kurtosis as well as the Sharpe ratio of the funds located on the map. The most interesting insights come from a superposition of the funds' standard deviations onto the trained map (see chart 2). It can be clearly observed that the upper central part of the map (the area where convertible arbitrage, fixed income, merger arbitrage and distressed securities funds cluster) feature funds with relatively low standard deviations whereas the lower left corner (the area where sector technology, sector healthcare and emerging markets and numerous managed futures are concentrated) can be described as the high volatility section of the map. Chart 2 clearly shows that funds pertaining to the managed futures style group cover the entire range of standard deviations. This should not surprise in the light of the fact that different extents of leverage can easily be established with futures instruments due to the built-in lever of this asset class: Initial margins set by futures exchanges are fairly small in comparison to the nominal value of the contracts. A small change in the futures price therefore corresponds to a significant positive or negative return on the money invested in this instrument. In the case of managed futures funds, this extreme volatility inherent in futures contracts is typically reduced by dedicating only a specified percentage of the fund's assets to margin payments and investing the remainder into riskless treasury bills. The combination of high volatility inherent in the futures instruments with individual degree of de-levering employed by each fund (via riskless investments) results in the diverse volatility spectrum of managed futures observable on the SOM.

As far as the mean return is concerned, no specific pattern can be observed.[17] Relatively high Sharpe Ratios can be observed for those fund groupings displaying particularly low standard deviations, i.e. in particular merger arbitrage, distressed securities, sector financial funds, and even more so for convertible arbitrage and fixed income funds (see table 3).

[Insert Chart 2 about here.]

In terms of the third moment we generally observe negative skewness in the upper left section and positive skewness in the lower right section of the map. Hence, broadly speaking, we find that while hedge funds exhibit negative skewness, in general, managed futures clearly feature positive skewness. This observation corroborates the findings by Kat (2005) who also notes that the average hedge fund's return exhibits significant negative skewness, while futures returns tend to be positively skewed. However, the distinctness of this simple pattern disappears when it comes to kurtosis: When we superimpose the funds' kurtosis onto the map, we are unable to discern any clear tendency.

[Insert Table 3 about here.]

[17] However, it should be noted that the cluster "sector healthcare" exhibits noticeably higher means than other sections of the map.

### 4.3 Style misclassification and fund survival

A question which arises naturally in this context is whether funds which ceased to exist during the sample period were particularly bad at declaring their true investment strategy. It could be argued that funds which misclassify themselves experience a substantial withdrawal of investor resources and therefore perish. We can find no evidence for this hypothesis in our data. Table 4 shows a cross-tabulation of self-declared strategies with empirically confirmed strategies based on our analysis with the SOM trained with the whole data sample of 2,442 funds discussed earlier; in contrast to table 2, only "dead" funds are included in the table. The results show that on the whole, dead funds (or, more precisely, funds which cease to report) do not exhibit a more prominent mis-declaration of their investment style.[18]

[Insert Table 4 about here.]

[18] As a caveat it should be noted that not all funds which cease to report their returns to the database are in fact discontinued. Unfortunately, our data does not allow us to distinguish between funds that stopped reporting voluntarily and those which were liquidated.

### 4.4 Style Creep

In order to analyze the tendency of hedge funds to change their (return-based) styles over time, we split our sample into two consecutive five-year periods. For both sub-periods, separate maps are trained and analyzed. We require all funds to have no more than 10 missing observations in the combined 10 year period so that both maps are trained with a similarly exhaustive data pool of fund returns. This ensures that differences in the two maps will be solely due to changing return characteristics of the funds included in the sample rather than due to differences in the scope of the data fed into the neural network. Therefore, we exclude funds with less than 110 data points from our analysis to be able to follow the performance history of hedge funds more closely over our two five-year sub-periods and to guarantee enough overlapping returns for computational robustness. This leaves us with 459 funds in the “style creep” sample.

Tables 5 and 6 show the cross-tabulations resulting from the two five-year period maps. As outlined above, we restricted our sample quite rigorously to track the history of fund self-declaration. This restriction led to a lower dimensional map (10x10 fields vs. 20x20 fields) and hence, to fewer precisely discernable style groups emerging from the SOM classification process (six instead of nine). Compared to the cross-tabulation in table 2 for the ten-year period, the identification of fund styles that perform well in their self-classification and those that do not is largely consistent. Futures, short-sell and sector financial funds take the lead again, with equity hedge funds spread all over the map. By splitting the sample into two time segments, we can generally observe that hedge fund categories that were good/bad “classifiers” in the first sub-period remained so in the second sub-period. However, the percentage of misclassified funds has increased over time within each style class in all but two cases where this percentage remained constant. This seems to indicate that overall style inconsistencies of hedge funds were on the rise. With all the necessary caveats, this conclusion is further supported by the results from earlier research. For example, while Barès et al. (2001), using hedge fund data up to 1999, find that self-declared styles are mostly consistent with the empirically detected groupings, Amenc and Martellini (2003) use a more recent data set and find serious misclassifications in self-declared hedge fund investment styles.

[Insert Table 5 about here.]

[Insert Table 6 about here.]

