D S E

Dipartimento   **Scienze**   **Economiche**

Working Paper

Department of Economics

Ca' Foscari University of Venice

Monica Billio

Mila Getmansky

Loriana Pelizzon

Non-Parametric Analysis of Hedge Fund Returns: New Insight from High Frequency Date

Non-Parametric Analysis of Hedge Fund Returns: New Insights from High Frequency Data

Monica Billio
University of Venice and SSAV, Scuola Studi Avanzati in Venezia

Mila Getmansky
Isenberg School of Management, University of Massachusetts

Loriana Pelizzon
University of Venice and SSAV, Scuola Studi Avanzati in Venezia

Abstract

This paper examines four different daily datasets of hedge fund return indexes: MSCI, FTSE, Dow Jones and HFRX, all based on investable hedge funds, and three different monthly datasets of hedge fund return indexes: CSFB, CISDM and HFR which comprise both investable and non-investable hedge funds. Our study, based on standard statistical analysis, non-parametric analysis of the distribution and non-parametric regressions with respect to the S&P500 index shows that key data biases and disparate index construction methodologies lead to different statistical properties of hedge fund databases. One key variable that highly affects the statistical properties of hedge fund index returns is the "investability" of hedge funds.

Address for correspondence:
Loriana Pelizzon
Department of Economics
Ca' Foscari University of Venice
Cannaregio 873, Fondamenta S.Giobbe
30121 Venezia - Italy
Phone: (++39) 041 2349164
Fax: (++39) 041 2349176
e-mail: loriana.pelizzon@unive.it

1. Introduction

In the last two decades we have observed a rapid acceleration in hedge fund investments, which has been particularly important in the 1990s. This has lead to desire for greater information on hedge funds and their historical performance, and has created the emergence of a number of Data Vendors that has satisfied the demand of hedge funds databases.

The main reason for the creation of these datasets is the need by hedge fund investors to track fund performance, and this has resulted in the creation of both aggregate indexes as well as sub-indexes reflecting various investment strategies. Most of current available hedge fund databases provide monthly returns. However, various data vendors follow very different methodologies in structuring hedge fund indexes. For example, they use different number of hedge funds, include or exclude funds based on various fund restrictions, fund size, the years in operation, liquidity provisions, and other parameters. Also, the aggregation techniques vary from equal-weighted, median-weighted to asset-weighted.

Recently, we observed the creation of datasets based on daily observations and this aspect has inspired this paper. Our goal is to examine four daily hedge fund index databases in order to highlight their statistical characteristics and reliability, and compare them to monthly datasets based both on investable and non-investable hedge funds.

This analysis is relevant since by using monthly data, most of hedge fund studies have shown that sample distributions of hedge fund returns and indexes constructed from them exhibit characteristics not typically displayed by traditional equity and fixed income investments. For instance, Brooks and Kat (2001) examine hedge fund indexes from multiple databases and find that for some investment strategies the return

distributions are highly skewed and kurtotic. Getmansky, Lo and Makarov (2004) show that the return series display strong first-order serial correlation. Fung and Hsieh (1997), Agarwal and Naik (2004), and Billio, Getmansky and Pelizzon (2007) show that risk factor exposures are highly non-linear and depend on the volatility of market risk factors.

However, Boyson, Stahel and Stulz (2006), using daily returns to examine whether hedge funds experience contagion, show that the daily data returns are more similar to traditional equity returns. Li and Kazemi (2006) investigate the same issue using a semi-non-parametric GARCH(1,1) model and thus consider the conditional distribution and correlations. They work with the Dow Jones hedge fund indexes based on daily observations and find similar results. Our paper is linked to this branch of the literature and extends the analysis to four different datasets: the MSCI Hedge Invest Indexes, the FTSE Hedge Fund Index family, the Hedge Fund Research Indexes (HFRX daily indexes), and the Dow Jones Hedge Fund Strategy Benchmarks. Our study is based on standard statistical descriptive analysis, non-parametric analysis of the distribution and non-parametric regressions with respect to S&P500 index.

We investigate differences among daily datasets that originate not only from different sample periods used, but mostly from the differences in database constructions. We compare these characteristics with three monthly databases based on non-investable indexes and one monthly database based on investable indexes and we show that "investability" is another important characteristic that affects hedge funds returns distributions.

The paper is organized as follows. Section 2 describes different datasets. Section 3 provides statistical descriptions of the daily and monthly datasets. Section 4 presents the results of the non-parametric analysis. Section 5 concludes.

2. Data Description

2.1. Daily Data

The daily hedge fund indexes include the MSCI Hedge Invest Indexes, the FTSE Hedge Fund Index family, the Hedge Fund Research Indexes (HFRX daily indexes), and the Dow Jones Hedge Fund Strategy Benchmarks. S&P 500 is the daily market index. All daily hedge fund indexes and S&P 500 are downloaded from the DataStream database. Dow Jones indexes are downloaded from the Dow Jones website. HFRX data is from April 1, 2003 through June 29, 2007. MSCI data is from July 15, 2003 through June 29, 2007. FTSE is available from July 1, 2004 through June 29, 2007. Down Jones data is available from January 2, 2004 through June 29, 2007.

We use both the composite-level MSCI Hedge Invest Index and seven strategy-level indexes: Equity Non-Directional, Convertible Arbitrage, Event-Driven, Systematic Trading, Discretionary Trading, Long Bias and Variable Bias.

The MSCI Hedge Invest Indexes are asset-weighted, and designed to be investable. Each investable hedge fund index contains funds with weekly liquidity and is designed to be licensed for use as a basis for tradable investment products. MSCI includes funds that have identified themselves as open or closed to new investment. It includes funds regardless of domicile location. MSCI does not require a minimum track record for database inclusion. If a fund has only one strategy, the designation of the strategy is used for the final classification of the fund. If a fund has more than one strategy, the fund will be classified within the strategy where at least 80% of the fund capital is being risked.

The FTSE Hedge Fund Indexes we use include a headline index, management style indexes (directional, event-driven and non-directional) and trading strategy

indexes (equity hedge, merger arbitrage, distressed and opportunities, convertible arbitrage, and equity arbitrage).

Trading strategy indexes are weighted by investability of all eligible hedge funds for that strategy. Management style indexes are weighted by the sum of their respective Trading Strategy indexes. No single constituent can represent more than 40% of its Trading Strategy Index. No single Trading Strategy Index can represent more than 30% of the FTSE Hedge Index. Constituent fund weightings are reviewed monthly. The FTSE Hedge Fund Indexes are designed to be investable.

The FTSE Hedge Index consists of a fixed number of 40 hedge funds. Each strategy index consists of a minimum of 3 hedge funds. Constituent funds must meet tightly defined quality, liquidity and capacity requirements, i.e. they need to have: (1) independent audited financial statements; (2) at least USD$50m of unleveraged AUM; (3) a minimum two year track record at the time of annual review; (4) a monthly reporting structure with quarterly liquidity screening; and (5) significant remaining investment capacity and be open to investor subscriptions.

The HFRX Indexes comprise nine single strategy indexes and an HFRX Equal Weighted Strategies Index. The HFRX strategy indexes include Convertible Arbitrage, Distressed Securities, Equity Hedge, Equity Market Neutral, Event Driven, Merger Arbitrage, Relative Value Arbitrage, Absolute Return, and Market Directional. The index weighting of HFRX Strategy Indexes are based on Representative Optimization. The HRFX indexes are investable.

The HFRX Equal Weighted Strategies Index is designed to be representative of the overall composition of the hedge fund universe. It is comprised of eight strategies: convertible arbitrage, distressed securities, equity hedge, equity market neutral, event

driven, macro, merger arbitrage, and relative value arbitrage. The strategies are equal-weighted.

HFR believes in a dynamic, bottom-up approach to index construction based on hard ???? numbers and style purity:

- HFR screens approximately 3,000 hedge funds to identify those firms with multi-year records, at least $50 million AUM, a minimum two-year track record, willing to trade on a transparent basis and are open to new assets.

- Cluster and correlation analyses are performed to group managers by true strategy categories and to eliminate outliers.

- Monte Carlo Simulation helps determine the adequate number and types of managers to replicate each strategy.

- Selected managers must provide daily transparency and pass extensive qualitative screening.

- Manager investments are then weighted to maximize correlation with their group.

Dow Jones Hedge Fund Strategy Benchmarks cover six style classifications, including Convertible Arbitrage, Distressed Securities, Event Driven, Equity Long/Short, Equity Market Neutral, and Merger Arbitrage. These strategy Benchmarks are investable. Each Benchmark is substantially equally weighted among its component managers, thereby minimizing the chances of one manager's performance skewing the performance of the Benchmark.

The Dow Jones Equal Weighted Hedge Fund Balanced Portfolio Index reflects the performance of a portfolio that is allocated approximately equally among the six strategies (target range: 14.67%-18.67%) on an asset management platform that seeks to

track the Dow Jones Hedge Fund Strategy Benchmarks. The Dow Jones Equal Weighted Hedge Fund Balanced Portfolio Index is investable.

The index universe is screened to arrive at a grouping of approximately six to eight managers for each style classification. In order to be selected for inclusion in a Benchmark, each prospective manager must:

- Have a minimum track record of two years.

- Have minimum assets under management of USD 50 mil.

- Undergo confirmatory cluster and other quantitative analyses.

- Undergo correlation analysis with related hedge fund indexes (e.g., EACM, CSFB, HFR).

- Undergo an on-site due diligence review and qualitative analysis.

- Be willing to run a managed account and to adhere to leverage limitations.

- Undergo Benchmark Advisory committee review.

- Gain final approval of Dow Jones Hedge Fund Indexes, Inc. ("DJHFI").

The S&P 500 index is weighted by market capitalization, and it is investable.

2.2. Monthly Data

For the monthly analysis, we use three vendors: the Hedge Fund Research Indexes (HFRX indexes), the CASAM CISDM Database (formerly the MAR Database), and Credit Suisse First Boston/Tremont (CSFB) indexes.[4]

The HFRX index construction and methodology have been described above. The data available is from April, 2003 through June, 2007.

The CISDM Equal Weighted Hedge Fund Index reflects the average performance of hedge fund managers reporting to the CISDM Hedge Fund/CTA

---

[4] For the sake of space, we refer to the Credit Suisse First Boston/Tremont indexes as Credit Suisse/Tremont or CSFB indexes.

Database. Its objective is to provide an estimate of the rate of return to an equally weighted portfolio of hedge fund managers who trade a wide variety of hedge fund strategies which are based on many different trading models. The index as well as all strategy indexes except Convertible Arbitrage Index start in January 1990. Convertible Arbitrage Index starts in January 1992. The strategy indexes are weighted by median performance.

The Hedge Fund Index has the following inclusion rules:

- Only funds that have reported monthly returns at the time the medians are calculated are included in the index.

- There is no minimum asset size required for inclusion in the hedge fund indexes.

- There is no length of time required that a fund must be reporting to the database to be included in the indexes.

- Once calculated, index performance numbers are never revised.

- Both offshore and onshore funds are included in the medians.

The Credit Suisse/Tremont Hedge Fund Index is the industry's first asset-weighted hedge fund index. The methodology utilized in the Credit Suisse/Tremont Hedge Fund Index starts by defining the universe it is measuring. The Index Universe is defined as funds with:

- A minimum of US $50 million assets under management ("AUM"),

- A minimum one-year track record, and

- Current audited financial statements.

Funds are separated into ten primary subcategories based on their investment style. The Index in all cases represents at least 85% of the AUM in each respective category of the Index Universe. Credit Suisse/Tremont analyzes the percentage of assets invested in each subcategory and selects funds for the Index based on those percentages,

matching the shape of the Index to the shape of the universe. The Index is calculated and rebalanced monthly. Funds are reselected on a quarterly basis as necessary. The Index uses a rules-based construction methodology, identifies its constituent funds, and minimizes subjectivity in the Index member selection process. It aims at a maximum representation of the Index Universe. To minimize survivorship bias, funds are not removed from the Index until they are fully liquidated or fail to meet the financial reporting requirements.

