The **HELLENIC OPEN BUSSINES ADMINISTRATION** *Journal*

Volume 1 - 2015, No 1 - ***Author Reprint***

Edited by: **Dimitrios A. Giannias**, *Professor*
HELLENIC OPEN UNIVERSITY

ISSN: 2407-9332

Athens 2015
Publisher: D. Giannias

The **HELLENIC OPEN**
BUSINESS ADMINISTRATION
Journal

Volume 1 - 2015, No 1

The HELLENIC OPEN BUSINESS ADMINISTRATION *Journal*
Publisher: D. Giannias / Athens 2015

ISSN: 2407-9332

www.hoba.gr

The HELLENIC OPEN BUSINESS ADMINISTRATION JOURNAL

AIMS AND SCOPE

The ***HELLENIC OPEN BUSINESS ADMINISTRATION Journal*** is published two times a year and focuses on applied and theoretical research in business Administration and economics.

The HELLENIC OPEN BUSINESS ADMINISTRATION JOURNAL

EDITOR'S NOTE

The ***HELLENIC OPEN BUSINESS ADMINISTRATION Journal*** is concerned with theory, research, and practice in business administration and economics (in its wider sense encompassing both private and public sector activities of profit-seeking ventures, as well as of governmental, private non-profit, and cooperative organisations) and provides a forum for academic debate on a variety of topics which are relevant to the journal's central concerns, such as:

- Administration of Businesses and Organizations
- Marketing
- Public Administration and Policy
- Accounting
- Financial Management
- Total Quality Management
- Law and Administration
- European Business
- Tourism Business Administration
- Cultural Organisations Management
- Health Care Management
- Environmental Management
- Industrial Organization
- Economic Analysis and Policy
- Money and Capital Markets
- Quantitative Methods

- Labour Economics

The ***HELLENIC OPEN BUSINESS ADMINISTRATION Journal*** also publishes special issues. A special issue focuses on a specific topic of wider interest and significance, which is announced through relevant call for papers.

The journal was established in 2014 following the completion of the HELLENIC OPEN BUSINESS ADMINISTRATION International Conference.

The ***HELLENIC OPEN BUSINESS ADMINISTRATION Journal*** (The HOBA Journal) is published two times a year, in January and July. These two issues constitute one volume. One or more issues may focus on a specific topic of wider interest and significance, which is announced through relevant call for papers.

The editorial process at The HOBA Journal is a cooperative enterprise. Articles received are distributed to the Editor for a decision with respect to publication. All articles are first reviewed to be judged suitable for this journal. The Editor arranges for refereeing and accepts and rejects papers or, alternatively, forwards the papers to a member of the Board of Editors. The member of the Board of Editors, then, arranges for refereeing and accepts or rejects papers in an entirely decentralized process. In any case, each submission is sent to two referees for blind peer review and the final decision is based on the recommendations of the referees. The referees are academic specialists in the article's field of coverage; members of the Board of Editors and/or members of the Editorial Advisory Board may act as referees in this process. Only when a paper is accepted for publication it is sent again to the Editor. Subsequently, the Editor sends the finally accepted paper to The HOBA Journal office for final editing and typesetting.

The Editor or the member of the Board of Editors who coordinates the decision with respect to publication of an article may send an article for refereeing to member(s) of the Editorial Advisory Board or cooperate with one or more of them to jointly assign referees who have some substantive knowledge of the topic and research in the relevant field and, finally, to jointly decide whether to accept or reject a paper.

The Editor, the members of the Editorial Board, and the members of the Editorial Advisory Board come from a breadth of fields designed to cover the largest substantive areas in economics and business administration from which we expect to receive submissions.

The above outlined co-editing process has major advantages. First, it is helpful in the assignment of referees and in the decision whether to publish a submission. Second, it avoids the apparent conflict of interest that results when an Editor handles a colleague's article. As a general rule the Editor and the members of the Board of Editors never assign papers written by authors at the same institution.

Finally, it provides an efficient way to handle about 200 submissions annually.

The editorial structure and process is reviewed annually.

While the Journal seeks to publish papers, which are academically robust, hence the rigorous review process (double blind peer review), it also seeks to publish papers that communicate effectively. It is interesting, well written and, therefore, readable papers that really contribute to the area of interest. Articles submitted should, therefore, keep technical jargon and statistical formulae within papers to a minimum and always aim to present material, however complex, simply and clearly.

As a forum, the Journal invites responses to articles that are published and is also willing to publish controversial articles to stimulate debate. To facilitate this, in addition to standard articles, the Journal also publishes "viewpoints" and "notes". These are short papers (up to 2,000 words), that explore, or comment on, an issue in a way which is useful, interesting, worthwhile, relevant and, ideally, provocative.

It will contain book reviews, and review essays designed to bring relevant literatures to the attention of a wider readership.

For libraries subscribed to the Journal, all printing or photocopying fees or any royalty payments for multiple internal library use are waived. Special arrangements exist for subscribers in low-income countries.

