ERASMUS UNIVERSITY ROTTERDAM

FACULTY OF ECONOMICS

BACHELOR THESIS IN FINANCIAL ECONOMETRICS

A THREE-STAGE EVALUATION OF HEDGE FUND PERFORMANCE

Denny Man

June 29, 2012

ABSTRACT

In this paper, I evaluate the performance of 955 hedge funds for the period January 2002 to December 2011, with use of four performance measures: the Sharpe ratio, Sortino ratio, Omega ratio and Jensen's alpha. The study is carried out in three stages. First, I assess whether hedge funds outperform the market, as proxied by the S&P 500. All performance measures suggest that approximately 84% of the hedge funds outperform the benchmark. The second stage investigates how the choice of performance measures impacts hedge fund rankings. While the different ratio's exhibit extremely high rank correlations mutually, they show lower rank correlation with Jensen's alpha. This deviates from previous work, though it is shown that the results are severely influenced by the choice of the sample period. In the third stage I determine significant performance persistency with use of two non-parametric tests. This means that some predictive power with regards to a fund's future performance is present.

INTRODUCTION

Since the inception of the first hedge fund, formed by Alfred W. Jones in 1949, hedge fund popularity has risen substantially. Especially during the late 1990s, when more readily available data led to an increased interest from high net worth individuals as well as institutional investors. According to recent numbers, the estimated size of the global hedge fund industry has risen to US$ 1.9 trillion, as opposed to US$ 50 billion in 1990. During that same period, the amount of individual hedge funds ascended from 500 to well over 9500.

In the past, potential hedge fund investors were attracted mostly by the promised superior returns. The common belief was that hedge fund managers' vast experience and proven investment skills would be reflected by superior returns. Many high net worth investors proved sensitive to these promises and they were responsible for much of the early growth of the hedge fund industry. However, near the end of the 1990s, disappointing hedge fund performances led to a change in sales tactics. More and more emphasis was put on the diversification properties of hedge funds. The typically low correlation with traditional asset classes (stocks and bonds), make hedge funds excellent portfolio diversifiers. This change of heart, together with the disappointing returns, led to the key question whether hedge funds still truly outperform other investment methods. Recent data from Morningstar shows that 79% of large-cap fund managers underperformed the S&P 500 index in 2011. Similar findings were reported in (Liang, 2001), who studied hedge fund performance from 1990 to mid 1999. In this period, hedge funds had an annual return of 14.2% compared with 18.8% for the S&P 500 Index, although the S&P 500 was much more volatile. These numbers make it interesting to reevaluate present day hedge fund performance. However, simply comparing hedge fund returns to benchmark returns is not sufficient, as it does not take into account the risk that was taken to achieve those returns. Therefore, a variety of measures have been developed that consider both the historical returns and the corresponding risk to assess hedge fund performance.

The most commonly used performance measures are the Sharpe ratio and Jensen's alpha. The Sharpe ratio, introduced by William F. Sharpe in 1966, considers the asset return in excess of the risk free rate per unit of deviation[1]. Originally, the Sharpe ratio assumed the risk free rate to be a constant. Later, a revised version of the performance metric acknowledged that the risk free rate changes with time (Sharpe, 1994). It has been applied in a respectable amount of studies concerning hedge fund performance evaluations (see e.g. (Brown, Goetzmann, & Ibbotson, 1999), (Liang, 1999) and (Edwards & Caglayan, 2000)). Invariably, the consensus is that hedge funds do indeed provide their investors with superior returns. The Jensen measure, most often referred to as Jensen's alpha, was first utilized by Michael Jensen in 1968 as performance measure for mutual funds. It is defined as the intercept derived from a regression of fund returns over a number of factors. Initially, Jensen's alpha

[1] For a more detailed description, see Section 2.I.1. or (Sharpe 1966, 1994)

was obtained through a single-factor model, in which only the excess market returns, as predicted by the Capital Asset Pricing Model (CAPM), were considered as a factor. In an attempt to improve on this model, multi-factor models have been applied (e.g. (Fung & Hsieh, 1997a), (Kat & Miffre, 2006) and (Brown, Fraser, & Liang, 2008)). Similarly as with the Sharpe ratio, researchers find that hedge funds routinely outperform the benchmark according to Jensen's alpha.

However, both performance metrics suffer from some drawbacks. Numerous studies have pointed out that hedge fund returns are typically not normally distributed. This in turn can lead to severe overestimations of the aforementioned performance measures (see e.g. (Brooks & Kat, 2001), (Fung & Hsieh, 2001) and (Agarwal & Naik, 2004)). With regards to factor models, (Kat & Palaro, 2006) comment that hedge fund returns may derive from exposure to highly unusual risk factors. Incorrectly specified or even completely omitted risk factors tend to show up in alpha, thereby implying superior performance where there in fact may be none. In addition, (Kat & Miffre, 2006) state that factor models tend to exhibit low coefficients of determination R^2, suggesting that multi-factor models have difficulty capturing the complex and highly opportunistic nature of hedge fund strategies. Furthermore, monthly hedge fund returns show significant serial correlation, which may lead one to seriously underestimate true volatility and correlation (Brooks & Kat, 2001). Moreover, (Kosowski, Naik, & Teo, 2005) and (Gregoriou & Zhu, 2005) underline that the dynamic trading strategies unique to hedge funds often challenge the assumption of constant coefficients.

In light of the previous, a considerable amount of alternative performance measures have been developed. These include the modified Sharpe ratio, conditional Sharpe ratio, Treynor ratio, Omega ratio, the Sortino ratio, the Calmar ratio, the Sterling ratio and the Burke ratio. Nonetheless, the use of these performance measures has been under debate as well. As a result, researchers have been creative in finding new ways to evaluate hedge fund performance. (Amin & Kat, 2001) suggest a so-called efficiency test, which does not require any assumptions about the hedge fund return distribution. They conclude that hedge funds do not offer a superior risk-return profile. (Murthi, Choi, & Desai, 1997) employ a technique called data envelopment analysis (DEA) on mutual funds, their research was extended to hedge funds by (Gregoriou & Zhu, 2005). Moreover, (Kosowski, Naik, & Teo, 2005) use a robust bootstrap procedure with a Bayesian approach to overcome the short-sample problem inherent in hedge fund returns. They report that hedge funds provide added value that cannot be explained by luck. In contrast with (Kat & Palaro, 2007), who find no superior performance for the majority of the hedge funds, using a replication method that involves generating returns with the same statistical properties as the returns generated by the hedge fund using traditional asset classes. They argue if the fund returns can be replicated, they cannot be superior and *vice versa.* Surely, the contradicting results solicit further research.