It is furthermore noteworthy to see that the overall style consistency is lower in the ten-year sample period for 1994-2004 (table 2) than in the five-year period from 1999 to 2004 (table 6). The number of hedge funds has risen dramatically since the beginning of the millennium, as is well documented by the literature[19]. As we required a minimum of only 24 monthly return observations for the sample used in section 4.1, the 2,442 hedge funds analyzed in that section include a very large number of recently issued hedge funds (which are obviously not included in the sample used in the current section, where we required a minimum of 110 observations). This fact, combined with the insights gained in this section, corroborate our reasoning that, overall, style inconsistencies have risen in the hedge fund industry in recent years and that they are mostly driven by more recently issued, fledgling hedge funds. This argument is further supported by a more detailed analysis of the self-classification consistency of fledgling hedge funds. Using the SOM trained with the whole data sample of 2,442 funds (as described in section

[19] According to Fung and Hsieh (1997), there were between 1,000 to 2,000 hedge funds with $100-$160 billion in assets under management in 1994. In 2004, the hedge fund industry comprised around 7,000 funds with $795 billion in assets (Risk Magazine, September 2004, p.9). See also ECB (2004) and SEC (2003) in this respect.

4.1), we looked at all hedge funds that started after the beginning of the year 2000 in order to determine their ability to correctly self-declare their investment style. Performing the cross-tabulation in analogy to table 2 with the resulting 505 fledgling hedge funds, we can see that except for the self-declared categories convertible arbitrage and distressed securities, the consistency values are far lower for fledgling hedge funds than for the whole sample. To quote one particularly striking example, out of the 11 fledgling sector technology funds started after January 2000, not a single one was captured in our SOM-based sector technology category, indicating a rather severe amount of mis-classification of the fledgling sector technology funds.

Despite the general rise of style inconsistencies over time, some fund styles (e.g. distressed securities and short selling funds) feature the same number of correctly declared funds in both sub-periods. The apparent congruence could be either due to funds indeed sticking to their declared style over time; or, alternatively, it could be caused by funds switching from one category to another in a fashion which leaves the aggregate picture unaffected. In order to clarify this point and to analyze style creep in more detail, we follow each fund individually to see whether there was a change in the SOM-based style classification from the first period to the second. Table 7 summarizes these results for individual fund groups. These results indicate that style creep is an issue in the hedge fund industry, with more than 23% of funds changing style over our observation period. However, style creep is not as prevalent as it is in the mutual fund industry (see e.g. Kim et al., 2000, or Gallo and Lockwood, 1999). Overall, it is noteworthy to see that a marked difference in the tendency towards style creep exists between funds that declare themselves correctly (fourth row in table 7) and all funds within a given style category (second row in table 7). The ex post observed probability of a style change is halved (23.3% against 11.7%) in the case of hedge funds that declare themselves correctly in the first sub-period, indicating that, overall, funds which misdeclare themselves once are prone to change their style again. In a nutshell, one is therefore tempted to conclude "don't trust a liar."

Style creep in the different fund categories corroborates this argument. Those fund classes which have high consistency values in their self-declared styles are less inclined to change style over time. Futures e.g. seem to be fairly consistent in their intertemporal investment style. Emerging market funds on the contrary seem to be quite inclined to alter their style, whereas for sector financial and short-sell hedge funds the style creep tendency is high for the entire sample but improves markedly for funds that correctly self-classify. As a caveat, it should, however, be considered that not all fund categories occupy the same surface on the map. As Euclidean distance serves as a proxy for similarity, comparatively minor deviations in return characteristics appear as style creep in fund classes spanning only a small surface on the map such as short sell (SS) and sector financial (SF). To sum up the evidence gathered, our analysis documents the presence of style creep in the hedge fund universe, with those funds that misclassify being more inclined to change style.

[Insert Table 7 about here.]

### 4.5 Style Gaming

As can be seen from the last section, our empirical findings clearly confirm the existence of style creep. However, the cause of this phenomenon is less evident. One reason for style creep could lie in funds gaming their declared style (see Brown and Goetzmann, 1997, for example), with false self-declaration of the investment strategy being attributed to the investment manager's hope to manipulate performance evaluation for the better. Style gaming could either manifest itself in misdeclaring funds clearly outperforming their self-declared peers or, less dramatically, in a less severe underperformance compared to peers. In the first case, we would expect returns to be higher for funds that declared themselves incorrectly compared to those which accurately declared their investment strategy. In other words, we would expect to see higher median returns for misdeclaring funds in any given category. Secondly, we also have to consider the possibility that misdeclaring funds engage in style gaming in order to improve their relative performance compared to their peers. To make this point more concrete, suppose a self-declared CA fund follows an investment strategy similar to SF funds and is therefore placed by the SOM in the SF cluster. According to our methodology, this constitutes a false self-declaration. Even if the fund manager cannot outperform other CA funds, he could be actually even worse off in terms of performance relative to peers if he correctly declared himself to follow a SF investment style. Hence he would rather stick to the false self-declaration. This would also constitute a form of style gaming.

In order to analyze the first scenario, we compare the funds' median returns by category using the whole sample of 2,442 funds. In this case, we do not find any evidence to support the style gaming hypothesis - the median of the arithmetic means of the individual funds' returns within a given category is in fact higher for funds which have correctly declared themselves[20] (see Table 8).

[Insert Table 8 about here.]