All articles must be submitted in WORD format to:
theHOBAjournal@gmail.com

Dimitrios A. Giannias, Editor
School of Social Sciences
Hellenic Open University

www.hoba.gr

Volume 1 - 2015, No 1

CONTENTS

AIMS AND SCOPE …………………………………………… 5

EDITOR'S NOTE ……………………………………………… 7

COMPLEXITY IN A MONOPOLY MARKET WITH NON LINEAR DEMAND AND COST FUNCTIONS ……..…… 13

IS HEDGE FUNDS DECLARED STRATEGY COMPATIBLE WITH WHAT ACTUALLY DATA SAY? ……………….. 23

JOB SATISFACTION AND DEMOFRAPHIC CHARACTERISTICS OF GREEK BANK EMPLOYEES .. 47

INVENTORY HOLDING AND INTERNATIONAL MIXED DUOPOLY …………………………………………………. 73

TOURIST DESTINATION MARKETING AND MEASUREMENT OF TOURIST SATISFACTION: THE CASE OF SANTORINI ISLAND …………………………………..91

Volume 1 - 2015, No 1

IS HEDGE FUNDS DECLARED STRATEGY COMPATIBLE WITH WHAT ACTUALLY DATA SAY?

Theodosios B. PALASKAS
Professor of Economic Analysis, Department of Economic and Regional Development, Panteion University of Athens and EAP, DEO34.

Chrysostomos E. STOFOROS
Assistant Prof. of Microeconomic Analysis, Department of Economic and Regional Development, Panteion University of Athens and EAP, MBA50.

Abstract

Constructing a diversified hedge funds portfolio is very much associated with the degree of complementarity/substitutability of alternative fund strategies/styles, which are self declared by fund managers. To this extend, an obvious and important question related to the safe pick by investors of a specific fund strategy is: 'if and to what extend the fund managers declared strategy and its investment style are compatible with what actually their performance states?'. If yes, then the investors' decision making is well informed, otherwise they run to risk of constructing a homogeneous rather than heterogeneous portfolio of funds. The purpose of this paper is, by utilizing clustering techniques -k-means and two-step -, to identify the compatibility between managers self-declaration of fund strategy/style and what the data actually reveal. The framework used is based on data driven classification, covering 2,853 hedge funds of all strategies and styles for the period 2000-2009 with monthly returns, standard deviations, skewness, kurtosis and % of positive months as inputs. The results suggest that both, k-means and two-step clustering, produce similar outcomes and that all funds strategies and styles are classified in eight clusters. Moreover, the results indicate that, other than convertible arbitrage, emerging markets and event driven fund strategies, which are mainly classified in one cluster, none of the other hedge fund strategies appear as a homogeneous category. The latter suggests that investors should be very careful when building up their portfolio of hedge funds because the self-declaration of fund managers' strategy could be misleading.

Keywords: Portfolio Analysis, Hedge Funds, Diversification, Cluster Analysis.

JEL Classification: C38, G11, G23

Introduction

Hedge funds have experienced a substantial growth over the last decade and have increased in popularity mainly because investors see them as an alternative vehicle which can provide investment opportunities in areas that is limited, due to the knowledge that is required, access. In addition, investors turned to hedge funds in order to alter their portfolios diversification and to increase the risk adjusted returns.

A key issue in the hedge fund industry is its heterogeneity in terms of investment strategies and styles, which combined with the absence of an industry standard for their classification (Das 2003) and the large number of hedge fund database providers, creates a blockage for investors to achieve their diversification targets. It should be noted that hedge fund database providers allow each hedge fund manager to self-declare their strategy with respect to the given classification schemes (Moerth 2007). Therefore, an important issue that arises due to the existence of various classifications for the hedge fund industry is the distance between what a specific classification 'promises' to investors and what it actually delivers. In other words, if an investor decides to invest in a Long Short Fund he has specific expectations from this decision regarding returns, risk, leverage exposure, skewness, kurtosis, etc. To this extend, an obvious and important question related to the safe pick by investors of a specific fund strategy is: 'if and to what extend the fund managers declared strategy and its investment style are compatible with what actually their performance states?'. If yes, then the investors' decision making is well informed, otherwise they run to risk of constructing a homogeneous rather than heterogeneous portfolio of funds.

The purpose of this paper is, by utilizing clustering techniques -k-means and two-step -, to identify the compatibility between managers self-declaration of fund strategy/style and what the data actually reveal. The framework used is based on data driven classification, covering 2,853 hedge funds for the period 2000-2009 (which includes periods of market turmoil, i.e. dot com bubble, twin

tours and the housing bubble) with monthly returns, standard deviations, skewness, kurtosis and percent of positive months as inputs.
The paper, apart from introduction and conclusions unfolds in three sections. The second section provides a brief literature review for hedge funds clustering. Section 3 explains the applied methodology, while Section 4 presents and discusses the empirical analysis.

2. Literature Review

A number of papers were dedicated over the last 14 years to the study of strategy and/or style classification of hedge funds.