Besides the issue of superior returns, researchers and practitioners are also interested in whether a certain fund outperforms others, i.e. a fund's relative performance. To this extent, it is common in literature to compute the hedge funds' Spearman rank correlation coefficients. This involves assigning a rank to each individual hedge fund based on their performance and calculating the correlation coefficient accordingly. (Eling & Schuhmacher, 2007) compared the Sharpe ratio to 12 other performance measures for a sample of 2,763 hedge funds. (Nguyen-Thi-Thanh, 2008) performed a comparative study on 149 Long/Short Equity funds. Both papers report that all the measures produce virtually identical fund rankings. (Nalpas, Vanhems, & Simar, 2011) compare free disposal hull (FDH) directional measures with DEA- and parametric measures. Although mostly significant, they find lower Spearman rank correlations. Moreover, hedge funds can be categorized based on their investment strategies. Although there exists no formal set of definitions with which to classify the hedge funds, data vendors typically divide hedge funds into a number of broad investment categories. As each strategy implies different expected returns, volatility and risk, it is interesting to investigate whether a certain strategy outperforms others.

Another point of interest is the persistence of the performance measures. In hedge fund analysis, performance persistency is described as the relation between a fund's performance during a defined first period and its performance during the next. If a hedge fund exhibits significant performance persistency, it means that some predictive power with regards to the fund's future performance is present. The most prominent method in literature to determine the presence of persistency is a comparison of Spearman rank correlations between two specified sub-periods. In addition, (Nguyen-Thi-Thanh, 2008) suggests a contingency table-based test. As is often the case in hedge fund analysis, existing literature is inconclusive. (Agarwal & Naik, 2000), (Edwards & Caglayan, 2001) and (Nguyen-Thi-Thanh, 2008) found significant persistency in their studies while (Brown, Goetzmann, & Ibbotson, 1997) and (Schneeweis, Kazemi, & Martin, 2001) did not.

This study is carried out in three stages. First, I assess whether hedge funds deliver superior returns by comparing the hedge fund performances against a market benchmark. Following (Amin & Kat, 2001), the S&P 500 will serve as the market proxy. Second, I investigate how the choice of the performance measure affects, if at all, relative hedge fund performance, i.e. hedge fund rankings. In addition, I research whether a specific investment strategy performs better than others. Third, I test for the stability of the performance measures. That is, examine the presence or absence of performance persistency. This paper contributes to existing literature in the following ways. In contrast to (Schneeweis, Kazemi, & Martin, 2001), I concentrate on persistency in the performance measures, as opposed to persistency in the fund returns. I extend the work of (Nguyen-Thi-Thanh, 2008), who exclusively focused on 149 Long Short Equity funds, while I consider 955 hedge funds across all investment categories. In contrast to (Gajda, 2009), I calculate the performance measures for all individual hedge funds separately instead of averaging the returns within a particular group, so monthly returns are not flattened. Finally, in contrast to many papers, I correct for biases that are inherently present in hedge fund data.

The remainder of this paper is organized as follows. Section 1 describes the data along with several well known biases that tend to appear in hedge fund data. Section 2 presents the methodology employed in this paper, while Section 3 reports the empirical results. They are divided into three subsections, dealing with the performance measures against the benchmark, hedge fund rankings and performance persistency, respectively. Section 4 contains concluding remarks and offers suggestions for further research.

1. DATA

To gather reliable hedge fund data, one has to overcome a number of hurdles. There is no central data bank that provides generally accepted and standardized information on the hedge fund universe. As a result, only a fraction of the hedge fund industry is observable[2]. Since hedge funds are essentially prohibited from advertising[3], they voluntarily report their performance data to a data vendor in order to reach potential investors. However, fund managers typically disclose information to only a single data vendor (Agarwal, Daniel, & Naik, 2004). Consequently, large differences can be found between data from different vendors. To elaborate, (Liang, 2000) compared databases from Hedge Fund Research (HFR) and Lipper Trading Advisor Selection System (TASS). Out of the 1162 and 1627 hedge funds that were included in the databases, respectively, only 465 hedge funds were reported in both. Similar conclusions were reached by (Gregoriou, 2002) and (Agarwal, Daniel, & Naik, 2004), who extended Liang's research to a comparison of the HFR, TASS and Zurich Capital Markets (ZCM) database. Clearly, the largely different databases impose some problems. As far as investors are concerned, the scattered hedge fund data make it more difficult to find the most beneficial investment opportunity. From an academic viewpoint, a researcher typically has access to one specific database, which is often not the same database as other researchers have used. Hence, studies are more difficult to reproduce.

In addition, because the decision to report performance data is entirely voluntary for hedge funds, a strategic element is added to the reporting process. The resulting biases, as discussed below, lead to an overstatement of performance, which contributes to the difficulty investors have in evaluating the funds[4].

[2] W. Fung, D.A. Hsieh, *Hedge-Fund Benchmark: Information Content and Biases,* Financial Analysts Journal, January/February 2002, p.23.
[3] Individual funds are allowed to disclose their own data, but not comparative data.
[4] G. Bhardwaj, *Alternative Investments Versus Indexing: An Evaluation of Hedge Fund Performance,* Vanguard Investment Counseling & Research, 2010, p.3.

Survivorship bias: Survivorship bias arises when a data vendor removes performance data from funds that have stopped reporting, i.e. defunct funds. Oftentimes this is due to liquidations, mergers or manager discretion, which tends to occur when a fund performs poorly. Thus, surviving funds have generally outperformed the defunct funds, causing an upward bias in the average returns. Several authors have estimated survivorship biases across different databases. (Liang, 2000) found an annual bias of 0.31% in the HFR database, as opposed to 1.49% in the TASS database. He continued his work in (Liang, 2001) and estimated the survivorship bias in the TASS database to be 1.69% annually for the period 1990 to 1999, and 2.43% for the period 1994 to 1999. Similar numbers were found by (Fung & Hsieh, 2000). (Schneeweis, Kazemi, & Martin, 2001) report that the impacts of survivorship bias can greatly vary over time. Evidently, hedge fund performance evaluation is strongly affected by the choice of the sample period and database.