The Credit Suisse/Tremont Investable Hedge Fund Index is designed to provide a transparent, representative and rules-based benchmark of the investable index universe. A hedge fund is eligible for membership in the Investable Index if the fund meets all of the following criteria:

- It is an existing member of the Original Index or the Proxy for an existing member of the Original Index.

- It is accepting new investments and redemptions.

- The minimum amount for initial investment is less than or equal to the greater of (i) the product of USD 50,000,000 and its prospective weight in the Index; and (ii) USD 100,000.

- If it imposes a minimum amount for subsequent investments, that minimum is less than or equal to the lesser of (i) the product of USD 10,000,000 and its prospective weight in the Index; and (ii) USD 200,000.

- It is not a US domiciled hedge fund.

- It has no investment lock-up period.

- It allows investments no less frequently than monthly.

- It allows redemptions no less frequently than monthly or, in the case of funds in the Convertible Arbitrage, Event Driven and Multi Strategy Sectors, no less frequently than quarterly.

- It requires notification of redemptions of one month or less or, in the case of funds in the Convertible Arbitrage, Event Driven and Multi Strategy Sectors, three months or less.

- It meets the reporting criteria of the Original Index.

- Neither it, its investment management company or any affiliate are, to the knowledge of the Calculation Agent, under investigation or review by a regulatory body or other authority for reasons of wrongdoing, breach of any law, regulation or rule, or any similar reason, which is deemed likely to be materially adverse to the fund by the Calculation Agent.

- The 60 Member Funds are determined in order of size as the largest six Eligible Funds by AUM in each of the ten Sectors comprising the Original Index, but on the basis that once an Eligible Fund is included in a particular Sector, another Eligible Fund managed by the same or an affiliated investment manager will not be included in the same Sector. If any Sectors do not have six Eligible Funds, only the Eligible Funds in those Sectors are included. In that case, the remaining Member Funds will be chosen in order of size as the largest remaining Eligible Funds by AUM in the other Sectors until a total of 60 Member Funds have been determined, but on the basis that once a Member Fund is chosen from a particular Sector, a further Member Fund will not be chosen from that Sector until all other Sectors with remaining Eligible Funds have contributed to the selection process.

The Credit Suisse/Tremont Sector Invest Indexes are designed to provide transparent, representative and objective benchmarks of the ten style-based investment strategies of the hedge fund universe. It has similar criteria to the CSFB Investable Hedge Fund index. The additional requirements are:

- It has an AUM of at least USD 50,000,000.

- For each Sector Invest Index, the Member Funds are determined as all Eligible Funds whose AUM in aggregate is equal to 70% of all AUM in the Investable Universe for that Sector, with funds selected in order from largest to smallest AUM and subject to a maximum of 25 funds in each Sector Invest Index, but on the basis that once an Eligible Fund is included in a particular Sector, another Eligible Fund managed by the same or an affiliated investment manager will not be included in the same Sector Invest Index.

The data for the CSFB Tremont Hedge Fund Index starts in January, 1994. The data for CSFB Tremont Investable Index and Sector Investable Indexes starts in January, 2000. Detailed strategy descriptions are found in the Appendix. In listing strategies, we concentrate on equity-based strategies (because we use S&P500 as the benchmark) and eliminate non-equity based strategies like Fixed Income, Global Macro, Managed Futures, and CTAs.

3. Statistical Analysis

In this section, we analyze statistical properties of hedge funds returns. We examine risk and return relationship and statistical characteristics of return distributions. We focus on whether different datasets show differences in terms of normality of the distributions and autocorrelations. Section 3.1 reports the summary of the different daily datasets, and Section 3.2 reports summary statistics for the monthly datasets based on non-investable indexes, the indexes that are mostly used in previous empirical

studies. Section 3.3 focuses on the main differences in terms of investable and non-investable hedge funds indexes at the monthly frequency. Section 3.4 presents summary statistics for monthly returns constructed from the daily datasets and compares daily and monthly characteristics. Since daily data considers only investable indexes, in this section we concentrate our analysis only on monthly investable indexes.

3.1. Daily Data

The summary statistics for four daily datasets used (MSCI, FTSE, Dow Jones (DJ), and HFRX) are presented in Table 1. Among the four datasets we observe that the mean returns and standard deviations are quite similar across different indexes. However, it is well known that means and standard deviations are not the most relevant indicators for hedge fund returns because hedge fund returns are usually not normally distributed. This feature is also confirmed at the daily level for all different datasets and strategies. The most interesting aspect is the deviation from normality and this is highlighted through the skewness and excess kurtosis measures. As Table 1 shows, the skewness largely differs from zero, it is always negative and the lowest is presented by the DJ index and the highest belongs to FTSE index. The same applies to excess kurtosis: it is largely different from zero, it is positive, that is the distribution has fat tails. The lowest excess kurtosis belongs to the DJ index and the largest to the FTSE index. This peculiarity cannot be related to the different sample periods because DJ index is related to a shorter sample period and MSCI and HFRX are related to a longer sample period and all of them present a lower skewness and excess kurtosis with respect to the FTSE. This peculiarity therefore has to be related to the dataset construction. Section 2.1 clearly shows that the construction of the FTSE indexes is the most restrictive, sometimes having only 3 hedge funds per strategy index compared to other more

inclusive indexes. The total FTSE hedge fund index has the constant number of only 40 funds.

If we investigate single strategies we have that the sample characteristics in terms of skewness and excess kurtosis are different among the different datasets. We prefer to avoid the aggregation among the different strategies in three main categories as in Boyson et al. (2007) paper because this aggregation may hide peculiarity of some strategies that are not included in the other datasets. For this reason, we concentrate on two strategies that are common to all four datasets: Convertible Arbitrage (CA) and Event Driven (ED). Moreover, we have a series of strategies that are similar and could be considered Directional strategies (i.e. FTSE. Directional.Strategies and HFRX Market.Directional) and we provide a broad comment about them.

Table 1 shows that the Convertible Arbitrage strategy presents a standard deviation in the FTSE index that is twice as large compared to a standard deviation in all other datasets, and skewness that is almost zero in the MSCI index, it is positive in the FTSE index and negative in the HFRX and DJ indexes. Regarding the excess kurtosis, we have that it is above 3 for MSCI, above 2 for DJ and FTSE and at 1.6 for HFRX. The same applies to the Event Driven strategy where skewness is almost zero for MSCI and FTSE and it is quite negative for HFRX and DJ indexes. Excess kurtosis then ranges from 1.47 in DJ to 12.9 in MSCI dataset. Directional strategies are more similar and present negative skewness and an excess kurtosis above 3. The Directional Strategy in FTSE has the largest negative return. It also has one of the largest minimum and maximum returns.

We also examine autocorrelation for each hedge fund index using Box-Pierce test with one lag. For all datasets, there is still evidence of autocorrelation at the index level, but evidence is quite different if we look at single strategies. As previous

literature has stressed, the general level of autocorrelation is lower for daily data (see Boyson et al., 2007) but this reduction in terms of autocorrelation is different among strategies and datasets. For example, we have that the CA strategy always presents autocorrelation, but for the ED strategy in the MSCI dataset returns are not autocorrelated. Also for the Equity Market Neutral strategy (or non-directional strategies) the hypothesis of autocorrelation is rejected for MSCI, HFRX and DJ; for FTSE instead we do find evidence of autocorrelation.[5] These results confirm again our finding that similar strategies constructed by different database providers display significant statistical differences. One possible reason of these differences among the datasets may come from different levels of selection bias. Since managers are willing to provide daily pricing, they are less likely to hold illiquid positions or have less discretion in pricing, but this feature may be different among hedge funds included in the different datasets.

If we compare the above statistics with those of the S&P500 daily returns we observe that the mean of the S&P500 is two or three times higher than the one of the hedge fund indexes and the standard deviation is more than three times higher. This clearly indicates that if we consider the Sharpe ratio of the S&P500, it would be lower than the Sharpe ratio of each of these indexes.

This aspect is well captured by cumulated returns presented in Figures 1 and 2 where we observe that the behavior of the different datasets is similar when the evolution of returns is compared; the returns of the S&P500 are largely higher than those of the hedge funds. Regarding the single indexes we observe that in the first part of the sample the FTSE index provides the highest returns, but in the second part of the

---

[5] FTSE index is the most restrictive compared to other four. For example, the total hedge fund index always has a constant of 40 funds, and individual strategy indexes can have as low as 3 funds in each index.

sample it provides the lowest returns. This provides additional evidence that this dataset

(FTSE) diverges the most from other datasets.

	n. obs.	Mean Returns %	Sd %	Skw	Excess kurt	Min	Max	Box-Pierce p-value	JB p-value
MSCI.Hedge.Invest.Index	**997**	**0.02**	**0.15**	**-0.9**	**3.54**	**-0.92**	**0.44**	**0.000**	**0.000**
MSCI.Hedge.Invest.Equity.Non.Dir	997	0.02	0.18	-0.59	4.27	-1.1	0.8	0.500	0.000
MSCI.Hedge.Invest.Conv.Arbitrage	997	0.01	0.18	0.03	3.87	-0.88	1.18	0.000	0.000
MSCI.Hedge.Invest.Event.Driven	997	0.02	0.22	-0.09	12.9	-1.94	1.98	0.002	0.000
MSCI.Hedge.Invest.Syst..Trading	997	0.02	0.48	-0.3	4.92	-3.46	3.08	0.029	0.000
MSCI.Hedge.Invest.Discr..Trading	876	0.01	0.28	0.05	2.83	-1.38	1.48	0.224	0.000
MSCI.Hedge.Invest.Long.Bias	997	0.04	0.27	-0.58	1.97	-1.41	1.04	0.000	0.000
MSCI.Hedge.Invest.Variable.Bias	752	0.04	0.31	-0.62	3.14	-1.56	1.29	0.000	0.000
FTSE.Hedge.Fund	**754**	**0.02**	**0.18**	**-1.28**	**6.41**	**-1.21**	**0.64**	**0.000**	**0.000**
FTSE.Directional.Strategies	754	0.02	0.35	-0.97	6.22	-2.38	1.58	0.000	0.000
FTSE.Equity.Hedge	754	0.04	0.38	-0.43	7.97	-2.35	2.58	0.078	0.000
FTSE.NonDir. Strategies	754	0.01	0.15	0.06	5.28	-0.83	0.99	0.000	0.000
FTSE.Equity.Arbitrage	754	0.01	0.18	0.35	3.89	-0.88	0.94	0.169	0.000
FTSE.Convertible.Arbitrage	754	0	0.4	0.14	2.15	-1.58	1.93	0.000	0.000
FTSE.Event.Driven.Strat.	754	0.03	0.21	-0.02	4.39	-0.95	1.08	0.222	0.000
FTSE.Merger.Arbitrage	754	0.02	0.23	-0.66	8.17	-1.79	1.04	0.018	0.000
FTSE.Distressed.Sec.and.Opport.	754	0.04	0.28	0.32	5.32	-1.35	1.76	0.286	0.000
HFRX.Equal.Weighted	**1070**	**0.02**	**0.13**	**-0.74**	**2.88**	**-0.76**	**0.59**	**0.000**	**0.000**
HFRX.Convertible.Arbitrage	1070	0.01	0.21	-0.33	1.6	-0.93	0.81	0.001	0.000
HFRX.Distressed.Securities	1070	0.04	0.15	0.44	3.15	-0.73	0.77	0.000	0.000
HFRX.Equity.Hedge	1070	0.03	0.32	-0.53	2.22	-1.82	1.51	0.000	0.000
HFRX.Equity.Market.Neutral	1070	0.01	0.18	0.19	1.09	-0.62	0.89	0.307	0.000
HFRX.Event.Driven	1070	0.04	0.24	-0.38	1.97	-1.21	1.04	0.001	0.000
HFRX.Merger.Arbitrage	1070	0.02	0.17	-0.55	2.71	-0.86	0.71	0.774	0.000
HFRX.Relative.Value.Arbitrage	1070	0.02	0.17	0.39	9.65	-0.95	1.63	0.005	0.000
HFRX.Absolute.Return	754	0.02	0.13	-0.17	1.46	-0.47	0.5	0.000	0.000
HFRX.Market.Directional	754	0.04	0.39	-0.88	3.58	-2.64	1.33	0.000	0.000
Dow.Jones.Equal Weighted	**627**	**0.03**	**0.15**	**-0.45**	**1.21**	**-0.6**	**0.64**	**0.000**	**0.000**
Dow.Jones.Convertible.Arbitrage	879	0.01	0.2	-0.53	2.47	-1.18	0.7	0.000	0.000
Dow.Jones.Distressed.Securities	879	0.05	0.18	0.04	4.97	-0.99	1.07	0.002	0.000
Dow.Jones.Event.Driven	879	0.04	0.24	-0.38	1.47	-1.12	0.98	0.006	0.000
Dow.Jones.Equity.Market.Neutral	879	0.01	0.12	0.71	1.91	-0.37	0.55	0.129	0.000
Dow.Jones.Merger.Arbitrage	879	0.03	0.18	-0.18	1.43	-0.74	0.63	0.115	0.000
Dow.Jones.Equity Long Short	670	0.04	0.44	-0.44	3.16	-2.32	2.45	0.001	0.000
S&P500	1070	0.06	0.71	-0.07	0.86	-3.46	2.63	0.023	0.000