Martin (2000) classified hedge funds via a medoid method using monthly returns as an input and concluded that there is significant heterogeneity in individual fund returns within clusters, so aggregate data are likely to be only weakly applicable to individual funds. Experimentation led to the conclusion that eight separate clusters generate the most useful results.

The same number of clusters, through a different clustering technique, was the outcome of the research of Brown and Goetzmann (2003) who studied the monthly hedge fund returns over the period 1989-2000. The authors suggested a generalised style classification model that is effectively a k-means cluster and modified it as a generalised least squares (GLS) procedure in order to take into account the time varying and fund specific residual return variance. The authors concluded that the return-base quantitative classification shows an agreement with the classification of the TASS database for a three-year period up to December 1999.

Miceli and Susinno (2003) used the Euclidean distance to classify hedge funds based on their returns in order to identify if funds are following the strategies they say they are following. The main finding of their research was that the less discretionary the hedge fund strategies are, the more similar their corresponding returns. An example of less discretionary hedge fund strategies according to the authors, are those that follow mathematical models implemented by software decisions.

Das (2003) utilized cluster analysis (k-means) to classify hedge funds. The classification was based on asset class, size of the hedge funds, incentive fee, risk level and liquidity. Distance measures and mean silhouette values for the six-cluster, seven cluster and nine-cluster classifications were compared. According to the author, the seven-cluster classification performed better both in terms of optimising the distance criterion and reducing misclassification. The new classification was then compared with the ZCM/Hedge fund database and the outcome was that the new classification did not kept intact any category of the ZCM database.

Maillet and Rousset (2003) classified hedge funds employing the Kohonen algorithm. The authors concluded that two separate groups of funds can be distinguished: two third of the data belongs to the first one, consisting of one class only, whilst one third belongs to the second one, consisting of nine other classes. The analysis was based on a relatively small data set of only 294 funds, which limits the significance of the results.

Das and Das (2005) presented a hedge fund classification technique using fuzzy neural networks. The classification was based on asset classes the hedge funds invest in, incentive fees, leverage, liquidity of the investment strategy and fund sizes. The study indicated that there are six possible hedge fund groups. As with the Das 2003 paper, the present classification has not kept intact any category of the existing self-classification.

Baghai-Wadji et al. (2005) utilized Self-Organizing Maps (SOM) to detect homogeneous groups of hedge funds based on similar return characteristics. The authors, identified nine hedge fund classes and they concluded that, in contrast to Brown and Goetzmann (2003), a number of declared hedge fund styles display no or very limited return similarities.

Gibson and Gyger (2007) examined the style classification and the style consistency of hedge funds using a hard clustering procedure as well as fuzzy cluster membership in order to estimate hedge funds' probabilistic exposure to various styles. As with previous researches, the authors concluded that certain managers do not follow their investment style consistently over time.

Moerth (2007) discussed the benefits of a quantitative k-means cluster analysis with respect to the construction of a portfolio of hedge funds using as an input the mean performance. Portfolio allocation methods were developed based on the results of the cluster analysis for 1,349 hedge funds grouped in four strategies -Tactical Trading, Equity Long/Short, Event Driven and Relative Value- with a minimum track record of 60 months in the period May 2000 to April 2005. The quantitative cluster classification was also compared with the qualitative self-reported classification of hedge fund managers. According to the results, Tactical Trading funds tend to form their own cluster while Equity Long/Short and Event Driven funds exhibit similar properties that are distinct from most Tactical Trading and Relative Value funds. Finally, Relative Value funds tend to be spread over several clusters.

Shawky and Marathe (2010) utilized two clustering techniques (k-means and Hierarchical Clustering) to provide an objective method for classification of hedge funds. The data used were from the CISDM database. As stated, a data driven classification framework that utilizes monthly returns as inputs, was shown to provide better comparisons among fund categories and could help investors in identifying common factors that can lead to better diversification strategies. According to their results, there were only three unique hedge fund styles, namely the Equity Hedge, Fund of Funds and Emerging Markets.

Although there are a number of papers dealing with the issue of hedge funds clustering, the present paper utilizes the two-step algorithm through which the determination of the number of clusters is based on statistical information criteria and is not ad hoc specified. In addition, the database is unique because it's a combination of available data from Bloomberg plus authors' data collection through interviews with a large number of hedge funds. Finally, the reference period includes the subprime crisis, which had a severe impact on funds performance, and in most cases, irrespective of the attained strategy (Palaskas et al., 2013).

3. Metodology

The most common approach used to define the degree of similarity among returns is the correlation coefficient (Miceli et al 2003). However, since the purpose of this paper is to identify the compatibility between managers' self-declaration of fund strategy/style and what the data actually reveal, the most suitable statistical technique that can be utilized is cluster analysis. This technique allows the determination and the visualization of a taxonomy implanted in hedge funds historical data (Miceli et al 2003) and is a useful instrument to identify homogeneous groups in a heterogeneous sample of funds. More specifically, cluster analysis aims at sorting different objects into groups in such a way that the degree of association between two objects is maximal if they belong to the same group and minimal otherwise (Shawky et al 2010).