Backfill bias: Also known as instant history bias, refers to when a data vendor includes a hedge fund that has already been operating for a while. Usually, fund managers are allowed to backfill their historical returns. Since inclusion in the database is often for the purpose of advertisement, managers will be reluctant to backfill data if the fund performed poorly. Consequently, only funds that have performed well will backfill data, resulting in a positive bias. (Fung & Hsieh, 2000) estimated the instant history bias of the TASS database to be 1.4% annually for the period 1994 to 1998.

Selection bias: Selection bias emerges due to the voluntary nature of inclusion in the databases. Presumably, only funds that have performed well wish to be included in the database. Therefore, hedge funds included in the database might not properly represent the complete hedge fund universe, supposedly leading to an upward bias. However, (Fung & Hsieh 1997a, 2000) point out anecdotal evidence that suggests selection bias could be limited. Managers with superior performance did not necessarily contribute to a database. For example, George Soros's Quantum Fund did not disclose information to data vendors, despite their outstanding performance record. This indicates that there are offsetting factors at work, limiting the magnitude of the selection bias.

For the empirical investigation, I use data obtained from the Lipper TASS. As of February 2012, the database contains monthly returns for 1143 hedge funds, divided into 1039 (90.9%) *live* funds and 104 (9.1%) *graveyard* funds, for the sample period January 2002 to December 2011 (120 observations). Live funds indicate hedge funds that currently still report performance data on a monthly basis, whereas graveyard funds (i.e. defunct funds) are funds that have stopped reporting[5]. The sample period is chosen in a way that allows for a relatively large amount of hedge funds to fit the selection criteria, while still maintaining a long enough horizon to conduct the sub-period analysis. Additionally, it

[5] As mentioned before, this can be due to any number of reasons, for example liquidations or when investment capacities are reached..

keeps in mind the study objective of performance evaluation of *present day* hedge funds, by only considering relatively recent data. This explains the small proportion of graveyard funds.

Because the Lipper TASS database includes performance data both on live funds as well as defunct funds, it is possible to determine and mitigate the survivorship bias, which is calculated as the difference in hedge fund returns between all funds and merely the live funds. Furthermore, the database provides additional information on each fund, such as assets under management, strategy description and inception date. The latter allows us to account for backfill bias, by only including observations of a fund after its inception date. It is not possible to remove selection bias from the data[6]. As a last criterion, I only consider funds that have consistently, without a break, reported monthly data during the entire study period. This leaves a total of 955 hedge funds, consisting of 917 (96.0%) live and 38 (4.0%) graveyard hedge funds.

Since the performance measures strongly rely on normality of the hedge fund returns, I conduct a Jarque-Bera test on all 955 funds. Table 1 summarizes the results, along with the mean of the first four moments of the return distribution: mean value, standard deviation, skewness and kurtosis. The first reveals that live funds generally exhibit higher mean returns than graveyard funds, taking the value (with standard error in parentheses) 0.649 (3.194) compared to 0.582 (3.453). Survivorship bias is extremely small, with an estimation of 0.003%. This is not surprising, given the low proportion of graveyard funds in the sample. Furthermore, both live- and graveyard funds show negative skewness of approximately -0.6 and are leptokurtic with average kurtosis estimated at 7.684 and 8.245 respectively. The final column brings to light that normality is rejected for 86.9% of the live funds and 76.3% of the graveyard funds, supporting findings by (Nguyen-Thi-Thanh, 2008), who argued that longer sample horizons lead to a substantially higher proportion of funds that fail the null hypothesis of normality.

Table 1: Summary statistics and test of normality for hedge fund returns distribution

This table reports the distributional properties of the hedge fund returns. Live funds indicate hedge funds that currently still report performance data on a monthly basis, whereas graveyard funds (i.e. defunct funds) are funds that have stopped reporting. Columns one to four report the first four moments of the return distribution, while column five reveals the percentage of funds for which the null hypothesis of normality is rejected.

	Nr. of funds	Mean [%]	Std. Dev. [%]	Skewness	Kurtosis	% of funds with Jarque-Bera $p < 0.1$
Live funds	917	0.649	3.194	-0.590	7.684	86.9%
Graveyard funds	38	0.582	3.453	-0.573	8.245	76.3%
All funds	955	0.646	3.204	-0.588	7.706	86.5%

[6] Fung & Hsiesh (2000) argue this can only be done with the input from investors of funds that generally do not disclose their data, which is unlikely to be achieved by academics alone.

2. METHODOLOGY

The purpose of this section is to describe the methods employed in this paper to evaluate hedge fund performance. It is divided into three subsections. The first subsection defines the performance metrics that are considered in this paper. These are the Sharpe ratio, Sortino ratio, Omega ratio and Jensen's alpha. The second subsection deals with the approach used to determine the impact of the choice of performance measures on relative hedge fund performance, i.e. hedge fund rankings. In addition, an outline of the various hedge fund strategies is given to allow for analysis on the relative performance of hedge funds in different categories. The third subsection conveys the tests that are used to determine the presence of performance persistency. More precisely, I discuss the contingency table-based test and the Spearman rank correlation coefficients test.

2.I. THE PERFORMANCE MEASURES

This subsection presents the performance measures. To determine whether hedge funds outperform the market, I will compute the measures for the 955 hedge funds and compare it to the performance measures of a benchmark. Following (Amin & Kat, 2001), the S&P 500 will serve this purpose. It is important to note however, that theoretically the S&P 500 should only be used as a market proxy to evaluate fund managers that trade in S&P 500 stocks exclusively. However, given the study objective of this paper and the widely varying hedge fund trading strategies, the S&P 500 suffices, at least from a practical standpoint.

2.I.1. SHARPE RATIO

In hedge fund analysis, the Sharpe ratio is the most commonly applied performance metric, despite its potential shortcomings. In studies, it is usually computed in order to compare hedge fund evaluations of other performance metrics to. Its use has been heavily criticized by a multitude of authors (e.g. (Kat & Palaro, 2007), (Brooks & Kat, 2001)), who argue that the non-normal distribution of hedge fund returns causes substantial overestimation of the Sharpe ratio. However, (Eling & Schuhmacher, 2007) put forth that the comparison of the Sharpe ratio to other performance measures resulted in virtually identical rank ordering across hedge funds, concluding that the use of the Sharpe ratio is justified, at least from a practical perspective. The Sharpe ratio calculates the risk premium per unit of deviation and is defined as:

$$Sharpe\ Ratio_i = \frac{r_i^d - r_f}{\sigma_i} \tag{1}$$

where $r_i^d = \frac{r_{i1}+\dots+r_{iT}}{T}$, represents the average monthly return for hedge fund i over period T and σ_i is the standard deviation of the monthly return r_{it}. The risk-free monthly interest rate r_f is represented by the three-month US Treasury bill rate[7].