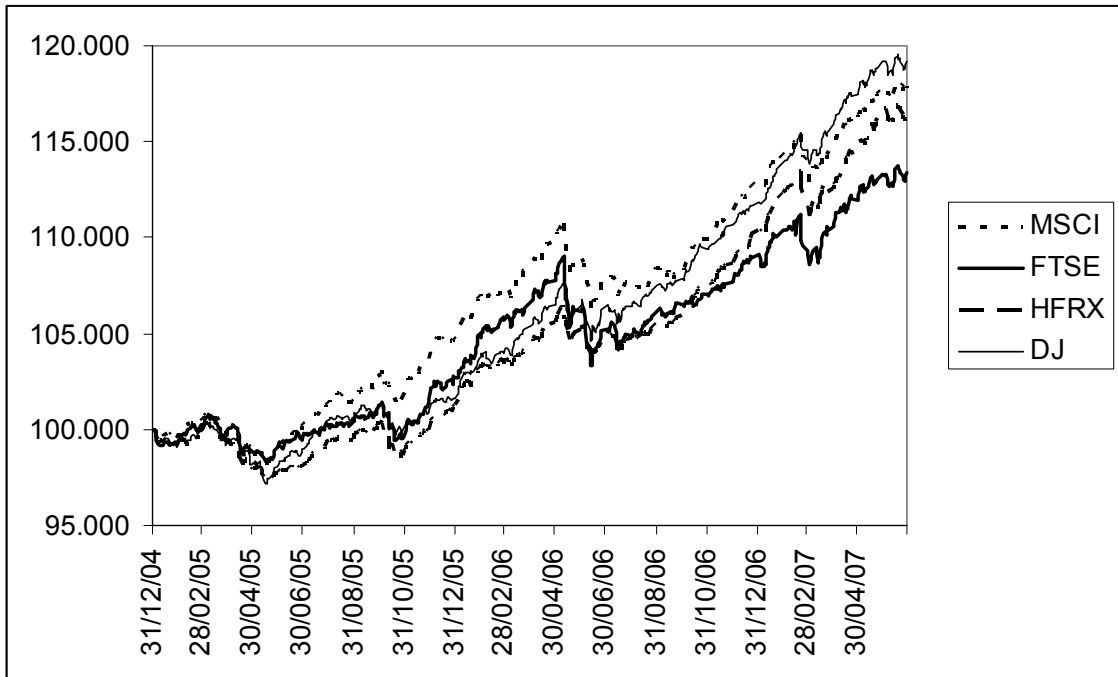Table 1. Descriptive statistics of daily datasets.

Figure 1. Cumulative returns of hedge fund indexes form January 2005 till June 2007.
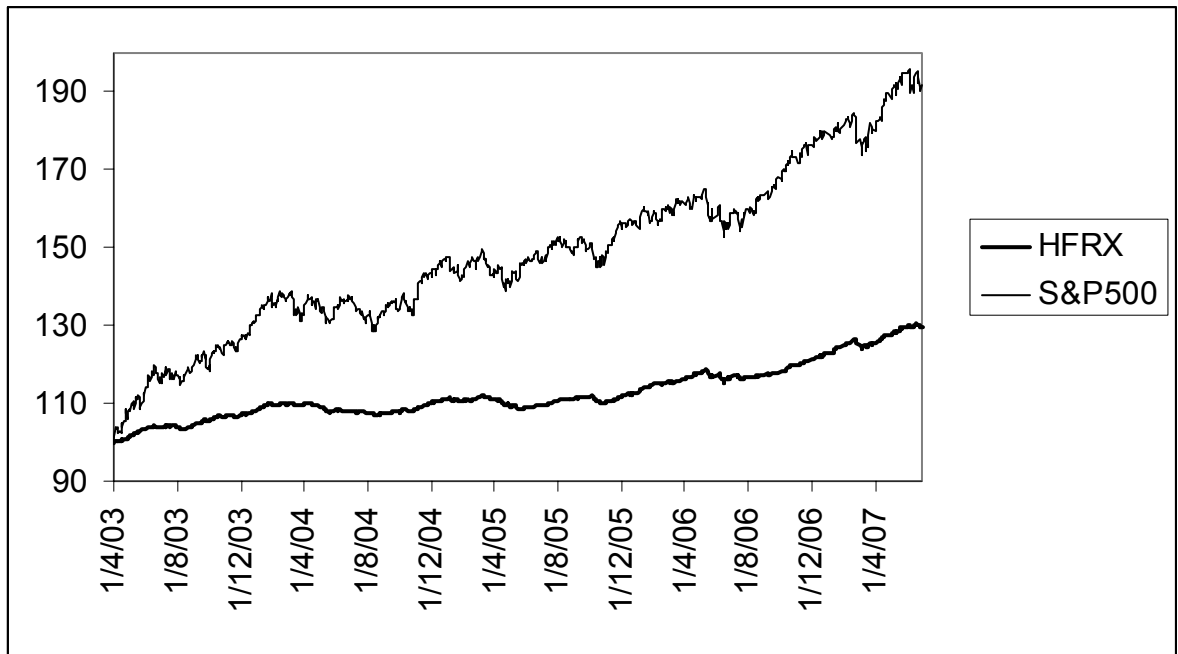

Figure 2: Cumulative returns of HFRX and S&P500 indexes from April 2003 till June 2007.

16

3.2. Monthly Data, Non-Investable Indexes

As described in the previous section, there are several different databases that collect monthly hedge fund data and use different methodologies to come up with monthly indexes. One aspect that it is crucial to consider when comparing databases is whether hedge fund indexes are investable. In terms of asset allocation and performance such characteristic is fundamental. However, in this work we are primarily interested in statistical properties of these datasets that are (in contrast to the daily datasets) largely used in the literature. In this section we concentrate on three different datasets considering only non-investable indexes: the Hedge Fund Research Indexes (HFRX indexes), the CASAM CISDM Database (formerly the MAR Database) and Credit Suisse/Tremont (CSFB) indexes. We investigate statistical properties of these indexes.

Table 2 provides summary statistics. CSFB and CISDM are more similar in terms of mean, standard deviation and skewness as well as minimum and maximum values. However, these features are not confirmed at the single strategy level. We observe that while Convertible Arbitrage (CA) presents similarities among the three datasets, the Event-Driven (ED) strategy largely differs in terms of excess kurtosis, ranging from 3.94 to 24.21. Almost all strategies present a distribution that is not normal and there is evidence of autocorrelation, with the only exceptions of directional strategies.

One aspect that it is fundamental to consider is that these databases are characterized by different sample periods which additionally contribute to varying properties among databases. We performed the same analysis using the identical sample period for all different datasets (in line with the HFR dataset) and the results are qualitatively equal. This confirms that statistical differences among the databases are mostly generated by the dataset construction rather than from different sample periods.

Nevertheless, we are interested in comparing daily and monthly datasets and for this reason we need to deeply investigate monthly investable indexes. The next two sections are dedicated to this issue.

	n. obs	Mean	Sd	Skw	Excess kurt	Min	Max	Box-Pierce p-value	JB p-value
CS.Trem.HF.Index	**162**	**0.91**	**2.17**	**0.07**	**2.43**	**-7.55**	**8.53**	**0.12**	**0.00**
CS.Trem.HF.Index.Convertible.Arbitrage	162	0.74	1.31	-1.38	3.39	-4.68	3.57	0.00	0.00
CS.Trem.HF.Index.Dedicated.Short.Bias	162	-0.09	4.82	0.84	2.13	-8.69	22.71	0.13	0.00
CS.Trem.HF.Index.Emerging.Markets	162	0.87	4.53	-0.71	4.89	-23.03	16.42	0.00	0.00
CS.Trem.HF.Index.Equity.Market.Neutral	162	0.80	0.82	0.33	0.41	-1.15	3.26	0.00	0.11
CS.Trem.HF.Index.Event.Driven	162	0.97	1.58	-3.39	24.21	-11.77	3.68	0.00	0.00
CS.Trem.HF.Index.Distressed	162	1.10	1.78	-2.95	19.57	-12.45	4.10	0.00	0.00
CS.Trem.HF.Index.ED Multi.Strategy	162	0.91	1.71	-2.46	16.16	-11.52	4.66	0.00	0.00
CS.Trem.HF.Index.Risk.Arbitrage	162	0.64	1.18	-1.16	6.30	-6.15	3.81	0.00	0.00
CS.Trem.HF.Index.Long.Short.Equity	162	1.02	2.85	0.19	3.97	-11.43	13.01	0.03	0.00
CS.Trem.HF.Index.Multi.Strategy	162	0.81	1.23	-1.21	3.29	-4.76	3.61	0.62	0.00
CISDM Equal Weighted Hedge Fund Index	**211**	**1.18**	**1.94**	**-0.33**	**3.64**	**-8.82**	**8.37**	**0.00**	**0.00**
CISDM Convertible Arbitrage Index	187	0.85	0.78	-1.07	2.52	-2.45	2.96	0.00	0.00
CISDM Distressed Securities Index	211	1.13	1.82	-0.86	5.07	-9.22	6.05	0.00	0.00
CISDM Emerging Markets Index	211	1.21	4.05	-1.18	11.80	-26.25	19.33	0.00	0.00
CISDM Equity Market Neutral Index	211	0.74	0.52	0.05	2.49	-1.33	2.76	0.00	0.00
CISDM Equity Long/Short Index	211	1.11	2.35	-0.28	2.55	-9.42	9.40	0.00	0.00
CISDM Event Driven Multi-Strat. Index	211	1.06	1.50	-1.06	3.94	-6.73	4.81	0.00	0.00
HFRX Equal Weigh. Strat. Index	**115**	**0.67**	**1.06**	**-0.84**	**4.38**	**-4.53**	**3.28**	**0.00**	**0.00**
HFRX Absolute Return Index	115	0.57	0.92	-1.27	4.01	-3.84	2.31	0.00	0.00
HFRX Convertible Arbitrage Index	115	0.54	1.12	-0.85	1.80	-3.44	3.17	0.00	0.00
HFRX Distressed Securities Index	115	0.79	1.62	-0.27	2.87	-5.96	6.11	0.00	0.00
HFRX Equity Hedge Index	115	0.99	2.20	0.64	2.69	-6.05	9.78	0.02	0.00
HFRX Equity Market Neutral Index	115	0.25	1.05	-0.10	0.33	-2.69	2.92	0.00	0.69
HFRX Event Driven Index	115	0.74	1.82	-1.47	6.58	-9.02	4.79	0.00	0.00
HFRX Market Directional Index	115	0.92	2.26	-0.40	1.39	-7.81	6.27	0.02	0.00
HFRX Merger Arbitrage Index	115	0.57	1.10	-1.39	4.58	-4.56	3.29	0.09	0.00
HFRX Relative Value Arbitrage Index	115	0.62	1.13	-1.45	6.74	-4.67	4.02	0.00	0.00

Table 2. – Descriptive statistics of monthly datasets (non-investable indexes).