Traditional clustering methods fall into two broad categories: relocation and hierarchical. Relocation clustering methods — such as k-means move records iteratively from one cluster to another, starting from an initial data point. In addition, the number of clusters must be specified in advance and it does not change during the iteration. Hierarchical clustering methods proceed by stages producing a sequence of partitions in which each one nests into the next partition in the sequence.

Another clustering method proposed by SPSS is the two-step clustering. Two-Step Cluster Analysis requires only one pass of data and it can produce solutions based on mixtures of continuous and categorical variables and for varying numbers of clusters. The clustering algorithm is based on a distance measure that gives the best results if all variables are independent, the continuous variables have a normal distribution, and the categorical variables have a multinomial distribution. Because cluster analysis does not involve hypothesis testing and calculation of observed significance levels, other than for descriptive follow-up, it's perfectly acceptable to cluster data that may not meet the assumptions for best performance. One of the advantages of the two-step clustering process is that the number of clusters is determined through the application of statistical information criteria and is not ad hoc specified. To this direction, to determine the number of clusters automatically, the method uses two stages. In the first stage the indicator BIC (Schwarz's Bayesian Information Criterion) or AIC (Akaike's Information Criterion) is calculated for each number of clusters from a specified range and then this indicator is used to find an initial estimation for the number of clusters (Schiopu 2010).

Regarding the cluster membership of the various units, the data points are allocated on the specifications of resolving atypical values and the options for measuring the distances. If the option of solving the atypical values is not used, the values are assigned to the nearest cluster, according to the method of distance measuring (Schiopu 2010). Otherwise, the values are treated differently as follows: For example, in the case of Euclidian method, a unit is assigned to the nearest cluster if the distance between them is smaller than a critical value (Schiopu 2010),

$$C = 2\sqrt{\frac{1}{JK^A}\sum_{j=1}^{j}\sum_{k=1}^{K^A}\hat{\sigma}^2_{jk}} \tag{1}$$

Otherwise, the item is declared as noise (outlier).

Silhouette values are calculated to capture the performance of two-step cluster analysis. The silhouette value for each observation is a measure of how similar that observation is to observations in its own cluster compared to observations in other clusters, and ranges from –1 to +1. It is defined as:

$$S_j = \frac{(\min b(j,k) - a(j))}{\max(a(j), \min(b(j,k)))} \tag{2}$$

where a(j) is the average distance from the observation j to the other observations in its cluster, and b(j,k) is the average distance from the observation j to observations in another cluster k. By using a variety of cluster members in the analysis, the average silhouette value can be used to determine the optimal number of clusters (Schiopu 2010).

4. Data and Empirical Estimation

The hedge fund data used in this study were obtained from the Bloomberg database and from authors' own research through in depth interviews of the manager's investment style and operations. The dataset covers 2,853 hedge funds from February 2000 through October 2009. The sample includes hedge funds domiciled in US, Europe, Asia, Latin America as well as offshore funds. The variables reported in the database are fund name, strategy and monthly returns. It is important to stress that the fund strategy is subjective and is based on each hedge fund own declaration. The 2000-2009 period is chosen for two reasons: First, the period is characterized by rapid growth of the hedge fund industry and second, the sample is long enough to cover more than one business cycle. In fact, the 2000-2009 period covers the dot-com bubble, the September 11th crisis as well as the 2008 subprime crisis.

The strategies included in the database are: Asset Backed, Convertible Arbitrage, Distressed, Emerging Markets, Event Driven (excluding distressed funds), Long-Short, Macro/CTA and Multistrategy.

The number of funds by strategy and their share to the total of the sample are reported in Table 1. Specifically, 1% of the hedge funds of the constructed database are self-reported as asset backed while 1/3 of them are reported as multistrategy. Convertible arbitrage, emerging markets, long-short and Macro/CTA funds have shares on total hedge funds in the database that range between 12,4% and 16,6%, while the 4,9% and 5,3% of the funds in the database belong to the distressed and event driven funds strategy respectively.

Table 1. Self-Declared Strategy Characteristics (raw data) (02/2000-10/2009)

Strategy	**Number of Funds**	**% of Total**
Asset Backed	27	1,0%
Convertible Arbitrage	374	13,3%
Distressed	136	4,9%
Emerging Markets	348	12,4%
Event Driven	149	5,3%
Long-Short	361	12,9%
Macro/CTA	465	16,6%
Multistrategy	943	33,6%
Total	**2803**	**100,0%**

Source: Bloomberg and Authors' Research.

The first step of the clustering analysis is to present summary statistics of the self reported clusters/strategies over the period February 2000 through October 2009. Mean returns, standard deviations, skewness, kurtosis and the percent of positive months for each of the strategies are estimated. It is noted that mean return provides heterogeneous results varying from –0,86 for the asset backed funds to 1,31 for the emerging market funds. The same holds for standard deviations that vary from 2,08 for event driven funds to 9,45 for emerging market funds. Skewness and kurtosis present also a large variation among the hedge fund strategies and it is likely that these third and fourth moments provide valuable information in the clustering process. Finally, the variable with the lowest variation is the percentage of positive months that varies from 56% for the Convertible Arbitrage funds to 76% for the event driven ones (Table 2).