2.I.2 SORTINO RATIO

One of the drawbacks of the Sharpe ratio is that it penalizes both upside and downside volatility equally. Typically, the upside volatility is not an unwanted treat and thus should not be penalized. To this extent, the Sortino ratio (Sortino & Price, 1994) provides a solution to differentiate between the good and bad volatility. The so-called downside risk measure is closely related to the Sharpe ratio and depends on the lower partial moments (LPM) function, which we will see is also the case for the Omega ratio, albeit of different order. The LPM of order n for hedge fund i is given by

$$LPM_{ni}(\tau) = \frac{1}{T}\sum_{t=1}^{T} \max[\tau - r_{it}, 0]^n \tag{2}$$

where τ denotes the minimal acceptable return, i.e. threshold return, T denotes the returns horizon and $n = \{0, 1, 2, \dots\}$. From (2) it should be clear that this represents the same standard deviation as in (1), but now only considering the negative deviations of the average return if we set $\tau = r_f$. The Sortino ratio uses a LPM of order $n = 2$ and is calculated as follows:

$$Sortino\ Ratio_i = \frac{r_i^d - \tau}{\sqrt{LPM_{2i}(\tau)}} \tag{3}$$

where r_i^d is the average monthly return of hedge fund i. Consequently, if $(r_i^d - \tau) > 0$, the Sortino ratio will provide a larger number than the Sharpe ratio, since we are dividing by a smaller number. For the same reason, if $(r_i^d - \tau) < 0$ the Sortino ratio penalizes the underperforming hedge funds more harshly. By definition, the Sortino ratio has the same sign (+/-) as the Sharpe ratio.

2.I.3. OMEGA RATIO

The Omega ratio, proposed by (Keatwick & Shadon, 2002) is a measure that does not rely on the assumption of a normal return distribution. They define it as follows:

$$\Omega(\tau) = \frac{\int_{\tau}^{\infty}[1 - F(r)]dr}{\int_{-\infty}^{\tau} F(r)dr} \tag{4}$$

[7] Treasury bills are considered to be (practically) risk-free, because they are backed by the U.S. governments.

Where $F(r)$ denotes the cumulative distribution function of the returns and τ is the threshold return. From (4) it should be clear that Omega represents the ratio of the probability of getting a return above the threshold against the probability of getting a return below the threshold. As opposed to the Sharpe ratio, the Omega ratio considers all moments of the distribution. (Kaplan & Knowles, 2004) derived an equivalent representation of (4) based on the LPM function as in (2) of order $n = 1$:

$$\Omega(\tau) = \frac{r_i^d - \tau}{LPM_{1i}(\tau)} + 1 \tag{5}$$

It is helpful to remember that the Sortino ratio only differs from the Sharpe ratio in magnitude and not in sign (+/-). A similar result can be derived for $\Omega(\tau) - 1$. Therefore, if a fund shows a positive Sharpe ratio, it will also have a Sortino ratio larger than zero and an Omega ratio larger than one.

2.I.4. JENSEN's ALPHA

A classic performance measure, with its origins dating back to 1968, Jensen's alpha is still widely being used. It was originally developed as a measure to evaluate the performance of mutual fund managers. Present day, Jensen's alpha is commonly applied in a variety of empirical finance researches, including hedge fund performance evaluations. It considers the abnormal return of a hedge fund over the theoretically expected return, as predicted by the Capital Asset Pricing Model (CAPM), or rather:

$$E(r_{CAPM}) = r_f + \beta_i(r_m - r_f) \tag{6}$$

where r_{CAPM} is the theoretical expected return, r_f the risk-free rate, β_i is a measure of an asset's volatility in relation to the market and r_m is the market return, represented by the S&P 500 index. The regression to obtain Jensen's alpha may now be written as:

$$\hat{\alpha}_i = r_i - \left[r_f + \hat{\beta}_i(r_m - r_f)\right] + e_i \tag{7}$$

in which $\hat{\alpha}_i$ is Jensen's alpha of hedge fund i, r_i is the hedge fund return, r_f the average risk-free rate and r_m the market return, $\hat{\beta}_i$ the regression coefficient and e_i the error. The intercept $\hat{\alpha}_i$ is interpreted as the value added by hedge fund managers; it is the return that cannot be explained by the independent variables.

2.II. HEDGE FUND RANKINGS

The adequateness of the Sharpe ratio as a hedge fund performance measure has been contested by many authors. The main argument is that non-normality of the hedge fund returns lead to a systematic overestimation of the ratio. However, if this overestimation occurs for all hedge funds, it might be an appropriate metric to measure the *relative* performance of the funds. Existing literature[8] favors this standpoint. They report that hedge fund rankings are approximately the same, regardless of the choice of the performance measure. To build on previous work, I will begin further analysis of the predefined measures by calculating the Spearman rank correlation coefficients. This entails assigning a rank to each hedge fund on the basis of their performance measure[9]. Then, the Spearman rank correlations are computed as follows

$$r_s = 1 - 6\sum_{i=1}^{n} d_i^2/(n^3 - n) \tag{8}$$

where n is the number of hedge funds and d_i is the ranked difference of hedge fund i. To test the significance of the correlation coefficients, a t-test is appropriate[10]:

$$t = r_s/\sqrt{(1 - r_s^2)/(n - 2)} \tag{9}$$

where r_s is the Spearman rank correlation coefficient and n the total number of hedge funds. The t-statistic follows a t-distribution with $n - 2$ degrees of freedom.

2.II.1. HEDGE FUND STRATEGIES

Hedge fund managers have a wide array of strategies at their disposal, that tend to be quite different from each other. Hence, hedge funds should be treated as a heterogeneous group. To do so, hedge funds are categorized with respects to their trading strategy. However, there exists no formal set of class definitions with which to classify the hedge funds. Nonetheless, data vendors usually divide hedge funds into a number of broad investment categories. Although these categories can vary across different databases, there are some commonly used classifications. Likewise, the Lipper TASS database has the 13 following hedge fund strategies:

Long/Short Equity	Managed Futures	Fund of Funds
Equity Market Neutral	Multi-Strategy	Event Driven
Global Macro	Emerging Markets	Fixed Income Arbitrage
Dedicated Short Bias	Convertible Arbitrage	Options Strategy
Other		

Note: *A more detailed description of each strategy is given in Appendix A.*

[8] See, for example, Eling & Schuhmacher (2007) or Nguyen-Tih-Thanh (2008)
[9] Tied values are assigned a rank equal to the average of their positions.
[10] See J.H. Zar, *Significance Testing of the Spearman Rank Correlation Coefficient*, Journal of the American Statistical Association, September 1972, p.578.