As has been outlined above, there could still be a case for style gaming even if misdeclaring funds do not outperform correctly declared funds. The possibility remains that they adopt / keep an incorrect style label because they would be even worse off if their performance was to be compared to their actual investment style peers identified by the SOM. In order to elaborate on this issue, we ranked the performance of misclassified hedge funds vis-à-vis their self-declared peers as well as vis-à-vis those funds which follow the same empirically confirmed investment strategy. The results can be seen in table 9: For each empirically confirmed investment style, we sorted the average monthly returns of funds into 4 quartiles, ranging from best performance (first quartile) to worst performance (fourth quartile). The column labeled "dec" contains the ranking of misdeclaring funds vis a vis their self-declared peers while the column "act" shows the ranking vis-à-vis funds following the same empirically confirmed strategy. For example, only 3 out of 24 hedge funds which falsely declared themselves to be emerging market (EM) funds were ranked in the top quartile in terms of return performance when compared with their self-declared emerging market peers. Suppose however that the falsely declared EM funds decided to correctly proclaim their investment style. In that case, 9 out of the 24 EM funds would boast a performance in the top quartile (when performance is measured relative to their actual investment style peers). The overall picture allows us to draw the conclusion that no form of style

[20] This is true for all style categories except "short selling". However, note that there are only three short selling hedge funds which have misdeclared themselves.

gaming is supported by our findings: In most cases, misdeclaring hedge funds would be better off rightfully disclosing their investment style since their performance relative to their actual investment style peers would be superior to their performance compared to their self-declared peers.[21]

[Insert Table 9 about here.]

As the observed underperformance of misdeclaring funds vis-à-vis their self-declared peers is not in line with the style gaming argument commonly cited in the literature, an alternative explanation is called for. An attempt to understand our empirical findings might be the following: Suppose, for example, that the majority of the fund managers who declared themselves incorrectly, i.e. proclaimed to pursue a certain investment strategy but in fact exhibit atypical investment behavior, are inexperienced or "untalented". In other words, they try to emulate a specific investment style but fail.[22] In that case, we would expect their performance compared to their peers who have correctly declared themselves and also know how to master a specific investment strategy to be inferior.

[21] To complete our investigation we also examined differences in other moments between correctly declaring and misdeclaring funds. However no useful pattern emerged from this analysis.

[22] This hypothesis is in fact largely supported by our data, if one considers that many fund managers of new (or "fledgling") hedge funds are inexperienced: As has been outlined in section 4.4, the 505 "fledgling" hedge funds in our sample do exhibit more prominent self-misdeclaration of their investment styles when compared to funds with longer return histories.

## 5. CONCLUSION

Despite having some common features, hedge funds remain an extremely diverse asset class. So far no commonly accepted hedge fund taxonomy has emerged. In this paper we provide a classification of hedge fund styles by detecting hedge fund groupings with similar return characteristics on the basis of Self-Organizing Maps (SOMs) that avoids the problems of factor based style analysis. Furthermore we analyze the phenomena of style creep and style gaming within the hedge fund universe.

Based on a ten-year sample of 2,442 active and dead hedge funds, we can identify nine hedge fund classes. Earlier findings which document a fairly adequate self-classification of hedge funds (such as Brown and Goetzmann, 2003, and Miceli and Susinno, 2003) can only be partially confirmed. The reliability of the declared classification substantially differs between various fund styles. Whereas managed futures and short-sell hedge funds are very consistent in their self-declared strategies, other hedge fund groups (such as fixed income, convertible arbitrage, merger arbitrage, distressed securities, sector technology and sector healthcare funds) exhibit an only moderate aptitude in correctly classifying themselves. Moreover, our results show that several declared hedge fund styles have hardly any similarities and are thus a rather useless label with very diverse return patterns incorporated in these funds (a case in point would be the equity hedge category). The SOM furthermore detects similarities in a number of declared strategies. No connection between mislabeling and fund survival could be found.

We also document that style inconsistencies have been on the rise over time, with young, fledgling hedge funds driving this increase. Furthermore, our results suggest that so-called style creep is an issue in the hedge fund universe. It is readily observable in the case of funds belonging to style categories which are particularly prone to erroneous self-classification, e.g. emerging market and equity hedge funds. It appears that hedge funds belonging to categories which are poor self-classifiers change their (return-based) investment style rather often whereas funds pertaining to more homogeneous categories, such as managed futures or short sell funds, exhibit more stable and consistent investment behavior. While we do find evidence for style creep, our data does not corroborate the hypothesis of funds strategically gaming their style in order to improve their track record vis-à-vis their peers.

Our results are important for a number of purposes. For instance, they can help avoid undiversified exposures to certain styles in the construction of fund of fund portfolios. Furthermore, a consistent classification can be useful in the construction of benchmarks and thus assist performance attribution. Moreover, fund investors might be interested in their exposure to different fund styles for risk management purposes. In this context, our results help in the construction of diversified portfolios and thereby enhance the risk-sharing among participants of financial markets.

## REFERENCES


Ackermann C, McEnally R, Ravenscraft D. 1999. The Performance of Hedge Funds: Risk, Return, and Incentives. *Journal of Finance* **54**(3): 833-874.

Agarwal V, Naik NY. 2000. Multi-Period Performance Persistence Analysis of Hedge Funds. *Journal of Financial and Quantitative Analysis* **35**(3): 327-342.