The next step of the analysis is the implementation of the clustering procedure through the adoption of the k-means and two-step algorithms.

Table 2. Self-Declared Strategy Characteristics (raw data) (02/2000-10/2009)

	Asset Backed	**Convertible Arbitrage**	**Distressed**	**Emerging**	**Event Driven**	**Long-Short**	**Macro /CTA**	**Multistrategy**
Mean	-0,86	-0,20	0,05	1,31	1,01	0,45	0,37	0,00
St. Deviation	3,76	3,73	3,25	9,45	2,08	3,72	4,91	2,94
Skewness	-0,99	-1,42	-0,97	-0,36	-0,08	-0,47	-0,30	-1,17
Kurtosis	3,67	5,54	4,11	0,73	2,23	2,74	2,31	4,50
Positive Months	0,67	0,56	0,62	0,61	0,76	0,61	0,58	0,62

Source: Authors estimations.

K-means Clustering

In the k-means clustering procedure hedge funds were placed in eight clusters. More precisely, the results of the k-means analysis show that six out of eight clusters present a positive monthly mean performance (Table 3) while clusters 4 and 6 have a considerable negative performance of –1,33 and –5,24 respectively. Standard deviations vary from 1,98 (Cluster 8) to 19,05 (Cluster 6).

Regarding skewness and kurtosis, seven out of the eight clusters are negatively skewed while two Clusters (2 and 7) present a kurtosis over 10. Finally, the cluster with the lowest percent of positive months is Cluster 6 (39%) while Cluster 8 presents the highest percent of positive months (73%).

Table 3. K-Means Clustering: Cluster Characteristics

	1	2	3	4	5	6	7	8
Monthly Mean (%)	0,16	0,07	0,39	-1,33	1,94	-5,24	1,07	0,69
Monthly St. Deviation (%)	3,12	2,89	11,40	4,54	8,30	19,05	4,81	1,98
Skewness	-0,61	-2,48	-0,19	-0,95	-0,35	-1,25	2,08	-0,58
Kurtosis	1,89	10,75	0,69	2,01	0,80	3,86	11,36	2,23
% of Positive Months	0,58	0,66	0,53	0,42	0,64	0,39	0,65	0,73

Source: Authors estimations.

The classification of the self-documented styles/strategies into the clusters that are produced through the k-means clustering procedure show the following (Tables 4 and 5):

- Less than half (48,1%) of the funds that their managers declare them as asset backed are classified in Cluster 8 while 29,6% in Cluster 4. The remaining asset backed funds are classified in Clusters 1, 2, 6 and 7. It is interesting to note that Cluster 8, where almost half of the asset backed funds are placed, has a 0,69 mean monthly performance and 1,98 standard deviation while the self declared asset backed strategy presents a –0,86 mean performance and 3,76 standard deviation (Tables 2 and 3).
- Almost 82% of the convertible funds are classified into two Clusters (1 and 2) while from the remaining 18%, 13,6% belong to Cluster 4. Both 1 and 2 clusters have a positive mean monthly performance while the self-declared convertible arbitrage strategy presents negative performance (Tables 2 and 3).
- Distressed funds show a large heterogeneity in their classification. More precisely, 27,9% of the distressed funds are classified in Cluster 8, 25% in Cluster 1, 18,4% in Cluster 2, 16,2% in Cluster 4 with the remaining 12,5% of the funds to be distributed in all other 4 Clusters. The large variation of the distressed fund classification or in other words, the absence of a representative cluster, does not permit the comparison with the self-declared Distressed strategy performance profile.
- More than 87% of the Emerging market funds are classified into two Clusters, 61,2% to Cluster 5 and 25,9% to Cluster 3. The remaining funds of this category are spread across all other 6 Clusters. The return profile of Cluster 5, where the majority of the Emerging market funds is classified, is consistent with the actual data from the self-declared Emerging market strategy (Tables 2 and 3).
- The vast majority of the event driven funds (78,5%) are classified in Cluster 8. The remaining 21,5% event driven funds are classified in Clusters 1, 2, 5 and 6. Cluster 8 presents similar return profile with the self-declared event driven strategy (Tables 2 and 3).
- 45,2% of the long short funds are classified in Cluster 1 and 25,8% in Cluster 8 while the remaining, 29,1%, funds are almost evenly distributed in Clusters 2, 3, 4 and 7. The return profile of Clusters 1 and 8 does not present variations from what is actually observed from the self-declared Long/Short strategy (Tables 2 and 3).
- Less than half (44,1%) of the Macro/CTA funds are classified in Cluster 1. From the remaining funds, 14,6% are classified in Cluster 8, 11,8% in

Cluster 5, 7,7% in Cluster l, 7,3% in Cluster 4 and 6,9% in Cluster 2. If Cluster 1 is considered to represent Macro/CTA funds then their return profile is not consistent with the self-declaration (Tables 2 and 3) of the Macro/CTA managers.