Each strategy imposes different expected returns, volatility and risk. Some strategies that are not correlated to equity markets can yield consistent returns with low risk, while others can be more volatile than mutual funds. For an investor, it is interesting to determine whether a certain strategy performs significantly better relative to other strategies. To examine this, I compute and compare the performance measures across the different investment objectives. Existing literature often calculates the performance measures by taking the average return of all hedge funds within a particular group. However, this causes the monthly returns to be flattened, which in turn leads to an underestimation of the standard deviation. Hence, the performance measures will be overestimated. To mitigate this problem, I will estimate the performance measures of each hedge fund within a category separately and take averages of the computed measures, not the returns. This provides more reliable results and gains in interpretative value.

2.III. PERFORMANCE PERSISTENCY

In hedge fund analysis, performance persistency is described as the relation between a fund's performance during a defined first period and its performance during the next. If a hedge fund exhibits significant performance persistency, it means that some predictive power with regards to the fund's future performance is present. To assess the presence of persistency, one can consider two types of performance: absolute performance and relative performance. The first refers to the exact value of the chosen performance measure, while the latter corresponds with the relative value of the performance measure, or rather, the ranking in relation with other hedge funds. As an extension of the previous section, this study focuses on the latter. In order to do so, I will divide the study period into two equal sub-periods of five years[11] and compute the performance measures accordingly.

To statistically determine the presence of persistence, I use two non-parametric tests, following (Nguyen-Thi-Thanh, 2008): the contingency table-based test and the Spearman rank correlation test. These tests are suggested because of their conceptual simplicity, their facility in application and the absence of econometric biases which involve parametric tests.

[11] First sub-period: Jan. 2002 – Dec. 2006, Second sub-period: Jan. 2007 – Dec. 2011

2.III.1 CONTINGENCY TABLE-BASED TEST

For the contingency table-based test, I divide the hedge funds into quartile based on their relative performance in the first period. Then, the following contingency table is formed:

Table 2: Example of a 4x4 contingency table for the performance persistency test

This table shows an example of a 4x4 contingency table. Hedge funds are divided into four quartiles (Q1 to Q4) for two sub-periods. Within each cell, n_{QxQy} denotes the number of funds that moved from Qx in the first sub-period to Qy in the second sub-period, for $x, y = 1,..,4$.

Sub-period 1	Sub-period 2			
	Q1	Q2	Q3	Q4
Q1	n_{Q1Q1}	n_{Q1Q2}	n_{Q1Q3}	n_{Q1Q4}
Q2	n_{Q2Q1}	n_{Q2Q2}	n_{Q2Q3}	n_{Q2Q4}
Q3	n_{Q3Q1}	n_{Q3Q2}	n_{Q3Q3}	n_{Q3Q4}
Q4	n_{Q4Q1}	n_{Q4Q2}	n_{Q4Q3}	n_{Q4Q4}

Hedge funds are divided into quartile one (denoted $Q1$) to quartile four (denoted $Q4$) based on their performance measure during a sub-period. The best performing funds are assigned to Q1, while the worst performing funds will be in Q4. Given that the sample consists of 955 hedge funds, the first three quartiles will contain 239 observations, while Q4 will have 238 observations. Within each cell, n_{QxQy} denotes the number of funds that moved from quartile Qx in the first sub-period to quartile Qy in the second sub-period, for $x, y = 1,..,4$. To statistically test the presence or absence of performance persistence, I will use the Chi-square statistic as suggested by Kahn and Rudd (1995):

$$\chi^2 = \sum_{i=1}^{l}\sum_{j=1}^{c}\frac{\left(\hat{n}_{ij} - n_{ij}\right)}{n_{ij}} \qquad \sim\chi^2_{(l-1)(c-1)} \tag{10}$$

where $\hat{n}_{ij}$ represents the observed frequency in the corresponding contingency table cell ij for $i, j = Q1, \ldots, Q4$ and n_{ij} denotes the theoretical frequency under the null hypothesis of no persistence. The degrees of freedom are $(l-1)(c-1)$, with l the number of rows and c the number of columns, both equal to four in our case.

2.III.2 SPEARMAN RANK CORRELATION COEFFICIENTS TEST

The choice of the number of quantiles in the contingency table-based test is somewhat arbitrary. One could argue that four quartiles for a large sample allows for too much movement within a quartile. A method that more closely monitors the individual hedge fund movement might provide additional insight. For this purpose, a Spearman rank correlation coefficients test is proposed. As also discussed in Section 2.II., this test involves assigning a rank to each individual hedge fund for both sub-periods, based on their performance. Then, the difference between the ranks, if any, is determined and finally, the Spearman rank correlation coefficient is computed:

$$r_s = 1 - 6\sum_{i=1}^{n} d_i^2/(n^3 - n) \tag{11}$$

where n now denotes the number of observations in each sub-period and d_i is the ranked difference of hedge fund i. A higher correlation means more similar rankings between both periods. Due to their nature, the contingency table-based test could be hinted at as a test on semi-strong persistence, while the Spearman rank correlation test is a test of strong form. Additionally, a t-test as in (9) is performed to determine the significance of the coefficients.

3. RESULTS

This section presents my findings using the methodology described in Section 2. First, I discuss results regarding the Sharpe ratio, Sortino ratio and Omega ratio, followed by discussion on Jensen's alpha. I compare the measures mutually for the 955 hedge funds and against the benchmark S&P 500. Second, I address how the choice of performance measures can affect hedge fund rankings. In addition an overview and interpretation of the average fund performance within a particular investment category is given. Third, I examine the results of the contingency table-based test and the Spearman rank correlation test on the presence or absence of performance persistency.