3.3. Monthly Data, Investable Indexes

The only data provider that has monthly indexes available for a long period of time for investable hedge funds is Credit Suisse/Tremont (CSFB). In this section we analyze the statistical characteristics of investable indexes and how they differ with respect to non-investable indexes at the monthly level. Table 3 provides summary statistics. We also report the statistics for the CSFB non-investable index in order to match the sample periods and avoid the possibility that statistical differences come from the sample period and not from the composition of the indexes.

	n. obs	Mean	Sd	Skw	Excess Kurt	Min	Max	Box-Pierce p-value	JB p-value
CS.Trem.HF.Index	**90**	**0.73**	**1.42**	**0.17**	**4.02**	**-4.63**	**6.49**	**0.42**	**0.00**
CS.Trem.HF.Index.Convertible.Arbitrage	90	0.79	1.22	-0.60	1.67	-3.15	3.57	0.00	0.00
CS.Trem.HF.Index.Dedicated.Short.Bias	90	-0.05	4.52	0.35	-0.04	-8.65	13.76	0.15	0.39
CS.Trem.HF.Index.Emerging.Markets	90	1.02	2.71	-0.89	0.89	-7.76	6.04	0.15	0.00
CS.Trem.HF.Index.Equity.Market.Neutral	90	0.72	0.59	0.55	-0.05	-0.34	2.24	0.02	0.10
CS.Trem.HF.Index.Event.Driven	90	0.94	1.13	-0.93	2.40	-3.10	3.27	0.01	0.00
CS.Trem.HF.Index.Distressed	90	1.02	1.27	-1.26	4.37	-4.27	3.91	0.00	0.00
CS.Trem.HF.Index ED Multi.Strategy	90	0.92	1.26	-0.27	2.09	-3.10	4.38	0.07	0.00
CS.Trem.HF.Index.Risk.Arbitrage	90	0.51	0.99	-0.35	2.33	-2.73	3.52	0.02	0.00
CS.Trem.HF.Index.Long.Short.Equity	90	0.65	2.35	0.49	4.85	-7.46	11.14	0.38	0.00
CS.Trem.HF.Index.Multi.Strategy	90	0.81	0.89	-0.24	-0.03	-1.82	2.76	0.00	0.66
CS.Trem.Investable.Index	**90**	**0.66**	**0.81**	**0.14**	**-0.43**	**-1.08**	**2.71**	**0.16**	**0.66**
CS.Trem.Investable.Index.Convertible.Arbitrage	90	0.80	1.01	-0.15	0.38	-2.36	3.40	0.00	0.57
CS.Trem.Investable.Index.Dedicated.Short.Bias	90	-0.23	4.22	-0.07	0.63	-13.57	10.88	0.16	0.37
CS.Trem.Investable.Index.Emerging.Markets	90	1.49	2.55	-0.33	-0.12	-5.07	7.40	0.78	0.43
CS.Trem.Investable.Index.Equity.Market.Neutral	90	0.50	0.56	0.99	1.05	-0.48	2.39	0.18	0.00
CS.Trem.Investable.Index.Event.Driven	90	0.89	0.96	-0.06	0.42	-1.76	3.48	0.02	0.61
CS.Trem.Investable.Index.Long.Short.Equity	90	0.42	2.71	0.29	3.33	-8.00	10.31	0.71	0.00
CS.Trem.Investable.Index.Multi.Strategy	90	0.76	0.99	0.61	2.08	-1.54	4.87	0.01	0.00
CS.Trem.Sector.Invest.Indexes.Convertible.Arb	90	0.71	0.98	-0.25	0.62	-2.25	3.28	0.00	0.25
CS.Trem.Sect.Inv.Indexes.Dedicated.Short.Bias	90	-0.17	4.13	-0.08	-0.13	-12.54	9.45	0.11	0.94
CS.Trem.Sect.Inv.Indexes.Emerging.Markets	90	1.37	2.64	-0.21	-0.41	-5.10	6.84	0.21	0.56
CS.Trem.Sect.Inv.Indexes.Equity.Market.Neutral	90	0.42	0.40	0.61	0.02	-0.43	1.45	0.00	0.05
CS.Trem.Sect.Inv.Indexes.Event.Driven	90	0.86	1.09	-0.59	1.13	-2.83	3.15	0.00	0.00
CS.Trem.Sect.Inv.Indexes.Long.Short.Equity	90	0.78	1.62	0.03	0.59	-3.35	5.68	0.27	0.43
CS.Trem.Sect.Inv.Indexes.Multi.Strategy	90	0.81	0.90	0.67	2.25	-1.43	4.57	0.00	0.00
S&P500	90	0.11	4.03	-0.30	0.42	-11.00	9.67		

Table 3.- Summary statistics of monthly CSFB investable and non-investable indexes.

As Table 3 shows, the two datasets have very different characteristics. The non-investable index presents a larger mean and a larger standard deviation compared to the investable index. Skewness is quite similar but excess kurtosis is quite different and it also has an opposite sign. Note that the investable index presents a level of excess kurtosis that is closer to zero and therefore resembles normality. Moreover, it is important to consider that the index is based not only on equity-linked strategies that we consider, but it is also including Global Macro and Fixed Income strategies that are not included separately in our analysis. Thus, the index is representative of different hedge fund strategies.

The Convertible Arbitrage (CA) strategy presents similar mean and standard deviation but again the main differences are in skewness and excess kurtosis measures. We do not observe large differences among the investable index and the investable sector index for this strategy, even if the number of hedge funds included in these two indexes is quite different. The investable index in fact includes only the few largest representative hedge funds (this index is usually refereed to as the "blue chip" hedge fund index) and the investable sector index is more inclusive.

The Event Driven (ED) strategy provides the same features. Overall we have that investable indexes have distributions with similar means and standard deviation but with skewness and excess kurtosis that are largely more close to normality compared to non-investable counterparts. This aspect is quite relevant if we consider that investable indexes are characterized by a lower number of hedge funds and therefore, if the law of the large numbers is true, we may expect the opposite. Specifically, we might expect that non-investable indexes might have distributions of returns closer to normal compared to investable indexes. Given these results it seems that the large excess kurtosis that characterizes the non-investable indexes is very common. Therefore, there

is a presence of common characteristics among non-investable hedge fund indexes that prevents the convergence to normality when the number of funds is increased.

In Table 3 we also document autocorrelation for each hedge fund index. The Table indicates that investable indexes present no autocorrelation (as evidenced by Box Pierce p-value) at the total index level (and also the non-investable index is showing this feature) and among strategies we have that there is evidence of autocorrelation only for three strategies: CA, ED and Multistrategy. The same applies, but at a lower level, to the investable sector strategies. It is interesting to observe that the Long-Short strategy, i.e., the most prevalent strategy in the hedge fund industry, presents no evidence of autocorrelation. This result could be interpreted as the signal that the managers of these hedge funds are less likely to hold illiquid positions or have less discretion in pricing since they may be affected by investment and disinvestment from their investors.

Normality tests indicate that most of investable strategy indexes have normal return distributions. However, we still have non-normality for ED, Equity Market Neutral and Multistrategy. There are no differences among the investable indexes and the investable sector datasets.

Overall, investable indexes, compared to non-investable counterparts, show evidence of normality in the distribution of returns.

3.4. Comparison Among Monthly and Daily Data

In the previous section we have separately analyzed daily and monthly returns. Tables 1 and 3 allow us to compare characteristics of daily and monthly returns for investable indexes. We find that at the monthly level the investable index has no autocorrelation, but we observe autocorrelation at the daily level for all the daily indexes. This result is in line with the literature of financial time series (Campbell, Lo and MacKinlay, 1997): autocorrelation at daily level that disappears at monthly level. In order to test this

conjecture, we construct monthly returns from our daily dataset. Construction of monthly returns from daily data allows us to keep the same sample of hedge funds in the analysis. Table 4 shows these results and Table 5 provides descriptive statistics for investable indexes with the same sample period.

We find that monthly returns constructed from the daily datasets for all indexes are normally distributed; therefore, these constructed returns present the same characteristics as the S&P500: at daily level normality is rejected, at monthly level normality is accepted. This result is also confirmed in terms of skewness and excess kurtosis: for all datasets these moments are close to zero. If we compare these statistics with those presented in Table 5, we observe that for the sample period considered, skewness and excess kurtosis are closer to zero, at least at index level; therefore, this result could be related to the relative stable period we are considering in the sample. However, if we consider single strategies, we have that CA strategy in Table 4 shows evidence of non-normality also at monthly level in two datasets: HFR and DJ. This result is not surprising given the crisis we observed in this strategy in the middle of 2005. On the contrary, it is surprising that this does not appear in the other two datasets. Both investable and non-investable indexes (Table 5) exhibit a large excess kurtosis for the CA strategy and the hypothesis of normality is rejected. However, skewness and excess kurtosis are close to zero with respect to daily data for all other datasets. Autocorrelation disappears at monthly level in all datasets except Convertible Arbitrage and Distressed Security strategies.

Lastly, taking into account the "investability" is important when comparing mean returns of hedge funds to the S&P500. We find that in the last three years the mean return of the S&P500 was higher than that for all investable indexes, but is lower for non-investable indexes.