- Multistrategy funds are mainly classified in three Clusters: Cluster 1 with 37,4% of the funds, Cluster 8 with 26,6% and Cluster 2 with 21,4% of the funds. The diversification of the multistrategy funds in the eight Clusters does not permit to draw any conclusions regarding their consistency with Multistrategy managers self-declaration.
- In total, more than 1/3 of the Funds are classified in Cluster 1, over 21% are classified in Cluster 8 and 15,4% in Cluster 2. The smallest Cluster is 6 with only 36 funds out of 2803 or 1,3% of the total funds. Therefore, the vast majority (70,7%) of hedge funds are classified in three Clusters irrespective of the strategy/style that initially belonged to, according to the declaration of their hedge fund managers.

Table 4. K-Means Clustering: Number of Funds in Each Cluster

	1	2	3	4	5	6	7	8	Total
Asset Backed	1	2	0	8	0	2	1	13	**27**
Convertible Arbitrage	168	137	4	51	0	2	3	9	**374**
Distressed	34	25	5	22	5	1	6	38	**136**
Emerging	23	1	90	3	213	6	7	5	**348**
Event Driven	10	9	1	0	7	0	5	117	**149**
Long-Short	163	23	14	22	28	0	18	93	**361**
Macro/CTA	205	32	36	34	55	16	19	68	**465**
Multistrategy	353	202	3	97	8	9	20	251	**943**
Total	**957**	**431**	**153**	**237**	**316**	**36**	**79**	**594**	**2803**

Source: Authors estimations.

Table 5. K-Means Clustering: Hedge Funds Strategy Taxonomy by Cluster

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Total**
Asset Backed	3,7%	7,4%	0,0%	29,6%	0,0%	7,4%	3,7%	48,1%	**100,0%**
Convertible Arbitrage	44,9%	36,6%	1,1%	13,6%	0,0%	0,5%	0,8%	2,4%	**100,0%**
Distressed	25,0%	18,4%	3,7%	16,2%	3,7%	0,7%	4,4%	27,9%	**100,0%**
Emerging	6,6%	0,3%	25,9%	0,9%	61,2%	1,7%	2,0%	1,4%	**100,0%**
Event Driven	6,7%	6,0%	0,7%	0,0%	4,7%	0,0%	3,4%	78,5%	**100,0%**
Long-Short	45,2%	6,4%	3,9%	6,1%	7,8%	0,0%	5,0%	25,8%	**100,0%**
Macro/CTA	44,1%	6,9%	7,7%	7,3%	11,8%	3,4%	4,1%	14,6%	**100,0%**
Multistrategy	37,4%	21,4%	0,3%	10,3%	0,8%	1,0%	2,1%	26,6%	**100,0%**
TOTAL	**34,1%**	**15,4%**	**5,5%**	**8,5%**	**11,3%**	**1,3%**	**2,8%**	**21,2%**	**100,0%**

Source: Authors estimations.

An important outcome of the k-means clustering analysis is that most of the hedge fund strategies/styles, with the exception of Emerging Markets and Event driven funds where almost 2/3 of the funds belong to one cluster, present more or less heterogeneous behaviour since more than half of the funds are spread in more than one clusters.

Two Step Clustering

In the two-step clustering procedure the algorithm produced, similarly to k-means, eight clusters where all observations were placed. The results show that five out of the eight clusters present a positive monthly mean performance (Table 6) while the rest, clusters 1, 6 and 7, have a negative performance of –4,10, -0,01 and –0,89 respectively. Standard deviations vary from 1,87 (Cluster 5) to 13,93 (Cluster 1). Regarding skewness and kurtosis, seven out of the eight clusters are negatively skewed and two Clusters (3 and 8) present a kurtosis over 10.

Table 6. Two Step Clustering: Cluster Characteristics

	1	2	3	4	5	6	7	8
Mean	-4,10	1,55	1,01	0,36	0,78	-0,01	-0,89	0,07
St. Deviation	13,93	9,54	5,64	2,97	1,87	3,14	4,11	2,75
Skewness	-1,45	-0,35	2,11	-0,37	-0,47	-1,91	-0,67	-3,55
Kurtosis	3,92	0,70	11,98	1,19	2,60	6,81	1,28	18,36
% of Positive Months	0,39	0,61	0,62	0,59	0,76	0,62	0,44	0,70

Source: Authors estimations.