The results of the analysis can be summarized as follows: According to the Sharpe ratio, Sortino ratio and Omega ratio, the vast majority (93.4%) of the hedge funds systematically deliver positive net-of-risk-free rate returns. The average Sharpe-, Sortino- and Omega ratio are (with standard error in parentheses) 0.253 (0.445), 0.384 (0.946) and 1.522 (0.919) respectively. Jensen's alpha is equal to 0.485 (0.251) on average and is significant for 54.8% of the hedge funds. Inspection of the residuals reveals that the vast majority of the funds fail the tests for normality (75.7%), homoskedasticity (50.6%) and absence of first-order autocorrelation (44.0%). In comparison with the S&P 500, hedge funds outperformed the benchmark according to all measures. With regards to hedge fund rankings, the Sharpe-, Sortino-, and Omega ratio exhibit extremely high rank correlations mutually (0.99 on average), indicating that these measures lead to nearly identical fund rankings. However, they show somewhat lower rank correlation with Jensen's alpha (0.78 on average). This is explained by the chosen sample period. In terms of investment categories, the performance measures generally provide different rankings. Nonetheless, some common preferences are observed. Multi-Strategy funds performs relatively well across all measures, while Long/Short Equity funds tend to perform poorly. Furthermore, better performing strategies characterize with relatively high standard deviations, supporting the economic theory that higher returns imply higher volatility. On performance persistency, both the contingency table-based test and the Spearman rank correlation test determine the presence of significant persistency. This implies that some predictive power is present with regards to a fund's future performance.

3.I. BENCHMARK COMPARISON

For the comparison and discussion of the different performance measures, first the Sharpe ratio, Sortino ratio and Omega ratio are considered. This is due to the different nature of the fourth performance metric, Jensen's alpha. The performance measures are computed for the 955 hedge funds and discussed mutually. To address the issue of whether hedge funds outperform the market, the measures are computed for and compared against the benchmark S&P 500.

3.I.1 SHARPE RATIO, SORTINO RATIO, OMEGA RATIO

Table 3 shows the average Sharpe-, Sortino- and Omega ratio for the 955 hedge funds during the period January 2002 to December 2011, along with some descriptive statistics. The mean monthly hedge fund return is equal to 0.646%, and the net-of-risk-free-rate monthly returns are on average positive with a value equal to 0.495%. The average standard deviation of the hedge fund returns, denoted σ_i, is equal to 3.204. Commonly, literature reports that hedge funds mostly have positive Sharpe ratio's. The findings in this study are in line with previous work. Out of the total of 955 hedge funds, 892 (93.4%) show a positive Sharpe ratio. Consequently, the same number of funds have a Sortino ratio larger than zero and an Omega ratio larger than one. In other words, the vast majority of hedge funds delivered returns that were systematically higher than the risk free rate. Moreover, 148 (278) funds have an annualized Sharpe (Sortino) ratio[12] larger than one, a treat often looked for by investors.

The average value (with standard error in parentheses) of the Sharpe ratio is 0.253 (0.445), while the average Sortino ratio is equal to 0.384 (0.946). The mean Omega ratio is estimated at 1.522 (0.919). On average, the hedge funds performed well in comparison with the S&P 500. Table 8 (see Appendix B) reports the performance metrics of the benchmark. The Sharpe-, Sortino- and Omega ratio of the S&P 500 take values of 0.042, 0.058 and 1.119, respectively. Additionally, Table 8 provides insight into the amount of hedge funds that had higher performance scores. On average, this was equal to about 84% for all measures.

However, the use of especially the Sharpe ratio for hedge fund performance evaluation has been the subject of intense criticism. For example, (Brooks & Kat, 2001) argue that Sharpe ratio's of hedge

[12] Annualized Sharpe- or Sortino ratio's are calculated by multiplying their monthly equivalent with the square root of 12.

funds are substantially overestimated due to non-normality of the returns. Given the previous, the relatively high number of well performing hedge funds should come as no surprise. Furthermore, despite the potentially erroneous estimations of the absolute value of the performance measures, possible usefulness could be found in the field of hedge fund ranking comparisons. A multitude of studies (e.g. (Eling & Schuhmacher, 2007), (Nguyen-Thi-Thanh, 2008) or (Nalpas, Vanhems, & Simar, 2011)) find highly positive correlation coefficients between the Sharpe ratio and numerous of other performance measures, indicating that the measures are appropriate for the relative comparison between hedge fund rankings. This subject is discussed more thoroughly in Sections 2.II. and 3.II.

Table 3: Summary Statistics of Sharpe- and Sortino- and Omega ratio

This table shows the mean of the returns r_i, the net-of-risk free rate returns $r_i - r_f$ and of the returns standard deviations σ_i of the 955 hedge funds. Columns four through nine present the mean and standard errors of the respective performance measures. The sample period is January 2002 to December 2011.

Mean								
(r_i) [%]	$(r_i - r_f)$ [%]	σ_i [%]	Avg. Sharpe Ratio	Std. Err. Sharpe Ratio	Avg. Sortino Ratio	Std. Err. Sortino Ratio	Avg. Omega ratio	Std. Err. Omega ratio
0.646	0.495	3.204	0.253	0.445	0.384	0.946	1.522	0.919

3.I.2. JENSEN's ALPHA

In order to obtain Jensen's alpha, separate regressions as described in Section 2.I.4. were estimated for all hedge funds. Additionally, tests of normality, heteroskedasticity and autocorrelation were performed on the hedge fund residuals to examine the behavior of fund returns. The achieved estimates and test results, along with a number of descriptive statistics are presented in Table 4. On average, alpha is positive with a value of 0.485% per month (5.82% annually). This can be interpreted as the average value added by a hedge fund manager. However, one needs to take into account that incorrectly specified or omitted variables will generally lead to an overestimation of alpha. Therefore, some caution is advised in the interpretation of the found values. The average standard error of alpha[13] is equal to 0.251, and alpha is statistically significant for 523 out of 955 (54.76%) hedge funds at the 10% confidence level.

Inspection of the residuals reveals that over 75% of the funds fail the Jarque-Bera test for normality, and that the residuals are skewed negatively (-0.213) and exhibit high kurtosis (6.705). Furthermore, the null hypothesis of homoskedasticity is rejected for over 40% of the funds according to the Breusch-Pagan test and significant first-order autocorrelation is found in 33.2% of the funds with use of the Ljung Box test-statistic. This indicates that the estimate of alpha is likely too optimistic.

[13] Note that this not the standard error of the average alpha, but rather the average of the standard errors of each individual $\hat{\alpha}_i$.