Agarwal V, Naik NY. 2004. Risk and Portfolio Decisions involving Hedge Funds. *Review of Financial Studies* **17**(1): 63-99.

Amenc N, Martellini L. 2003. The Alpha and Omega of Hedge Fund Performance Measurement. Working Paper: EDHEC.

Barès PA, Gibson R, Gyger S. 2001. Style Consistency and Survival Probability in the Hedge Funds' Industry. Working Paper: Swiss Institute of Banking.

Brealey RA, Kaplanis E. 2001. Hedge Funds and Financial Stability: An Analysis of their Factor Exposures. *Journal of International Finance* **4**(2): 161 – 187.

Brittain WHB. 2001. Hedge Funds and the Institutional Investor. *Journal of International Financial Management and Accounting* **12**(2): 225-234.

Brown SJ, Goetzmann WN. 1997. Mutual fund styles. *Journal of Financial Economics* **43**(3): 373-399.

Brown SJ, Goetzmann WN. 2003. Hedge Funds with Style. *Journal of Portfolio Management* **29**(2): 101-113.

Buetow Jr GW, Johnson RR, Runkle DE. 2000. The Inconsistency of Return-Based Style Analysis. *Journal of Portfolio Management* **26**(3): 61-77.

Carhart M. 1997. On Persistence in Mutual Fund Performance. *Journal of Finance* **52**(1): 57-82.

Deboeck G, Ultsch A. 2000. Picking stocks with emergent self-organizing value maps. *Neural Net World* **10**: 203-216.

Deboeck G. 1998. Picking mutual funds with self-organizing maps. In *Visual Explorations in Finance with Self-Organizing Maps*, Deboeck G, Kohonen T (eds). Springer; 39-57.

Deboeck G, 1999. Self-organizing Maps Facilitate Knowledge Discovery in Finance. *Financial Engineering News* **8**.

Deboeck G, Kohonen T (eds). 1998. Visual Explorations in Finance with Self-Organizing Maps. Springer.

De Bondt E, Cottrell PG and M. 1998. Projection of long-term interest rates with maps. In *Visual Explorations in Finance with Self-Organizing Maps*, Deboeck G, Kohonen T (eds). Springer; 24-38.

diBartolomeo D, Witkowski E. 1997. Mutual Fund Misclassification: Evidence Based on Style Analysis. *Financial Analysts Journal* **53**(5): 32-43.

Ennis RM, Sebastian MD. 2003. A Critical Look at the Case for Hedge Funds. *Journal of Portfolio Management* **29**(4): 103-112.

European Central Bank, ECB. 2004. Growth of the Hedge Fund Industry: Financial Stability Issues. *ECB Financial Stability Review* 2004: 123-132.

Fama EF, French KR. 1992. The Cross-Section of Expected Stock Returns. *Journal of Finance* **47**(2): 427-465.

Fung, W., Hsieh, D. A., 1997: Empirical Characteristics of Dynamic Trading Strategies: The Case of Hedge Funds. Review of Financial Studies **10**(2), 275-302.

Fung W, Hsieh DA. 1998. Performance Attribution and Style Analysis: From Mutual Funds to Hedge Funds. Working Paper: Duke Fuqua School of Business.

Fung W, Hsieh DA. 2000. Performance Characteristics of Hedge Funds and Commodity Funds: Natural vs. Spurious Biases. *Journal of Financial and Quantitative Analysis* **35**(3): 291-307.

Fung W, Hsieh DA. 2002. Asset-Based Style Factors for Hedge Funds. *Financial Analysts Journal* **58**(5): 16-27.

Furfine C. 2001. The costs and benefits of moral suasion: Evidence from the rescue of LTCM. BIS Working Paper No. 103.

Gallo JG, Lockwook LJ. 1999. Fund management changes and equity style shifts. *Financial Analysts Journal* **55**(5): 44-52.

Glosten L, Jagannathan R. 1994. A contingent claim approach to performance evaluation. *Journal of Empirical Finance* **1**(2): 133-160.

Goetzmann WN, Ingersoll J, Ross SA. 2003. High Water Marks and Hedge Fund Management Contracts. *Journal of Finance* **58**(4): 1685-1717.

Haykin S. 1998. Neural Networks: A Comprehensive Foundation. Prentice Hall.

Humayun KM, Hassan M.K. 2004. The near collapse of LTCM, US financial stock returns, and the fed. *Journal of Banking and Finance* **29**(2): 441-460.

Jorion P. 2000. Risk management lessons from long-term capital management. *European Financial Management* **6**(3): 277-300.

Kandel E, Kandel S, Wermers R. 2004. Endogenous Benchmarks. Working Paper: University of Maryland.

Kat HM. 2005. Integrating Hedge Funds into the Traditional Portfolio. *Journal of Wealth Management* **7**(4): 51-58.

Kho BC, Lee D, Stulz R. 2000. U.S. banks, rises and bailouts: From Mexico to LTCM. *American Economic Review* **90**(2): 28-31.

Kim M, Shukla R, Tomas M. 2000. Mutual fund objective misclassification. *Journal of Economics and Business*, **52**(4): 309-323.

Kohonen T. 1997. Self-organizing maps. Springer.