The classification of the hedge funds into the clusters that are produced through the two-step clustering procedure show the following (Tables 7 and 8):

- Less than half (48,1%) of the funds that their managers declare them as asset backed are classified in Cluster 5 while 18,5% in Cluster 7 and 14,8% in Cluster 1. The remaining asset backed funds are classified in Clusters 3, 4 and 6. It is interesting to note that cluster 5, where almost half of the asset backed funds are placed, has a 0,78 mean monthly performance and 1,87 standard deviation (results similar to the k-mean analysis) while the self declared asset backed strategy presents a –0,86 mean performance and 3,76 standard deviation (Tables 2 and 6).
- Almost 80% of the convertible funds are classified into two Clusters (4 and 6) while 15,8% are classified in Cluster 7. Cluster 6 where 58,3% of the convertible funds are placed has a negative mean monthly performance and 3,14 standard deviation which is more or less consistent with the self-declared strategy results (Tables 2 and 6).
- Distressed funds present, as in the case of k-means clustering, a large heterogeneity in their classification. More precisely, 22,8% of the distressed funds are classified in Cluster 6, 21,3% in Cluster 4, 21,3% in Cluster 5, 18,4% in Cluster 7 while the remaining 16,2% of the funds are distributed in all other four Clusters. As in the case of the k-means clustering, the variation of the distressed fund classification does not permit the comparison with the self-declared Distressed strategy performance profile.
- More than 83% of the Emerging market funds are classified in Cluster 2. The remaining funds of this category are classified in almost all other six Clusters (with the exception of cluster 6). The return profile of Cluster 2, where the majority of the Emerging market funds is classified, is consistent, as in the case of k-means analysis, with the data from the self-declared Emerging Markets strategy (Tables 2 and 6).
- The vast majority of the event driven funds (73,8%) are classified in Cluster 5. The remaining 26,2% event driven funds are classified in Clusters 2, 3, 4 and 6. Cluster 5 presents, as in the k-means case, similar return characteristics with the self-declared Event Driven strategy (Tables 2 and 6).
- 65,4% of the long short funds are classified in Clusters 4 and 5 (48,8% in Cluster 4 and 16,6% in Cluster 5) while from the remaining, 34,6% of the funds, 28,5% are almost evenly distributed in Clusters 2, 6 and 7. The return profile of Clusters 4 and 5 does not present great variations from

what is actually observed from the self-declared Long/Short strategy (Tables 2 and 6), a result that is similar to the k-mean outcome.

- Less than half (46,9%) of the Macro/CTA funds are classified in Cluster 4. From the remaining funds, 14,8% are classified in Cluster 2, 10,5% in Cluster 6, 9,7% in Cluster 5, 9,0% in Cluster 7, 4,3% in Cluster 1, 3,9% in Cluster 3 and 0,9% in Cluster 8. If Cluster 4 is considered to represent Macro/CTA funds then their return profile is consistent, contrary to the k-means results, with the return profile of Macro/CTA managers' self-declaration (Tables 2 and 6).
- Multistrategy funds are mainly classified in three Clusters: Cluster 4 with 35,5% of the funds, Cluster 6 with 25,0% and Cluster 5 with 17,2% of the funds. The diversification of the multistrategy funds in the eight Clusters does not permit to draw any conclusions regarding the consistency of the clustering return profile with the Multistrategy managers' self-declared strategy.
- In total, almost 1/3 of the Funds are classified in Cluster 4, over 20% are classified in Cluster 6, 15,2% in Cluster 5 and 14,8% in Cluster 2. The smallest Cluster is 1 with only 70 funds out of 2853 or 2,5% of the total funds.

Table 7. Two Stage Clusters: Number of Funds in Each Cluster

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Total**
Asset Backed	4	0	1	1	13	3	5	0	**27**
Convertible Arbitrage	3	3	4	75	2	218	59	10	**374**
Distressed	3	7	1	29	29	31	25	11	**136**
Emerging	8	291	7	27	4	0	10	1	**348**
Event Driven	0	6	4	22	110	7	0	0	**149**
Long-Short	0	33	19	176	60	35	35	3	**361**
Macro/CTA	20	69	18	218	45	49	42	4	**465**
Multistrategy	32	6	18	335	162	236	86	68	**943**
Total	**70**	**415**	**72**	**883**	**425**	**579**	**262**	**97**	**2803**

Source: Authors estimations.

Table 8. Two Stage Clustering: Hedge Funds Strategy Taxonomy by Cluster

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Total**
Asset Backed	14,8%	0,0%	3,7%	3,7%	48,1%	11,1%	18,5%	0,0%	**100,0%**
Convertible Arbitrage	0,8%	0,8%	1,1%	20,1%	0,5%	58,3%	15,8%	2,7%	**100,0%**
Distressed	2,2%	5,1%	0,7%	21,3%	21,3%	22,8%	18,4%	8,1%	**100,0%**
Emerging	2,3%	83,6%	2,0%	7,8%	1,1%	0,0%	2,9%	0,3%	**100,0%**
Event Driven	0,0%	4,0%	2,7%	14,8%	73,8%	4,7%	0,0%	0,0%	**100,0%**
Long-Short	0,0%	9,1%	5,3%	48,8%	16,6%	9,7%	9,7%	0,8%	**100,0%**
Macro/CTA	4,3%	14,8%	3,9%	46,9%	9,7%	10,5%	9,0%	0,9%	**100,0%**
Multistrategy	3,4%	0,6%	1,9%	35,5%	17,2%	25,0%	9,1%	7,2%	**100,0%**
Total	**2,5%**	**14,8%**	**2,6%**	**31,5%**	**15,2%**	**20,7%**	**9,3%**	**3,5%**	**100,0%**

Overall, the outcome of the two-step clustering is similar with the results from the k-means analysis although in the former, the number of clusters was the result of statistical inference and not ad hoc specification as in the latter. More precisely, only in three hedge fund strategies, Convertible Arbitrage, Emerging Markets and Event Driven, the funds where mainly classified in one cluster, in all other strategies hedge funds were distributed in more clusters.