Table 4: Summary statistics of alpha

This table reports the results from the alpha regression for the 955 hedge funds. In Panel A, the first three columns report the mean of alpha, the standard error of alpha and the t-statistic, respectively. The fourth column reveals the percentage of funds for which the null hypothesis of normality is rejected. Panel B shows the distributional properties of the hedge fund return residuals. Columns one and two report average skewness and kurtosis, while columns three to five report the percentage of fund residuals that fail the test normality, heteroskedasticity and autocorrelation at the 10% significance level, respectively. The sample period is January 2002 to December 2011.

Panel A: Descriptive statistics of alpha

Avg. alpha [%]	Avg. Standard Error [%]	Avg. T-stat	% of funds with alpha $p < 0.1$
0.485	0.251	2.95	54.76

Panel B: Distributional properties of residuals

Mean		Test of normality	Test of heteroskedasticity	Test of autocorrelation
Skewness	Kurtosis	% of funds with Jarque-Bera $p < 0.1$	% of funds with Breusch-Pagan $p < 0.1$	% of funds with Ljung-Box $p < 0.1$
-0.213	6.705	75.71	50.58	43.98

Table 9 (see Appendix B) displays a summary of the alpha regression performed on the S&P 500. The estimate of alpha is equal 0.163 and it is significant at the 1% confidence level. Thus, similarly to what we saw earlier, hedge fund managers are able to provide added value over the market. Furthermore, the residuals show high skewness (1.568) and kurtosis (23.946), leading to a strong rejection of normality in the fund residuals. Additionally, the null hypothesis of homoskedasticity and no first-order autocorrelation are strongly rejected as well. This strongly suggests that the 1-factor model captures neither the hedge fund returns nor the market returns well.

3.II. HEDGE FUNDS RANKINGS

This section provides answer to the question of how the choice of performance measures affects hedge fund rankings. In order to do so, the Spearman rank correlations as described in 2.II. are computed. Table 5 shows the results.

Interestingly, a clear distinction is observed between the Spearman rank correlations of the Sharpe ratio, Sortino ratio and Omega ratio with Jensen's alpha. While the different ratio's exhibit extremely high rank correlations mutually, they show lower rank correlation with Jensen's alpha. This follows the results of (Nalpas, Vanhems, & Simar, 2011) more so than the findings of (Eling & Schuhmacher, 2007). As we will see in Section 3.III.2, this is explained by the chosen sample period. Nonetheless, the correlation coefficients are still extremely high and statistically significant at the 1% confidence level. This implies that nearly identical fund rankings are achieved, regardless of the choice of the performance measure. Hence, one could conjecture that the Sharpe ratio appropriately captures hedge fund performances. However, rank correlations merely give insight into the relative performance of the funds and do not take into account the absolute values of the measures. Potentially, a problem could arise, for example, when an investor wants to compare hedge fund performances to mutual fund performances. If the distribution of the hedge fund returns leads to a systematic overestimation of a specific performance measure (Brooks & Kat, 2001), while the mutual fund returns do not, incorrect conclusions could be drawn. Therefore, the results from the Spearman rank correlations alone are not sufficient to disregard other performance measures. Further investigation is necessary to that extent, but is beyond the scope of this paper.

Table 5: Spearman Rank Correlation Coefficients

This table shows the Spearman rank correlation coefficients from the different performance measures for the sample period January 2002 to December 2011.

Performance Measure	Sharpe Ratio	Sortino Ratio	Omega Ratio	Jensen's alpha
Sharpe Ratio	1			
Sortino Ratio	0.991	1		
Omega Ratio	0.993	0.981	1	
Jensen's alpha	0.776	0.801	0.772	1

Note: *All values in this table are significant at the 1% confidence level.*

3.II.1. HEDGE FUND STRATEGY PERFORMANCE

In this section I examine potential differences in performance between the hedge fund strategies. The funds were broken down by the investment categories as described in Section 2.II.1. Table 10 (see Appendix C) shows the results corresponding to the Sharpe-, Sortino- and Omega ratio, while Table 11 (see Appendix D) covers Jensen's alpha.

A quick glance at Table 10 shows that the measures generally provide different strategy rankings, although some common preferences can be found. Both the Sharpe- and Sortino ratio agree on Multi-Strategy hedge funds as the best performing category. The Omega ratio puts forth Event Driven funds as top category, but has Multi-Strategy funds in second place. Long/Short Equity funds appear to perform poorly across all performance measures. It is interesting to note that better performing strategies characterize with relatively high standard deviations of the performance metric, with the exception of Emerging Markets in the Omega ratio. This reflects the more volatile nature of the strategies; e.g. a fund either performs very well or very poorly.

Table 11 showcases Jensen's alpha per investment category, along with descriptive statistics and tests of normality, heteroskedasticity and serial correlation on the residuals. The second column reveals that Jensen's alpha provides a different insight into hedge fund strategy ranking. The trio of best ranked strategies consists of Emerging Markets, Managed Futures and Global Macro, with alpha's of 1.03% (12.4%), 0.70% (8.4%), 0.67% (8.0%) per month (year) respectively. Of these top performing strategies, only Managed Futures shares a top position in the Sharpe, Sortino- and Omega ratio ranking. As before, it is observed that well performing strategies exhibit relatively high standard deviations, encouraging economic theory that higher returns imply higher volatility. Moreover, no particular investment objective is ranked consistently across all performance measures. Clearly, the deviating results leave strong suggestion that the choice of performance measure exercises influence on the outcome of the strategy ranking. Additionally, the largest negative skewness is observed for Fund of Funds residuals, suggesting greater exposure to negative returns. Hedge funds in the category Other are strongly leptokurtic (fat tails), implying large returns with a higher frequency than expected under the normal distribution. The null hypothesis of normality is rejected for over 50% of the fund residuals in all categories, with the exception of Managed Futures, which also shows less funds with significantly heteroskedastic residuals. Finally, column 9 brings to light that residuals of the fund categories Other, Event Driven and Multi-Strategy are more prone to significant first-order autocorrelation than Global Macro and Managed Futures funds.

3.III. PERFORMANCE PERSISTENCE

This subsection is devoted to the findings on performance persistency. To reiterate, the sample period is divided into two sub-periods of equal horizon. The first sub-period ranges from January 2002 to December 2006, while the second sub-period reaches from January 2007 to December 2011. Then, the performance measures are computed for all 955 hedge funds separately for both sub-periods. Finally, two non-parametric tests are performed to test for the presence of performance persistency. These are the contingency table-based test and the Spearman rank correlation coefficients test.