	n. obs	Mean	Sd	Skw	Excess kurt	Min	Max	Box-Pierce p-value	JB p-value
MSCI.Hedge.Invest.Index	**47**	**0.52**	**0.8**	**-0.29**	**-0.85**	**-1.15**	**2.19**	**0.0491**	**0.4063**
MSCI.Hedge.Invest.Equity.Non.Dir.l	47	0.42	0.66	0.54	1.25	-0.94	2.56	0.2299	0.0383
MSCI.Hedge.Invest.Conv.Arbitrage	47	0.19	0.78	-0.53	0.16	-1.77	2	0.0478	0.2801
MSCI.Hedge.Invest.Event.Driven	47	0.53	0.95	0.2	0.72	-1.49	3.21	0.3442	0.3966
MSCI.Hedge.Invest.Systematic.Trad.	47	0.48	1.96	-0.11	-0.35	-4.42	4.78	0.5891	0.9041
MSCI.Hedge.Invest.Discretionary.Trad.	41	0.3	1	-0.29	-0.92	-1.85	1.97	0.149	0.412
MSCI.Hedge.Invest.Long.Bias	47	0.85	1.35	-0.54	-0.51	-2.2	3.19	0.0783	0.2589
MSCI.Hedge.Invest.Variable.Bias	35	0.9	1.59	-0.95	0.8	-3.86	3.42	0.6124	0.0268
FTSE.Hedge.Fund	**35**	**0.41**	**0.85**	**-0.46**	**-0.37**	**-1.49**	**2.01**	**0.6566**	**0.4986**
FTSE.Directional.Strategies	35	0.53	1.43	-0.69	-0.06	-2.9	3.11	0.5442	0.2218
FTSE.Equity.Hedge	35	0.81	1.6	-1.01	0.88	-3.5	3.29	0.4167	0.0157
FTSE.NonDirectional.Strategies	35	0.15	0.51	0.52	0.44	-0.74	1.56	0.65	0.3155
FTSE.Equity.Arbitrage	35	0.25	0.62	0.17	-0.54	-1.05	1.52	0.0438	0.8164
FTSE.Convertible.Arbitrage	35	0.06	1.05	-0.21	-0.11	-2.81	2.17	0.847	0.8716
FTSE.Event.Driven.Strategies	35	0.67	0.94	-0.76	0.38	-2.02	2.29	0.5189	0.1218
FTSE.Merger.Arbitrage	35	0.37	1.02	-0.11	0.77	-2.02	3.08	0.458	0.4703
FTSE.Distressed.Sec.and.Opportunities	35	0.88	1.24	-0.2	-0.25	-2.04	3.91	0.2797	0.8737
HFRX.Equal.Weighted.Strategies	**51**	**0.52**	**0.85**	**-0.28**	**-0.55**	**-1.47**	**1.99**	**0.0202**	**0.5662**
HFRX.Convertible.Arbitrage	51	0.17	1.18	-0.85	0.78	-3.44	2.75	0.0008	0.0154
HFRX.Distressed.Securities	51	0.77	1.07	-0.09	-0.51	-1.71	2.92	0.0092	0.8057
HFRX.Equity.Hedge	51	0.67	1.53	-0.2	-0.42	-2.64	4.23	0.1439	0.7475
HFRX.Equity.Market.Neutral	51	0.18	0.82	0.02	0.27	-1.94	2.45	0.4558	0.8387
HFRX.Event.Driven	51	0.86	1.31	-0.21	-0.64	-2.05	3.53	0.0867	0.5987
HFRX.Merger.Arbitrage	51	0.52	0.91	-1.09	2.34	-2.85	2.32	0.1841	0
HFRX.Relative.Value.Arbitrage	51	0.44	0.69	-0.14	-0.37	-1.23	1.84	0.0045	0.8503
HFRX.Absolute.Return	35	0.44	0.8	-1.06	1.04	-1.99	1.59	0.1017	0.0087
HFRX.Market.Directional	35	0.83	1.99	-0.33	-0.75	-3.03	4.34	0.9461	0.5364
Dow.Jones.Bal.Portf.Index.EW	**30**	**0.59**	**0.82**	**-0.48**	**-0.2**	**-1.37**	**2.15**	**0.2018**	**0.5336**
Dow.Jones.Convertible.Arbitrage	42	0.2	1.13	-1.29	1.93	-3.68	1.84	0.006	0
Dow.Jones.Distressed.Securities	42	1.01	1.03	0.24	-0.17	-0.81	3.58	0.4494	0.8087
Dow.Jones.Event.Driven	42	0.84	1.09	-0.21	-0.46	-1.92	3.1	0.1933	0.7624
Dow.Jones.Equity.Market.Neutral	42	0.28	0.41	0.04	-0.39	-0.58	1.18	0.8358	0.9364
Dow.Jones.Merger.Arbitrage	42	0.6	0.96	0.53	1.28	-1.29	3.68	0.0613	0.0497
Dow.Jones.Equity.Long.Short	32	0.78	1.91	-0.3	-0.54	-3.53	4.82	0.8454	0.7002
S&P500	51	1.31	2.39	0.3	0.04	-3.31	8.24	0.8287	0.6538

Table 4. Summary statistics for monthly returns reconstructed from daily data.

	n. obs	Mean	Sd	Skw	Excess Kurt	Min	Max	Box-Pierce p-value	JB p-value
CS.Trem.HF.Index	**35**	**1.02**	**1.10**	**-0.48**	**-0.03**	**-1.46**	**3.23**	**0.28**	**0.17**
CS.Trem.HF.Index.Convertible.Arbitrage	35	0.51	1.08	-1.35	3.25	-3.13	2.75	0.00	0.00
CS.Trem.HF.Index.Dedicated.Short.Bias	35	-0.19	3.51	0.03	-0.44	-7.71	6.96	0.00	0.69
CS.Trem.HF.Index.Emerging.Markets	35	1.57	1.96	-1.29	3.43	-5.02	5.76	0.39	0.00
CS.Trem.HF.Index.Equity.Market.Neutral	35	0.73	0.58	0.62	0.18	-0.34	2.13	0.01	0.05
CS.Trem.HF.Index.Event.Driven	35	1.21	1.04	-0.38	1.24	-1.82	3.27	0.39	0.02
CS.Trem.HF.Index.Distressed	35	1.22	0.75	-0.06	-0.55	-0.10	2.67	0.20	0.56
CS.Trem.HF.Index.E.D.Multi.Strategy	35	1.24	1.37	-0.45	2.42	-3.10	4.31	0.67	0.00
CS.Trem.HF.Index.Risk.Arbitrage	35	0.59	0.77	0.91	2.24	-1.12	3.01	0.50	0.00
CS.Trem.HF.Index.Long.Short.Equity	35	1.21	1.65	-0.74	-0.01	-2.84	4.18	0.25	0.02
CS.Trem.HF.Index.Multi.Strategy	35	0.97	0.98	-0.37	-0.39	-1.10	2.76	0.01	0.27
CS.Trem.Investable.Index	**35**	**0.67**	**0.83**	**-0.27**	**-0.65**	**-1.08**	**2.04**	**0.29**	**0.33**
CS.Trem.Investable.Index.Conv.Arbitrage	35	0.44	0.89	-0.80	2.13	-2.36	2.47	0.00	0.00
CS.Trem.Investable.Index.Dedicated.Short.Bias	35	-0.55	3.16	-0.02	-0.75	-6.95	5.31	0.02	0.42
CS.Trem.Investable.Index.Emerging.Markets	35	1.76	2.67	-0.42	1.13	-5.07	7.40	0.13	0.05
CS.Trem.Investable.Index.Equity.Market.Neutral	35	0.44	0.42	0.88	0.65	-0.26	1.51	0.02	0.00
CS.Trem.Investable.Index.Event.Driven	35	1.07	1.11	-0.03	0.12	-1.34	3.48	0.21	0.97
CS.Trem.Investable.Index.Long.Short.Equity	35	0.85	1.39	-0.69	0.58	-2.77	3.43	0.57	0.03
CS.Trem.Investable.Index.Multi.Strategy	35	0.88	0.96	-0.24	-0.83	-1.18	2.34	0.01	0.24
CS.Trem.Sector.Invest.Indexes.Convertible.Arb	35	0.41	0.89	-0.66	1.58	-2.25	2.49	0.00	0.00
CS.Trem.Sect.Inv.Indexes.Dedicated.Short.Bias	35	-0.43	3.26	-0.07	-0.70	-7.60	5.61	0.03	0.45
CS.Trem.Sect.Inv.Indexes.Emerging.Markets	35	1.55	2.42	-0.52	1.36	-5.10	6.84	0.31	0.01
CS.Trem.Sect.Inv.Indexes.Equity.Market.Neutral	35	0.43	0.39	1.04	0.32	-0.06	1.35	0.06	0.00
CS.Trem.Sect.Inv.Indexes.Event.Driven	35	1.03	1.02	-0.25	-0.01	-1.24	2.98	0.23	0.68
CS.Trem.Sect.Inv.Indexes.Long.Short.Equity	35	1.06	1.58	-1.09	1.15	-3.35	4.03	0.63	0.00
CS.Trem.Sect.Inv.Indexes.Multi.Strategy	35	0.82	0.86	-0.24	-0.58	-1.13	2.40	0.00	0.42
S&P500	**35**	**0.91**	**2.03**	**-0.27**	**-0.84**	**-3.09**	**4.33**	**0.28**	**0.15**

Table 5. Summary statistics for CSFB monthly returns with the sample period from August 2004 till June 2007.

4. Non-Parametric Analysis

In this section we investigate deeper the characteristics of the daily dataset with the aim of analyzing the distribution of these returns. We know that they are not normally distributed, and therefore we study the shape of the distribution of the daily returns without imposing any model. Secondly, we investigate the link between hedge fund returns and the S&P500, and again we would like to leave the data to tell us the

peculiarity of this relationship. For this reason we use both the distribution analysis, regression analysis, and a non-parametric approach.

4.1. The Non-Parametric Technique

Many stochastic models used in financial studies are chosen to facilitate mathematical derivations and statistical inferences; often they are not derived from any economic theory and hence they cannot fit all financial data. These are called "*parametric*" models because they make some assumptions about the underlying process that generates the data, imposing on them a formal structure which can be fully described by a finite set of parameters: in this way, readily understandable summary statistics can be easily calculated.

Unfortunately, the strength of parametric models is also their weakness because by linking inference on an *a-priori* model, great gains in efficiency are possible only if the assumed model is, at least approximately, correctly specified.

Non-parametric techniques (also called *smoothing methods*) are useful instruments for this type of research since they allow relaxing the hypotheses on the data of non-serial correlation, normal distributions and linear dependence with risk factors. Indeed these methods are flexible tools in analysing unknown regression relationships; furthermore, they can show the detailed shapes of returns distributions.

Non-parametric techniques, in short, provide a bridge between making no assumption on the formal structure (a purely non-parametric approach) and making very strong ones (a parametric approach), moreover they lead to *robust* results.

The costs of such flexible statistical methods are computational, and, in some cases, difficulty in interpreting results because these methods provide graphical estimators and not numerical ones.

4.2. Kernel Smoothing: Density Probability Distribution

Let Y be a random variable with p.d.f. $f(y)$. The construction of the non-parametric estimator begins by dividing the sample space $\{y_1,...,y_n\}$ in a finite number of equally sized intervals. For each interval a rectangle with a height equal to the number of observations in the interval is drawn. y denotes the point at which the density $f(y)$ can be estimated with the kernel estimator of the following form:

$$\hat{f}(y) = \frac{1}{n}\sum_{i=1}^{n} w(y - y_i; h)$$

where w is itself a density, symmetric with zero mean, whose variance is controlled by the parameter h (usually h is the *standard deviation* of the kernel density). It is generally agreed that the particular shape of the kernel is not important, so it is often convenient to work with a Normal kernel, that is

$$w(y - y_i; h) = \phi(y - y_i; h)$$

where $\phi(z; h)$ indicates a Gaussian density in z with mean 0 and standard deviation h.

h is called a *smoothing parameter* or *bandwidth* as it controls the amount by which each observation is smoothed, i.e. locally averaged with different Gaussian-distributed weights depending on the distance form y_i, to produce $\hat{f}(y)$, which is also known as the "*naive*" estimator.

When h is too small, the estimate usually reflects variations associated with a single observation rather than the whole sample structure; on the other hand, if it is too large, important features of some sample regions might not be revealed as they are overshadowed by other observations.

The bandwidth can be chosen either by a trial and error approach looking at several different plots of $\hat{f}(y)$ against y for different values of h, or by mathematical derivation starting from the mean and the variance of $\hat{f}(y)$.

A measure of the efficiency of $\hat{f}$ in estimating f is the *Mean Integrated Squared Error* (MISE), which is in the one-dimensional case:

$$MISE\left(\hat{f}\right) = E\left\{\int\left[\hat{f}(y) - f(y)\right]^2 dy\right\} =$$
$$= \int\left[E\left\{\hat{f}(y)\right\} - f(y)\right]^2 dy + \int \mathrm{var}\left\{\hat{f}(y)\right\} dy$$

and can be approximated as:

$$MISE\left(\hat{f}\right) \approx \frac{1}{4}h^4\sigma_w^4\int f''(y)^2 dy + \frac{1}{nh}\alpha(w).$$

One of the most useful and logical way to find the correct value of h is to find the value h_{opt} that minimizes the MISE, which is:

$$h_{opt} = \left\{\frac{\gamma(w)}{\beta(f)n}\right\}^{-1/5}$$

where $\gamma(w) = \alpha(w)/\sigma_w^4$ and $\beta(f) = \int f''(y)^2 dy$. Unfortunately this formula still involves the unknown density function f; nevertheless, it is very informative because it tells us how the bandwidth should decrease, i.e. proportionally to $n^{-1/5}$.