Kohonen T, Hynninen J, Kangas J, Laaksonen J. 1995. SOM_PAK, The self-organizing map program package. Helsinki University of Technology, Laboratory of Computer and Information Science.

Kumar A, Pons V. 2002. Behavior and performance of investment newsletters. Working Paper: Yale University.

Lamont O. 1997. Cash flow and investment: Evidence from internal capital markets. *Journal of Finance* **52**(1): 83-110.

Mangiameli P, Chen SK, West D. 1996. A comparison of SOM neural network and hierarchical clustering methods. *European Journal of Operational Research* **93**(2): 402-417.

Maillet B, Tousset P. 2003. Classifying Hedge Funds with Kohonen Maps: A First Attempt. Computational Economics, Series in Advances in Computational Economics and Management Sciences, Kluwer Academic Press.

McGuire P, Remolona E, Tsatsaronis K. 2005. Time-varying exposures and leverage in hedge funds. *BIS Quarterly Review* (March 2005): 59-72.

Merkl D, Rauber A. 2001. Automatic Labeling of Self-Organizing Maps for Information Retrieval. *Journal of Systems Research and Information Systems* **10**(10): 23-45.

Mitchell M, Pulvino T. 2001. Characteristics of Risk and Return in Risk Arbitrage. *Journal of Finance* **56**(6): 2135-2175.

Miceli MA, Susinno G. 2003. Using trees to grow money. *Risk* **16**(11): 11-12.

Moreno D, Olmeda I, Marco P. 2002. A new look on classification of Spanish mutual funds. Working Paper: Universidad de Alcala.

Rauber A, Pampalk E, Merkl D. 2003. The SOM-enhanced JukeBox: Organization and Visualization of Music Collections based on Perceptual Models. *Journal of New Music Research* **32**(2): 193-210.

Risk Magazine, September 2004.

Sharpe WF. 1992. Asset allocation: Management style and performance measurement. *Journal of Portfolio Management* **18**(2): 7-19.

SEC, Securities and Exchange Commission. 2003. Implications of the Growth of Hedge Funds. Staff Report to the United States Securities and Exchange Commission.

Ultsch A, Vetter C. 1994. Self-Organizing-Maps versus Statistical Clustering Methods: A Benchmark. Research Report 0994, FG Neuroinformatik und Künstliche Intelligenz, University of Marburg.

**Table 1:** Hedge fund sample and summary statistics by self-declared investment strategy.

| **Self-Declared Investment Style** | **Number of Funds** | **Median Return (in %)** | **Standard Deviation (in %)** | **Skewness** | **Kurtosis** |
|---|---|---|---|---|---|
| Convertible Arbitrage | 136 | 0.9678 | 1.5529 | 0.0403 | 5.1186 |
| Distressed Securities | 74 | 1.0662 | 2.9520 | -0.1006 | 5.1746 |
| Equity-Hedge | 825 | 1.0723 | 4.0331 | 0.1777 | 4.6194 |
| Emerging Markets | 133 | 1.2577 | 6.0945 | -0.0822 | 5.8200 |
| Managed Futures | 821 | 1.0074 | 5.5850 | 0.5036 | 4.3447 |
| Fixed Income | 80 | 0.7400 | 1.8282 | -0.8294 | 6.2998 |
| Global Macro | 76 | 1.0026 | 4.5128 | 0.1337 | 4.5637 |
| Merger Arbitrage | 114 | 0.7601 | 1.7013 | -0.0568 | 5.3237 |
| Sector Financial | 26 | 1.4382 | 3.9381 | -0.3343 | 5.6261 |
| Sector Healthcare | 28 | 2.1160 | 7.8693 | 1.3684 | 6.8277 |
| Short Selling | 25 | 0.4583 | 6.7149 | 0.0521 | 4.2446 |
| Sector Technology | 46 | 1.4915 | 8.5734 | 0.3997 | 3.5473 |
| Sector Multi Sector | 27 | 1.1548 | 4.0377 | 0.2504 | 4.7079 |
| Long Only | 17 | 1.5720 | 7.1496 | 0.0046 | 4.3718 |
| Sector Energy | 7 | 1.3833 | 8.4937 | -0.0078 | 3.8907 |
| Sector Real Estate | 7 | 1.0052 | 2.3364 | -0.6346 | 6.6510 |

Notes: The numbers given correspond to the monthly category medians (e.g. “Median Return” denotes the category median of the arithmetic means of the time series of the individual hedge funds).