Finally, it is important to stress that if an investor aims to construct a diversified hedge funds portfolio and picks a fund irrespective of the managers self-declared strategy, there is more than 30% probability that he will end up with the same result in terms of risk adjusted returns. The latter, emphasizes the fact that investors run to risk of constructing homogeneous rather than heterogeneous portfolio of funds.

5. Conclusions

The purpose of the present paper was to examine the distance between what a specific strategy/style classification 'promises' to investors and what it actually delivers. In other words, the aim of the paper was to identify homogeneous groups in a heterogeneous sample of hedge funds. To this respect, the statistical technique that was considered the most appropriate was cluster analysis, k-means and two-step, because its target is to sort different objects into groups in such a way that the degree of association between two objects is maximal if they belong to the same group and minimal otherwise. Although there are a number of papers dealing with the issue of hedge funds clustering, the present paper utilizes the two-step algorithm through which the determination of the number of clusters was based on statistical information criteria and was not ad hoc specified.

The data driven classification implemented in this paper covered 2,853 hedge funds for the period 2000-2009 with monthly returns, standard deviations, skewness, kurtosis and % of positive months as inputs. It is worth mentioning that the database is unique because it's a combination of available data from Bloomberg plus authors' data collection through interviews with a large number of hedge funds.

According to the results both k-means and two-step clustering produced similar outcomes and all funds were classified in eight clusters. Moreover, the results indicated that other than convertible arbitrage, emerging markets and event driven funds which are mainly classified in specific separate clusters, all other strategies present more or less increased heterogeneity in their classification. The later proves that investors should be very careful when designing the structure of their portfolio of hedge funds because the self-declaration of fund managers can be misleading and thus the aim of diversification might not be achieved.

References

Baghai-Wadji, R., El-Berry, R., Klocker, S., and M. Schwaiger, (2005), The Consistency of Self-Declared Hedge Fund Styles – A Return-Based Analysis with Self-Organizing Maps, *Financial Stability Report*, 9, Österreichische Nationalbank, 64-76.

Bianchia, R., Drewa, M., Veeraraghavanb, M., and P. Whelana, (2005), An Analysis of Hedge Fund Styles using the Gap Statistic, *Working Paper, School of Economics and Finance*, Queensland University of Technology.

Brown, S. J. and W. N. Goetzmann, (2003), Hedge Funds with Style, *Journal of Portfolio Management*, 29 (2): 101-112.

Das N. (2003), Hedge Fund Classification Using K-Means Clustering Method, *9th International Conference on Computing in Economics and Finance*, July 11-13 2003, University of Washington, Seattle.

Das, N., and R. Das, (2005), Hedge Fund Classification Technique using Self-Organizing Feature Map Neural Network, *Annual Meetings of the Financial Management Association International*, Chicago.

Gibson R. and Gyger S. (2007), The Style Consistency of Hedge Funds, *European Financial Management*, 13 (2): 287-308.

Maillet B., and P. Rousset, (2003), Classifying Hedge Funds with Kohonen Maps, Connectionist Approaches in Economics and Management Sciences, *Series in Advances in Computational Science,* 6: 233-259.

Martin G., (2000), Making Sense of Hedge Fund Returns: What Matters and What Doesn't, *Derivatives Strategy*.

Miceli M. A. and Susinno G. (2003), Using Trees to Grow Money, www.risk.net, November 2003 RISK.

Moerth P. (2007), Hedge Funds: Performance Analysis, Strategy Classification, and Portfolio Construction, *PhD Thesis*, University of St. Gallen, Graduate School of Business Administration, Economics, Law and Social Sciences, Dissertation N. 3355.

Palaskas T., Stoforos C. and Drakatos C. (2013), Hedge Funds Development and their Role in Economic Crises, Scientific Annals of the"Alexandru Ioan Cuza" University of Iasi (SAAIC)*, 60 (1): 173-186.*

Schiopu D. (2010), Applying TwoStep Cluster Analysis for Identifying Bank Customers' Profile, *Buletinul, Universitatii Petrol, Gaze din Ploiesti*, LXII (3): 66-75.

Shawky H. A. and Marathe A. (2010), Stylistic Differences Across Hedge Funds as Revealed by Historical Monthly Returns, *Technology and Investment*, 2: 26-34.