The last step to find the value of h_{opt} is to assume that f is Normal. Therefore, the *normal optimal smoothing parameter* is:

$$h = \left(\frac{4}{3n}\right)^{1/5}\sigma = 1.059\sigma n^{-1/5}$$

where σ is the standard deviation of f that can be easily substituted with its estimate.

This approach requires very little calculation and being the Normal one of the smoothest possible distributions, the optimal value of h is large enough to induce some over-smoothing when applied to non-Normal data. This discourages over-interpretation of features due to sampling variation. For these reasons all the non-parametric estimates of this paper will be produced with a Gaussian kernels and Normal optimal smoothing parameters.

4.3. Kernel Smoothing: Non-Parametric Regression

The most widely used statistical procedure to describe the relationship between two (or more) variables is a linear regression. The linear regression can be defined as a parametric model where the target variable Y is linearly dependent on a set of regressors X_i, allowing for the prediction of future values of Y and the construction of tests and confidence intervals for predictions and parameters. The regression model can be written as:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \qquad i = 1,...,n$$

However, this model has very strong assumptions about the data generating process, i.e., $(x_i; y_i)$ is a bivariate Normal density, errors (ε) are independent and identically distributed with zero mean and finite variance. If these assumptions are not met, the estimators are biased and inconsistent (well-known problem of the misspecification in the functional form) and this is especially relevant for hedge fund returns which exhibit non-normal distributions, time-varying betas and non i.i.d. error terms.

Thus, a linear model should be substituted by a more general functional model like the *non-parametric regression model*:

$$y_i = m(x_i) + \varepsilon_i$$

where $m(x_i) = E(Y|X=x)$ does not have to have a linear form, $E(\varepsilon|X=x) = 0$ and $V(\varepsilon|X=x) = \sigma^2(x)$ are not necessarily constant. Removing the parametric restrictions allows for more flexible structural relationships.

The *local mean estimator* computes a local mean (applying a kernel to each observation) of the response variable (y) for each value of the covariate variable (x):

$$\tilde{m}(x)_{NW} = \frac{\sum_{i=1}^{n} w(x_i - x; h) y_i}{\sum_{i=1}^{n} w(x_i - x; h)}$$

where $w(z;h)$ is a kernel function, thus it is smoothed. This ensures that most weight is given to the observations whose covariates values x_i lie close to the point of interest x. As in the previous case, a Normal density function with a standard deviation h is used as the kernel, so the observations are over an effective range of 4σ. As h increases, the regression estimator becomes smoother (i.e. it misses some details in the curvature of the data). On the contrary, if the bandwidth decreases, the estimator tracks the data more closely (for a very small value it provides an uninformative interpolation of each observation).

The idea underlying the local linear approach is to estimate a linear regression locally for different regions of the sample, while the local mean computes only the average value for the same sub-sample of the data. For a high enough value of h the $\tilde{m}_{NW}$ is a straight line parallel to the x axis, with ordinate equal to the sample mean, and $\hat{m}(x)$ is equal to the common regression line. For this reason, and for its better performances in the extreme regions of the sample, the local regression method is used in this paper. The choice of h is obtained by the software "R" using an improved Akaike Information Criterion, even if it must have similar properties as the smoothing parameter described for the p.d.f. estimation.

Finally, the *variability bands*[6] that indicate the degree of variability of the estimates are drawn with the estimators. Regions with few significant observations have large variability bands for $\hat{m}(x)$. Alternative approaches to construct non-parametric regression estimators exist: *Gauss-Muller* method, *orthogonal series*, *Locally Varying Bandwidth* (also known as *loess*), and *splines*[7].

4.4 Non-Parametric Distribution of Returns

We analyze the non-parametric distributions of all the datasets we presented above. For the sake of space we only present results for the few of them: the most representative of different datasets and the ones that underlie peculiarity of the data.

Starting with the global indexes, we consider the FTSE and the Dow Jones that respectively present the largest and the lowest excess kurtosis. Figure 3 depicts the estimated conditional distributions of the indexes along with normal distributions that have the same mean and standard deviation as the corresponding hedge fund indexes. It is quite clear that the fitted distributions are more kurtotic than normal distributions, while it is not immediately clear if the fitted densities are symmetric. Moreover, this figure shows that the distributions of hedge fund index returns are well behaved and do not show multi-modality, even if in both cases they are not normal.

---

[6] Variability bands require very careful interpretation as they are computed differently than common confidence intervals. For more details see Bowman and Azzalini (1997).
[7] We verified for all sample used in this paper that the *smoothing splines* method often produces similar estimators to the local linear regression.
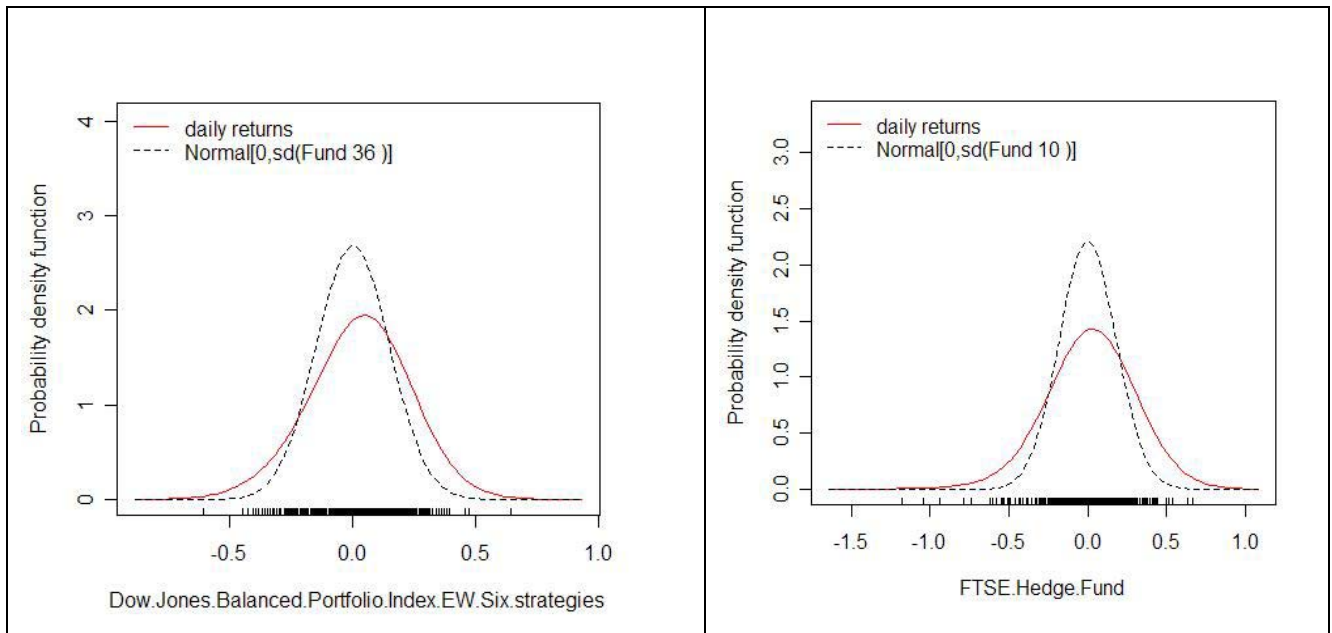
Figure 3. This figure depicts estimated conditional distributions of the hedge fund indexes (Dow Jones Balanced Portfolio Index and FTSE Hedge Fund Index) and normal distributions that have the same mean and standard deviation as the corresponding hedge fund indexes. Daily data is analyzed.

Results are similar when single strategies are analyzed. We show results for the Convertible Arbitrage strategy that experienced a crisis and more financial troubles than other strategies in the sample period considered. Similar to Figure 4, the estimated conditional distributions are well behaved and do not display any peculiarities.
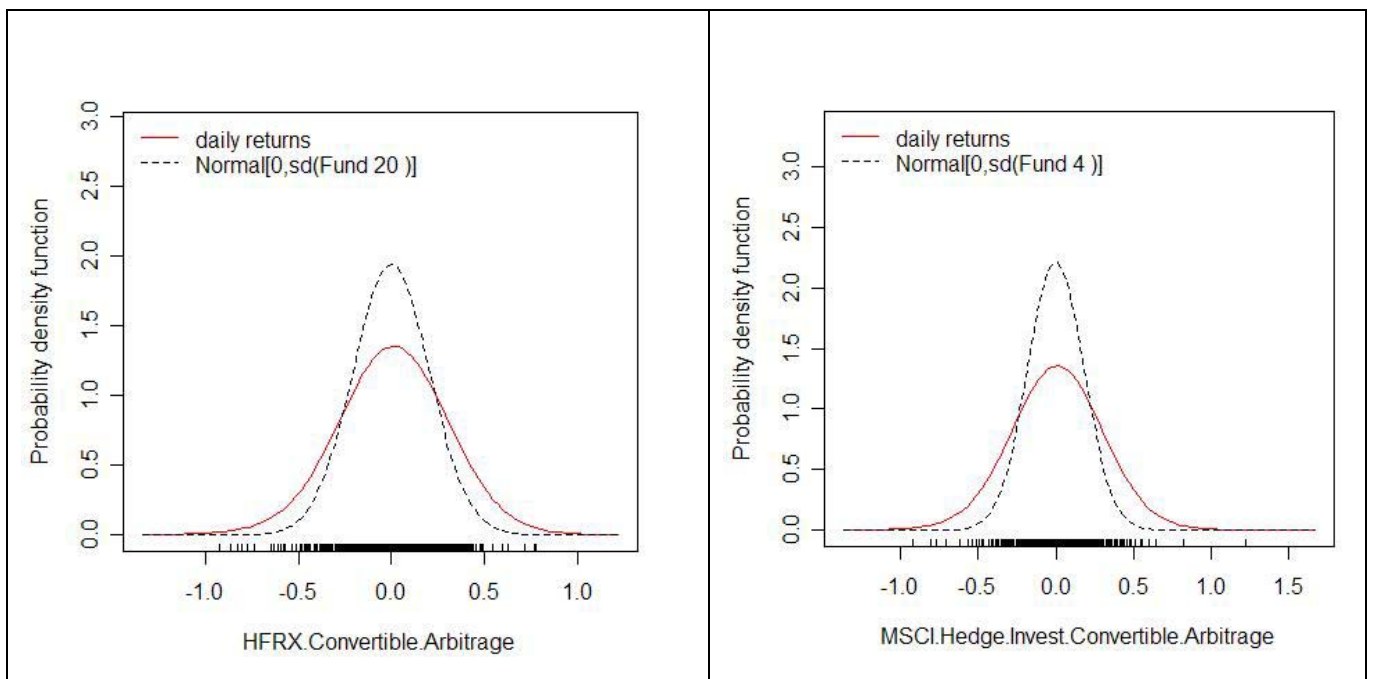
Figure 4. This figure depicts estimated conditional distributions of the Convertible Arbitrage indexes (HFRX and MSCI Investable Convertible Arbitrage Indexes) and normal distributions that have the same mean and standard deviation as the corresponding hedge fund strategy indexes. Daily data is analyzed.