**Table 2:** Cross-tabulation of self-declared strategies (columns) with empirically confirmed strategies (rows).

| | CA | DS | EH | EM | F | FI | GM | MA | SF | SH | SR | SS | ST | SMS | LO | *Total*** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **CA&FI** | 54.4 | 21.6 | 4 | 8.3 | 1.8 | 57.5 | 10.5 | 8.8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | *213* |
| **DS&MA** | 11 | 28.4 | 5.8 | 5.3 | 1.3 | 3.8 | 7.9 | 50.9 | 0 | 0 | 28.6 | 0 | 0 | 0 | 0 | *171* |
| **EM** | 1.5 | 1.4 | 4.3 | 42.1 | 0.4 | 0 | 2.6 | 0.9 | 0 | 0 | 0 | 0 | 6.5 | 0 | 5.9 | *105* |
| **F** | 1.5 | 4.1 | 6.1 | 1.5 | 79.5 | 11.3 | 35.5 | 2.6 | 3.8 | 0 | 14.3 | 0 | 8.7 | 0 | 0 | *755* |
| **SF** | 0 | 0 | 1.6 | 0.8 | 0.2 | 0 | 1.3 | 0 | 65.4 | 7.1 | 14.3 | 0 | 0 | 0 | 0 | *37* |
| **SH** | 0 | 0 | 1.2 | 0 | 0 | 0 | 1.3 | 0 | 0 | 53.6 | 0 | 0 | 0 | 14.8 | 0 | *30* |
| **SS** | 0 | 0 | 1.8 | 0 | 0.5 | 2.5 | 0 | 0 | 0 | 0 | 0 | 88 | 0 | 0 | 0 | *43* |
| **ST** | 0 | 1.4 | 1.9 | 0.8 | 0 | 0 | 1.3 | 0.9 | 0 | 3.6 | 0 | 0 | 39.1 | 3.7 | 23.5 | *44* |
| **other** | 31.6 | 43.1 | 73.3 | 41.4 | 16.3 | 24.9 | 39.6 | 35.9 | 30.8 | 35.7 | 42.8 | 12 | 45.7 | 81.5 | 70.6 | *1044* |
| **Sum*** | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | |
| ***Total**** | *136* | *74* | *832* | *133* | *821* | *80* | *76* | *114* | *26* | *28* | *7* | *25* | *46* | *27* | *17* | ***2442*** |

* in percent

** total number of funds in a given category

Notes: The numbers given are percentage points. The abbreviations denote the following: CA (Convertible Arbitrage), DS (Distressed Securities), EH (Equity Hedge), EM (Emerging Markets), F (Managed Futures), FI (Fixed Income), GM (Global Macro), MA (Merger Arbitrage), SF (Sector Financial), SH (Sector Healthcare and Biotechnology), SS (Short Sell), SMS (Sector Multi-Sector), SR (Sector Real Estate), ST (Sector Technology), LO (Long Only).

**Table 3:** Sharpe Ratios of empirically confirmed hedge fund strategies.

| Category | Number of Funds | Annualized Sharpe Ratios | Monthly Sharpe Ratios |
|---|---|---|---|
| Convertible Arbitrage & Fixed Income | 213 | 1.6689 | 0.4818 |
| Merger Arbitrage & Distressed Securities | 172 | 0.9766 | 0.2819 |
| Emerging Markets | 105 | 0.4701 | 0.1357 |
| Futures | 759 | 0.4640 | 0.1339 |
| Sector Financial | 37 | 1.0843 | 0.3130 |
| Sector Healthcare | 30 | 0.7251 | 0.2093 |
| Sector Technology | 58 | 0.2471 | 0.0713 |
| Short Selling | 43 | 0.1723 | 0.0498 |

Notes: The numbers given are the category medians.

**Table 4:** Cross-tabulation of self-declared strategies (columns) with empirically confirmed strategies (rows); only dead funds are considered.

| | CA | DS | EH | EM | F | FI | GM | MA | SF | SH | SR | SS | ST | SMS | LO | *Total*** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **CA&FI** | 39.3 | 17.6 | 3.4 | 2.4 | 1.6 | 71.4 | 9.1 | 8.6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | *56* |
| **DS&MA** | 14.3 | 23.5 | 7.4 | 7.3 | 1.9 | 0 | 9.1 | 54.3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | *61* |
| **EM** | 7.1 | 5.9 | 4.3 | 39.1 | 0.3 | 0 | 3 | 2.9 | 0 | 0 | 0 | 0 | 5.3 | 0 | 11.1 | *37* |
| **F** | 0 | 0 | 5.7 | 0 | 73.8 | 0 | 21.2 | 0 | 0 | 0 | 50 | 0 | 0 | 0 | 0 | *256* |
| **SF** | 0 | 0 | 2.4 | 0 | 0 | 0 | 3 | 0 | 33.3 | 0 | 0 | 0 | 0 | 0 | 0 | *9* |
| **SH** | 0 | 0 | 2.4 | 0 | 0 | 0 | 3 | 0 | 0 | 85.7 | 0 | 0 | 0 | 0 | 0 | *14* |
| **SS** | 0 | 0 | 1.3 | 0 | 0.3 | 0 | 0 | 0 | 0 | 0 | 0 | 87.5 | 0 | 0 | 0 | *12* |
| **ST** | 0 | 5.9 | 4.3 | 2.4 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 57.9 | 50 | 33.3 | *32* |
| **other** | 39.3 | 47.1 | 68.8 | 48.8 | 22.1 | 28.6 | 48.6 | 34.2 | 66.7 | 14.3 | 50 | 12.5 | 36.8 | 50 | 55.6 | *367* |
| **Sum*** | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | |
| ***Total**** | *28* | *17* | *297* | *41* | *313* | *28* | *33* | *35* | *3* | *7* | *2* | *8* | *19* | *4* | *9* | ***844*** |

* in percent

** total number of dead funds in a given category

Notes: The numbers given are percentage points.

**Table 5:** Cross-tabulation of self-declared strategies (columns) with empirically confirmed proprietary strategies (rows) for the balanced sample of funds from May 1994 to April 1999.

| | CA | DS | EH | EM | F | FI | GM | MA | SF | SH | SS | ST | SMS | LO | *Total*** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **CA, DS & MA** | 95 | 79 | 31 | 0 | 3 | 25 | 0 | 83 | 0 | 0 | 0 | 0 | 20 | 0 | *92* |
| **EM** | 0 | 0 | 1 | 78 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | *16* |
| **F** | 0 | 7 | 5 | 6 | 91 | 75 | 30 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | *224* |
| **SF** | 5 | 0 | 9 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | *17* |
| **SS** | 0 | 0 | 6 | 6 | 1 | 0 | 0 | 0 | 0 | 0 | 83 | 0 | 0 | 0 | *15* |
| **Other** | 0 | 14 | 48 | 11 | 6 | 0 | 70 | 17 | 0 | 100 | 17 | 100 | 80 | 100 | *95* |
| **Sum*** | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | |
| ***Total*** | *19* | *14* | *116* | *18* | *232* | *4* | *10* | *23* | *6* | *1* | *6* | *3* | *5* | *2* | ***459*** |

* in percent

** total number of funds in a given category

**Table 6**: Cross-tabulation of self-declared strategies (columns) with empirically confirmed proprietary strategies (rows) for the balanced sample of funds from May 1999 to April 2004.

| | CA | DS | EH | EM | F | FI | GM | MA | SF | SH | SS | ST | SMS | LO | *Total*** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **CA, DS & MA** | 79 | 79 | 17 | 17 | 5 | 25 | 0 | 74 | 0 | 0 | 0 | 0 | 0 | 0 | *79* |
| **EM** | 0 | 0 | 5 | 50 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | *16* |
| **F** | 16 | 0 | 11 | 6 | 86 | 75 | 30 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | *224* |
| **SF** | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 83 | 0 | 0 | 0 | 0 | 0 | *11* |
| **SS** | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 83 | 0 | 0 | 0 | *9* |
| **Other** | 5 | 21 | 58 | 28 | 9 | 0 | 70 | 17 | 17 | 100 | 17 | 100 | 100 | 100 | *120* |
| **Sum*** | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | |
| ***Total*** | *19* | *14* | *116* | *18* | *232* | *4* | *10* | *23* | *6* | *1* | *6* | *3* | *5* | *2* | ***459*** |

* in percent

** total number of funds in a given category

**Table 7**: Style Creep by Hedge Fund Class.

| | **EM** | **SF** | **CA, MA, DS** | **SS** | **F** | **other** | **OVERALL** |
|---|---|---|---|---|---|---|---|
| **Number Creep**[1] | 7 | 8 | 31 | 6 | 27 | 28 | 107 |
| **Percentage Creep**[2] | 43.80% | 47.10% | 33.70% | 40% | 12.10% | 29.47% | 23.31% |
| **Number Declared Creep**[3] | 5 | 1 | 7 | 0 | 20 | … | 33 |
| **Percentage Declared Creep**[4] | 35.70% | 16.70% | 14.60% | 0% | 9.50% | … | 11.66% |
| **Number Correctly Declared**[5] | *14* | *6* | *48* | *5* | *210* | … | 283 |
| ***Total***[6] | *16* | *17* | *92* | *15* | *224* | *95* | *459* |

[1] Based on the mapping results for the 1994–1999 sub-period, number of funds within a given category which changed their affiliation in the 1999–2004 sub-period.

[2] Percentage of funds that changed their affiliation in the 1999–2004 sub-period.

[3] Number of funds which correctly classified themselves in the 1994–1999 sub-period and subsequently changed their affiliation in the 1999-2004 sub-period.

[4] Percentage of funds (with respect to the number of correctly classified funds in a given category) which correctly classified themselves in the 1994–1999 sub-period but changed their affiliation in the 1999–2004 sub-period.

[5] Number of funds which correctly classified themselves in the 1994-1999 sub-period.

[6] Total number of funds within a given category.