If we apply the same analysis for the monthly datasets for investable indexes we observe that the shape is quite peculiar and is not well-behaved as shown in Figure 3.
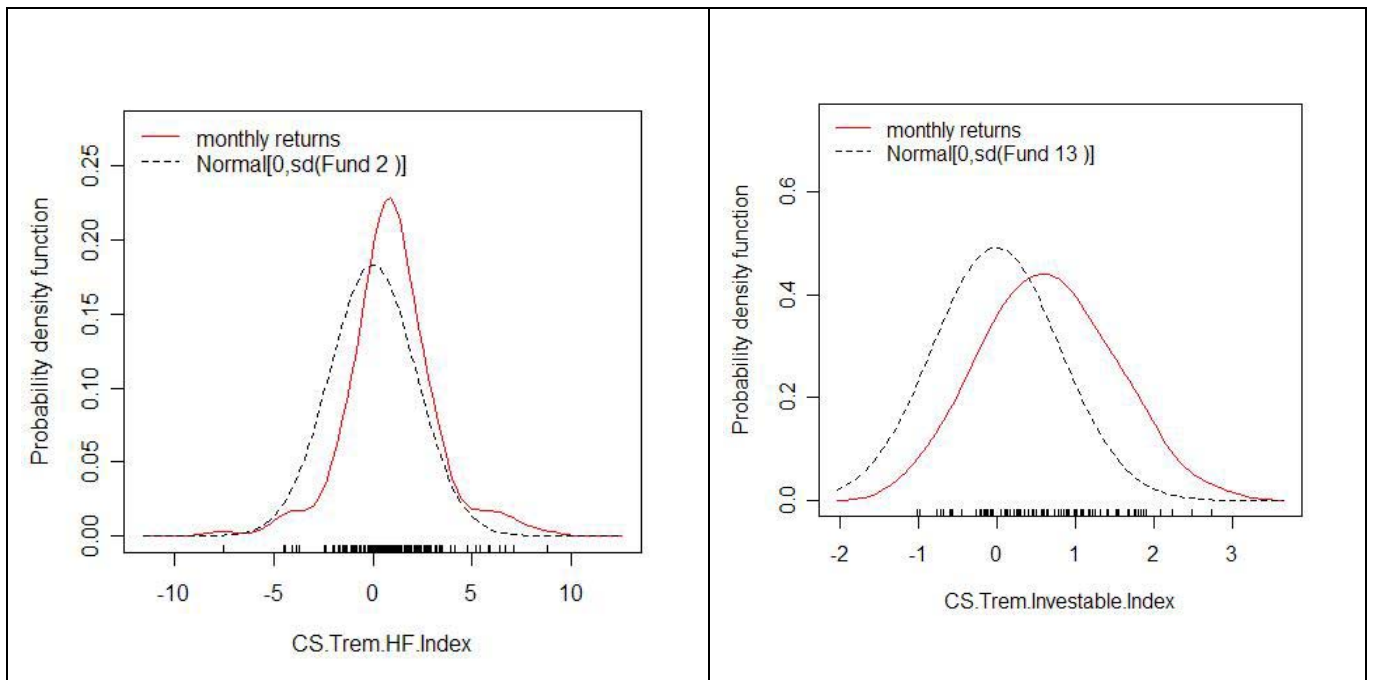
Figure 5. This figure depicts estimated conditional distributions of the hedge fund indexes (CS Tremont Hedge Fund and CS Tremont Investable Indexes) and normal distributions that have the same mean and standard deviation as the corresponding hedge fund indexes. The data for the hedge fund indexes is provided by the Credit Suisse First Boston (CSFB). The CS Tremont HF index is not investable and the CS Tremont Investable Index is an investable index. Monthly returns are analyzed.

On the left hand side of Figure 5 we present the non-parametric distribution of the CS Tremont Hedge Fund index using the sample period since the index's formation (1994). The sample period includes financial crises in the 1990s and at the beginning of this decade, and the Internet Bubble of 1998-2000. The density seems to pinpoint the presence of several bubbles. On the right hand side we present the non-parametric distribution of the CS Tremont Investable Index. The sample period for this index is

shorter (it starts in 2000), but we still find that the distribution largely differs with respect to the one presented in Figure 3.

By applying the same analysis on the monthly returns constructed from daily data we find that the distribution, even if normal, presents similar characteristics of the CS Tremont Investable Index as shown in Figure 6.
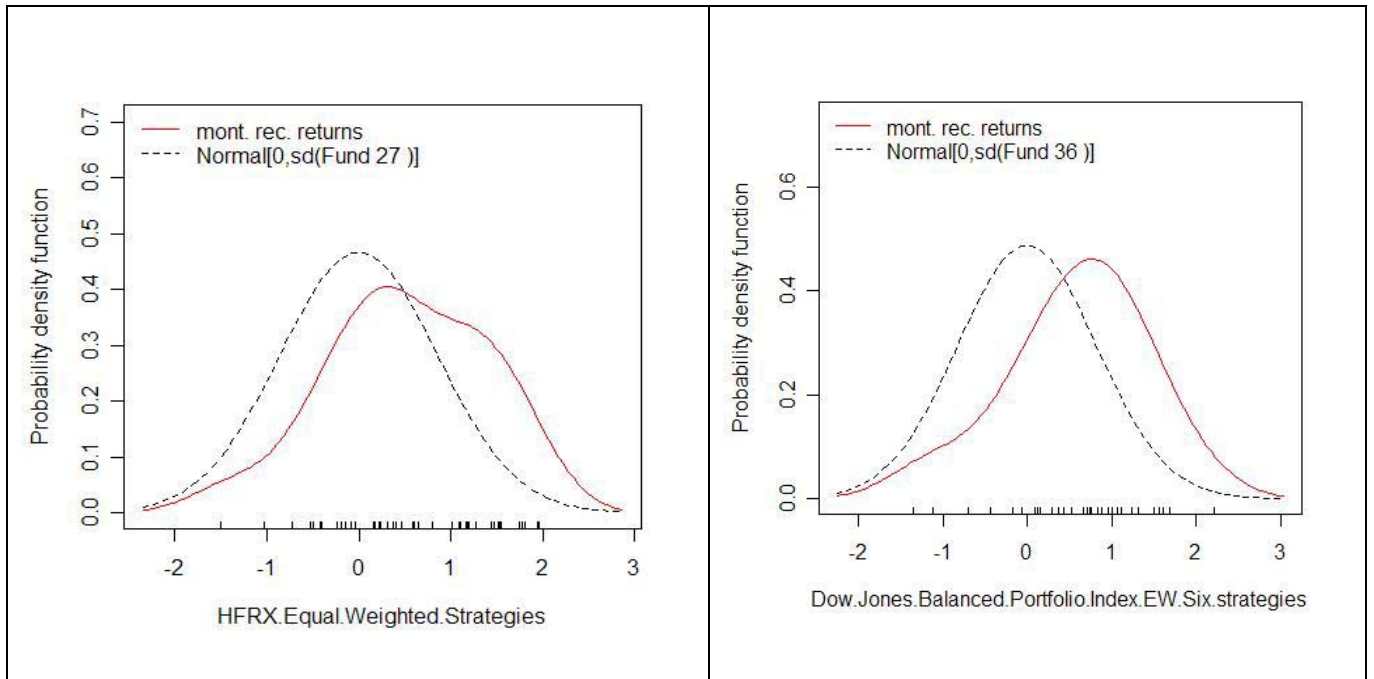

Figure 6. This figure depicts estimated conditional distributions of the hedge fund indexes (HFRX Equal Weighted and Dow Jones Balanced Portfolio Equal Weighted Indexes) and normal distributions that have the same mean and standard deviation as the corresponding hedge fund indexes. Monthly index returns are constructed from the daily returns provides by the HFR and Dow Jones.

4.5. Non-Parametric Regression

In this section we investigate the relationship between hedge fund indexes and the S&P500 index. In the previous sections we showed that hedge fund index densities are affected by the sample period, the frequency and the dataset provider, and we examined changes in the moments of the hedge fund returns for different conditioning information sets (different datasets, different frequencies, different sample periods, and

investability). In this section, we extend our analysis by examining the link between hedge funds and equity index: the S&P500.

Since the results of the previous section showed that hedge fund returns can be modeled with a non-parametric approach, we employ a non-parametric regression in the analysis of the hedge fund index exposure to the S&P500 index. This analysis is robust under various distributions.
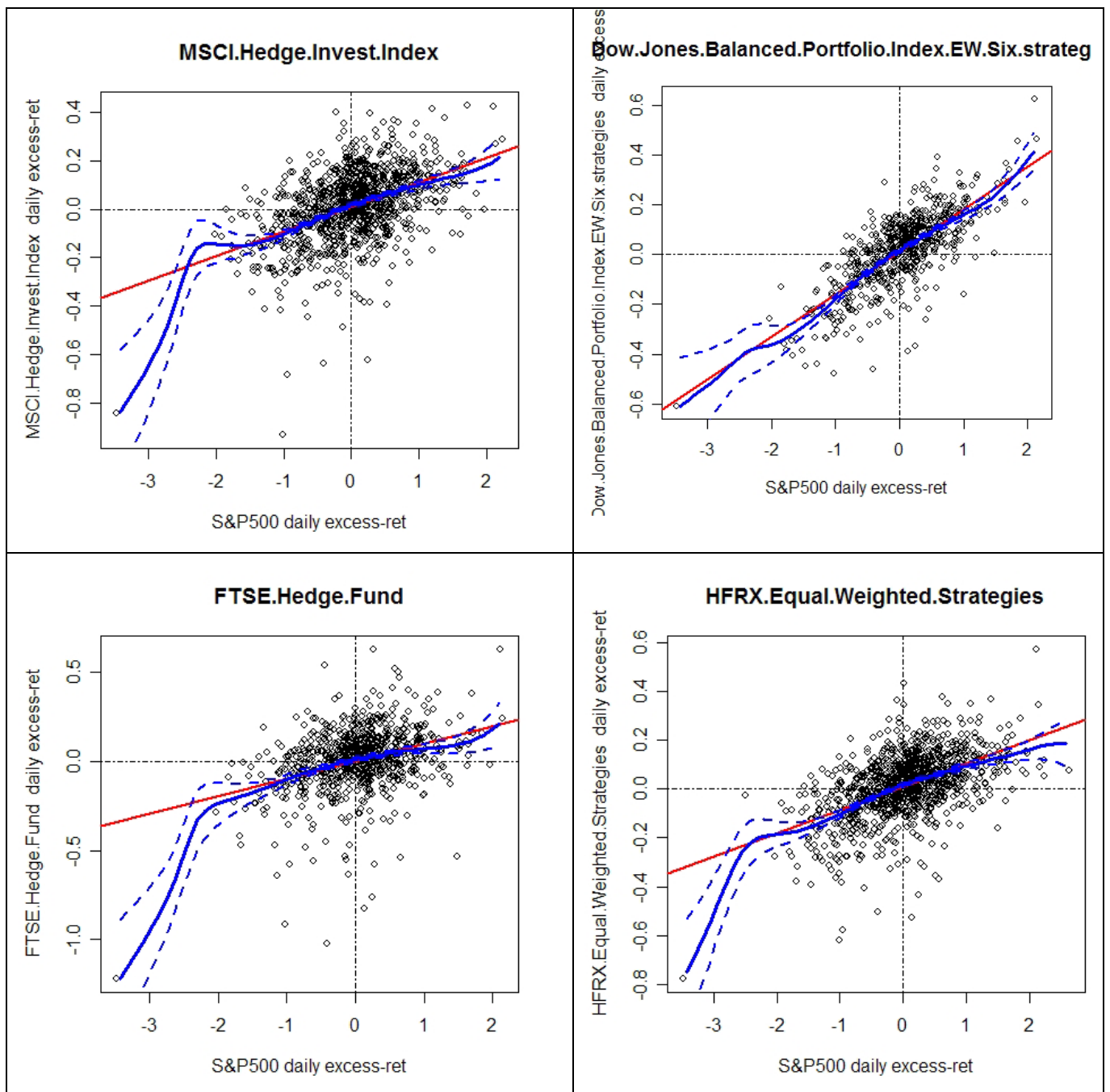

Figure 7. Non-parametric regression of the daily hedge fund return indexes on the S&P500. MSCI, Dow Jones, FTSE and HFRX indexes are considered.

According to the Figure 7, the link between the S&P500 and all daily hedge fund indexes (MSCI, Dow Jones, FTSE and HFRX) is positive and significant. The lower tail of the hedge fund index distributions exhibit a different exposure to the S&P500 compared to the most of observations that are concentrated in the center of each graph. The level of different exposures is statistically significant as shown by the variability bands (the dashed lines), that do not include the linear regression line (the thin solid line) on the left-hand side of the index return distributions. We find this difference even if the number of observations on the extremes is low and therefore the variability bands are larger. However, if we concentrate on daily strategy returns, we find that in most cases the exposure of these indexes to the S&P500 is not statistically different among mean and extreme returns of the strategy indexes. Figure 8 presents a few exceptions where tails of the hedge fund strategy return distributions have different exposure to the S&P500 compared to the mean returns.
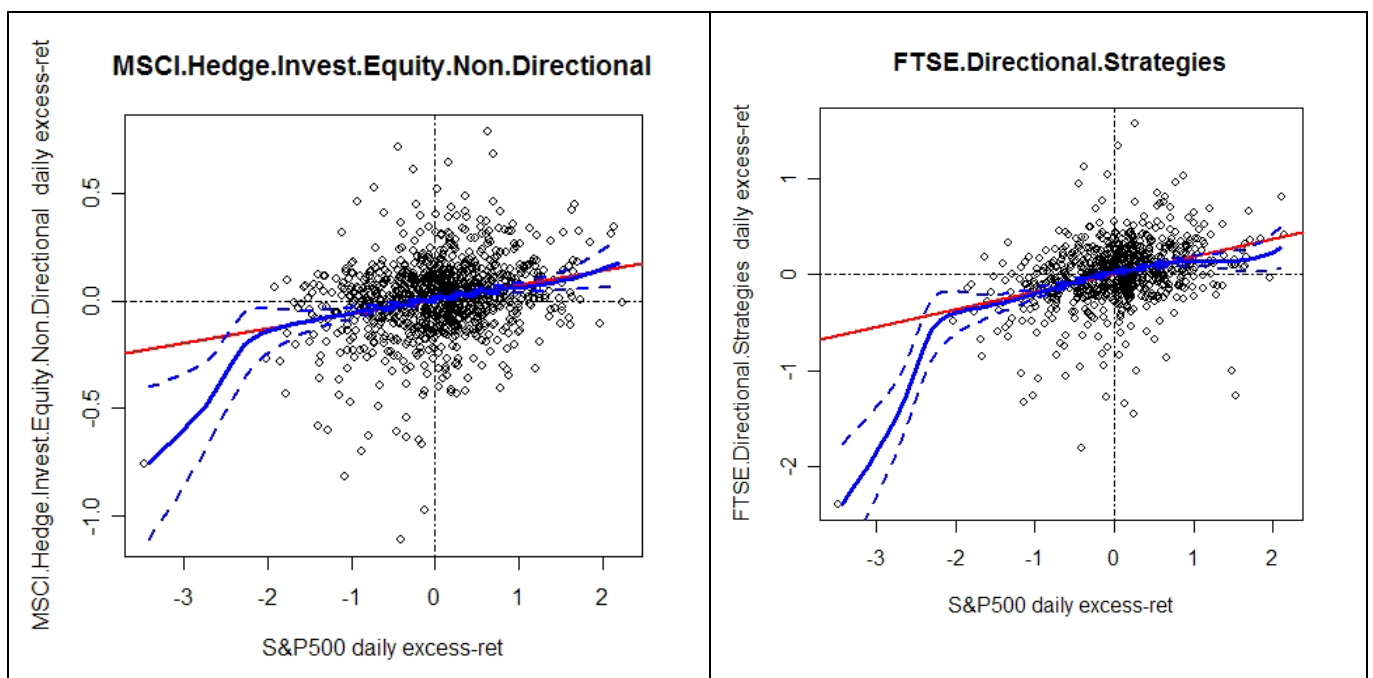

Figure 8. Non-parametric regression for MSCI Non-Directional and FTSE Directional strategies. Daily returns are analyzed.

5. Conclusion

We have examined the statistical properties of four hedge fund datasets: MSCI, DJ, HFR, and FTSE. Our analysis could be summarized in the following five main results. First, we find that FTSE is the dataset that highlights strong differences with respect to the other three datasets both at the total index level as well as for single strategies. This can be explained by the highly restrictive construction level of this database, where the number of funds in the overall index is always constant and stays at 40, and the number of hedge funds in individual strategy indexes can be as low as 3 funds.

Second, we observe the presence of statistical differences among investable and non-investable indexes, especially in terms of autocorrelation where investable indexes present lower evidence of autocorrelation.

Third, monthly returns constructed using daily data show almost no evidence of non-normality and autocorrelation in line with classical equity datasets, where autocorrelation on daily frequency can be diminished on monthly frequency.

Fourth, using the more flexible non-parametric approach, we are able to obtain estimates of conditional densities of daily returns and analyze exposure of hedge fund indexes to the S&P500 using the non-parametric regression. The analysis confirms that these datasets exhibit characteristics that are more similar to equity instruments.

Our analysis also confirms that a great variability in the statistical properties of different indexes is attributable to the existence of data biases in the database, like (i) selection bias: different number of hedge funds included in the database (and exclusion of funds not qualifying under selection criteria); funds are selected based on liquidity; (ii) survivorship bias (exclusion of defunct or under-performing funds); (iii) limitation in the sample size of data; (iv) frequency of observations; (v) investability of hedge funds and (vi) varying methodologies in constructing indexes.

Bibliography

Agarwal, V. and N. Naik, 2004, "Risks and Portfolio Decisions Involving Hedge Funds," *Review of Financial Studies* 17, 63-98.

Billio, M, Getmansky, M. and L. Pelizzon, 2007 "Dynamic Risk Exposure in Hedge Funds," UMASS Amherst Working Paper.

Boyson, N., Stahel, C. and R. Stulz, 2006, "Is There Hedge Fund Contagion?," Northeastern University Working Paper.

Bowman A. W., Azzalini A., 1997, "Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-PLUS Illustrations," Oxford University Press, Oxford.

Brooks, C. and H. M. Kat, 2001, "The Statistical Properties of Hedge Fund Index Returns and Their Implications for Investors," University of Reading and City University, London Working Paper.

Campbell, J.Y., Lo, A. And C. MacKinlay, 1997, "The Econometrics of Financial Markets," Princeton, NJ: Princeton University Press.

Fung, W. and D. Hsieh, 1997, "Empirical Characteristics f Dynamic Trading Strategies: The Case of Hedge Funds," *Review of Financial Studies* 10, 275-302.

Getmansky, M., Lo, A. and I. Makarov, 2004, "An Econometric Analysis of Serial Correlation and Illiquidity in Hedge-Fund Returns," *Journal of Financial Economics* 74, 3, 529-610.

Li, Y. and H. Kazemi, 2006, "Conditional Properties of Hedge Funds: Evidence from Daily Returns," UMASS Amherst Working Paper.

Appendix

Strategy Descriptions:

Systematic Trading utilizes computer models, mainly based on technical analysis of market data or fundamental economic data, to identify and make trades, with limited manager intervention.

Discretionary Trading seeks to opportunistically participate in market-driven price actions. The final decision about trading is made at the discretion of the fund manager.

Long Bias seeks to maintain a net long exposure to the market.

Variable Bias seeks to be more opportunistic about net market exposure, and has no intention to remain neutral or to maintain a particular directional bias.

Equity Hedge includes hedge funds that consist of a care holding of long equities hedged at all times with tactical short sales of stocks and/or stock index options. In addition to equities, some hedge funds many have limited assets invested in other types of securities.

Non-Directional Strategies includes convertible arbitrage, equity arbitrage and fixed income relative value strategies.

Equity Arbitrage is a market neutral strategy that seeks to profit by exploiting pricing inefficiencies between related equity securities, neutralizing exposure to directional market risk by combining long and short positions in broadly equal amounts.

Relative Value Arbitrage is a multiple investment strategy approach. The overall emphasis is on making "spread trades" which derive returns from the relationship between two related securities rather than from the direction of the market.

Absolute Return is a strategy that seeks stable performance regardless of market conditions.

Market Directional (also known as Directional Strategies) includes hedge funds that add value by participating in moves in the financial markets.

Event Driven Multi-Strategy has an objective to provide an estimate of the rate of return to event driven multi-strategy managers who attempt to seek to capitalize on investment opportunities relating to specific corporate events, such as spin-offs and restructurings, and other potential firm based actions.

Convertible Arbitrage managers seek to profit from investments in convertible securities employing both single security and portfolio hedging strategies. Managers typically build long positions of convertible and other equity hybrid securities and then hedge the equity component of the long securities positions by shorting the underlying stock or options of that company. Interest rate, volatility and credit hedges may also be employed. Hedge ratios need to be adjusted as markets move and positions are typically designed with the objective of creating profit opportunities irrespective of market moves.

Dedicated Short Bias managers seek to profit from maintaining overall net short portfolios of long and short equities. Detailed individual company research typically forms the core alpha generation driver of short bias managers, and a focus on companies with weak cash flow generation is common. Risk management consists of offsetting long positions and stop-loss strategies. The fact that money losing short positions grow in size for a short bias manager makes risk management challenging.

Emerging Markets managers seek to profit from investments in currencies, debt instruments, equities and other instruments of "emerging" markets countries (typically measured by GDP per capita). Emerging Markets include countries in Latin America, Eastern Europe, Africa, and Asia. There are a number of sub-sectors, including arbitrage, credit and event driven, fixed income bias, and equity bias.

Equity Market Neutral managers seek to profit from exploiting pricing relationships between different equities or related securities while typically hedging exposure to overall equity market moves. There are a number of sub-sectors including statistical arbitrage, quantitative long/short, fundamental long/short and index arbitrage. Managers often apply leverage to enhance returns.

Event Driven managers seek to profit from the potential mispricing of corporate securities. There is a wide range of sub-sectors within the Event Driven sector with a common theme of corporate activity or creditworthiness. Sub-sectors include mergers and acquisitions; special situations equity trading, distressed investing and credit oriented trading. Many managers use a combination of strategies; adjusting exposures based upon the opportunity sets in each sub-sector.

Distressed/High Yield Securities (also known as Distressed Security and Opportunities) fund managers in this non-traditional strategy invest in the debt, equity or trade claims of companies in financial distress or already in default. The securities of companies in distressed or defaulted situations typically trade at substantial discounts to par value due to difficulties in analyzing a proper value for such securities, lack of street coverage, or simply an inability on behalf of traditional investors to accurately value such claims or direct their legal interests during restructuring proceedings. Various strategies have been developed by which investors may take hedged or outright short positions in such claims, although this asset class is in general a long-only strategy.

Risk (Merger) Arbitrage invests simultaneously long and short in the companies involved in a merger or acquisition. Risk arbitrageurs are typically long the stock of the company being acquired and short the stock of the acquirer. By shorting the stock of the acquirer, the manager hedges out market risk, and isolates his exposure to the outcome of the announced deal. In cash deals, the manager needs only long the acquired

company. The principal risk is deal risk, should the deal fail to close. Risk arbitrageurs also often invest in equity restructurings such as spin-offs or 'stub trades'.

Long/Short Equity (also known as Equity Hedge) managers seek to profit from investing on both the long and short sides of equity markets. Managers have the ability to shift from value to growth, from small to medium to large capitalization stocks, and from net long to net short. Managers can change their exposures from net long to net short or market neutral at times. In addition to equities, long/short managers can trade equity futures and options as well as equity related securities and debt. Manager focus may be global, regional, or sector specific, such as technology, healthcare or financials. Managers tend to build portfolios that are more concentrated than traditional long-only equity funds.

Multi-Strategy managers seek to profit from allocating to a number of different strategies and adjusting their allocations based upon perceived opportunities. Many Multi-Strategy managers began as convertible arbitrage managers that diversified into other strategies. Because each strategy is not in a separate fund, these managers often have the ability to run higher leverage levels than single strategy managers.