**Table 8:** Group median of the monthly mean return (in percent) by accuracy of self-declaration.

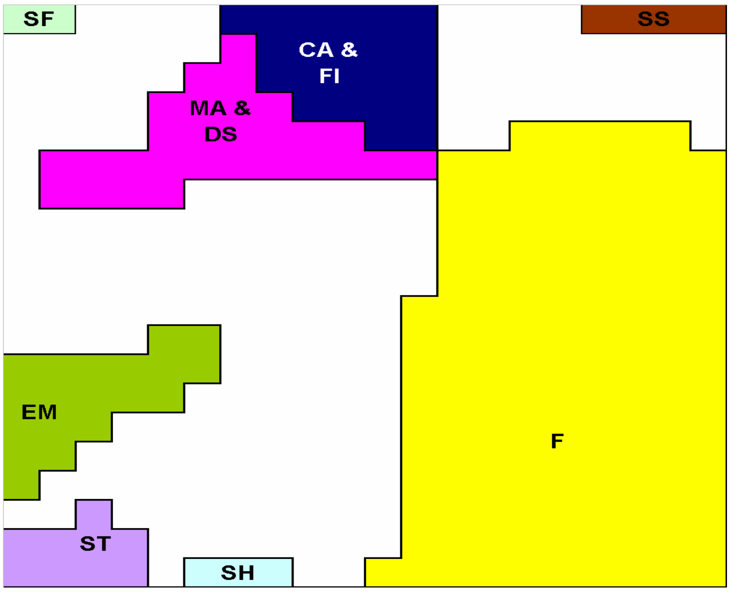
| | Right Self-Declaration | | False Self-Declaration | |
|---|---|---|---|---|
| **Self-Declared Investment Style** | **Number of Funds** | **Median Return (in %)** | **Number of Funds** | **Median Return (in %)** |
| Convertible Arbitrage | 74 | 1.0021 | 62 | 0.9029 |
| Distressed Securities | 21 | 1.4382 | 53 | 1.0255 |
| Emerging Markets | 56 | 1.4768 | 77 | 1.1566 |
| Managed Futures | 655 | 1.0614 | 166 | 0.7890 |
| Fixed Income | 46 | 0.8149 | 34 | 0.7147 |
| Merger Arbitrage | 58 | 0.7768 | 56 | 0.7222 |
| Sector Financial | 17 | 1.4679 | 9 | 1.4085 |
| Sector Healthcare | 15 | 2.3194 | 13 | 1.5893 |
| Short Selling | 22 | 0.4222 | 3 | 0.9081 |
| Sector Technology | 19 | 1.8213 | 27 | 1.1486 |

**Table 9:** Relative performance of falsely declared hedge funds vis a vis self-declared ("dec") and actual ("act") peers. The 1[st] Quartile corresponds to the best performing funds, the 4[th] Quartile to worst performing funds.

| | **CA & FI** | | **EM** | | **F** | | **MA & DS** | | **SF** | | **SH** | | **SS** | | **ST** | | **OVERALL** | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **dec** | **act** | **dec** | **act** | **dec** | **act** | **dec** | **act** | **dec** | **act** | **dec** | **act** | **dec** | **act** | **dec** | **act** | **dec** | **act** |
| **1st Quartile** | 9 | 6 | 3 | 9 | 5 | 9 | 1 | 2 | 0 | 0 | 2 | 2 | 0 | 0 | 1 | 1 | 21 | 29 |
| **2nd Quartile** | 4 | 8 | 6 | 6 | 12 | 12 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 24 | 28 |
| **3rd Quartile** | 11 | 10 | 10 | 6 | 10 | 5 | 2 | 2 | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 2 | 35 | 26 |
| **4th Quartile** | 9 | 9 | 5 | 3 | 8 | 9 | 10 | 10 | 1 | 0 | 1 | 1 | 0 | 0 | 4 | 3 | 38 | 35 |
| **Sum** | 33 | 33 | 24 | 24 | 35 | 35 | 15 | 15 | 1 | 1 | 3 | 3 | 0 | 0 | 7 | 7 | 118 | 118 |

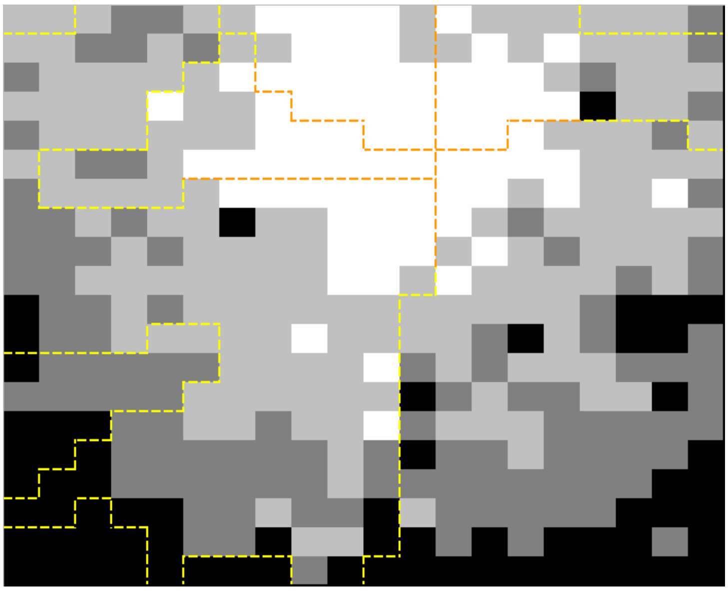
Notes: For clarity of exposition, only those hedge funds were included in the analysis which misdeclared their investment style *and* clustered with some other style category. Hence, funds which misdeclared themselves and fell into the "others" category on the SOM were not included in the present analysis.

**Chart 1:** A stylized representation of a 20x20 field map (i.e. 400 nodes) trained with our data sample. One square corresponds to one node of the SOM.




Notes: This particular map has been obtained with the following parameter specifications: *rough tuning:* Training cycles 13,000, $\alpha(0) = 0.06$, training radius 11; *fine tuning:* training cycles 4,000, $\alpha(0) = 0.01$, training radius 3. However, the results were very stable with regard to changes of parameter settings.

**Chart 2:** The 20x20 map with superimposed monthly standard deviations (SD). One square corresponds to one node of the SOM.



Notes: white SD 0-2.5%, light grey SD 2.5-5%, dark grey SD 5-7.5%, black SD > 7.5%. The dotted lines indicate the borders of the empirically confirmed strategies discussed earlier.