3.III.1 CONTINGENCY TABLE-BASED TEST

Table 6 displays the four 4x4 contingency tables representing the predefined performance measures, with the percentage deviation[14] of each cell given in parentheses. The Chi-square statistic as in (10) is reported in the bottom row of each contingency table. Based on the Chi-square values of 96.13, 119.83, 136.61 and 182.52 for the Sharpe ratio, Sortino ratio, Omega ratio and Jensen's alpha, respectively, the null hypothesis of the absence of persistence is strongly rejected at the 1% significance level for all performance measures. Thus, in terms of the relative performance of the hedge funds, there seems to be some predictive power. That is, if a fund performed well during a first period, it is likely to do so again in the next and *vice versa.*

A closer inspection of quartile movement within a specific performance measure is given as follows. The cells on the diagonal of the contingency tables indicate the number of funds that remained in the same quartile during both sub-periods. The largest positive percentage deviation is found in cell Q1Q1 across all measures. That is, the top performing funds show the highest persistency regardless of the choice of the performance measure. With regards to the Sharpe ratio, Sortino ratio and Omega ratio,

[14] The percentage deviation is a measure that specifies to what degree an observed value differs from the expected value based on the null hypothesis.

second- and third-quartile funds have the tendency to move into a lower quartile over time, as represented by the relatively large percentage deviations of 30.4%, 37.1% and 35.4 in cell Q2Q3 and 46.1%, 52.8% and 46.1% in cell Q3Q4. For Jensen's alpha, the highest percentage deviations of each row are situated on the table diagonal, suggesting the strongest performance persistence, which is also reflected by the highest Chi-square value.

Table 6: Contingency tables per performance measure

This table displays the four 4x4 contingency tables for each performance measure. The number in cell $QxQy$ denotes the number of funds that moved from quartile Qx in the first sub-period to quartile Qy in the second sub-period. The percentage deviation is given in parentheses. The sample period is January 2002 to December 2011.

Panel A: Sharpe ratio contingency table

	Sub-period 2				
Sub-period 1	Q1	Q2	Q3	Q4	Total
Q1	110 (+83.9%)	51 (-14.7%)	43 (-28.1%)	35 (-41.2%)	239
Q2	43 (-28.1%)	58 (-3.0%)	78 (+30.4%)	60 (+0.7%)	239
Q3	35 (-41.5%)	57 (-4.7%)	60 (+0.3%)	87 (+46.1%)	239
Q4	51 (-14.2%)	73 (+22.6%)	58 (-2.6%)	56 (-5.6%)	238
Total	239	239	239	238	955

Note: *The null hypothesis of the absence of persistence was rejected at the 1% significance level, with a Chi-square statistic of 96.13*

Panel B: Sortino ratio contingency table

	Sub-period 2				
Sub-period 1	Q1	Q2	Q3	Q4	Total
Q1	114 (+90.6%)	48 (-19.7%)	44 (-26.4%)	33 (-44.6%)	239
Q2	39 (-34.8%)	65 (+8.7%)	82 (+37.1%)	53 (-11.0%)	239
Q3	31 (-48.2%)	54 (-9.7%)	63 (+5.3%)	91 (+52.8%)	239
Q4	55 (-7.7%)	72 (+20.9%)	50 (-16.1%)	61 (+2.8%)	238
Total	239	239	239	238	955

Note: *The null hypothesis of the absence of persistence was rejected at the 1% significance level, with a Chi-square statistic of 119.83*

Panel C: Omega ratio contingency table

	Sub-period 2				
Sub-period 1	Q1	Q2	Q3	Q4	Total
Q1	122 (+104.0%)	50 (-16.4%)	35 (-41.5%)	32 (-46.3%)	239
Q2	36 (-39.8%)	63 (+5.3%)	81 (+35.4%)	59 (-0.9%)	239
Q3	35 (-41.5%)	53 (-11.4%)	64 (+7.0%)	87 (+46.1%)	239
Q4	46 (-22.8%)	73 (+22.6%)	59 (-0.9%)	60 (+1.2%)	238
Total	239	239	239	238	955

Note: *The null hypothesis of the absence of persistence was rejected at the 1% significance level, with a Chi-square statistic of 136.61*

Panel D: Jensen's alpha contingency table

	Sub-period 2				
Sub-period 1	Q1	Q2	Q3	Q4	Total
Q1	126 (+108.9%)	59 (-1.4%)	30 (-49.4%)	24 (-59.7%)	239
Q2	61 (+1.1%)	71 (+18.7%)	59 (-0.5%)	48 (-19.4%)	239
Q3	24 (-60.2%)	54 (-9.7%)	86 (+45.0%)	75 (+25.9%)	239
Q4	30 (-50.1%)	55 (-7.7%)	62 (+5.0%)	91 (+53.4%)	238
Total	239	239	239	238	955

Note: *The null hypothesis of the absence of persistence was rejected at the 1% significance level, with a Chi-square statistic of 182.52*

3.III.2 SPEARMAN RANK CORRELATION COEFFICIENTS TEST

The Spearman rank correlation coefficient test results are reported in Table 7. Similarly to the contingency table-based test, the null hypothesis of no persistency is rejected for all measures. The largest correlation coefficient is observed for Jensen's alpha (0.368), followed by the Omega ratio

(0.259), the Sortino ratio (0.246) and the Sharpe ratio (0.225). Interestingly, the results contradict earlier work done by (Nguyen-Thi-Thanh (NTT), 2008), who detected no rank stability for the Sharpe- Sortino- and Omega ratio. However, Liang (2000) points out that contradicting results in previous studies were often caused by the different datasets that were employed. It applies to this paper as follows. This study considers 955 hedge funds across all investment strategies, while NTT focused exclusively on 149 Long/Short Equity hedge funds. Furthermore, NTT extracted funds from the Center for International Securities and Derivatives Markets (CISDM), while this report employs data obtained from Lipper TASS. In addition, NTT studies funds over a different sample period[15]. The next paragraph shows how this can severely influence the correlation coefficients. These findings once more emphasize the pitfalls and difficulties of hedge fund analysis.

Panel B shows similar tables as in Section 3.II., albeit now for the two sub-periods of 2002 to 2006 and 2007 to 2011. Surprisingly, the Spearman rank correlations of Jensen's alpha with the other measures are substantially lower during the first sub-period than during the second sub-period, with an average value of 0.56 as opposed to 0.92. This finding provides a first evidence of erroneous rankings if the performance measure is not properly selected. The rank correlations between the Sharpe ratio, Sortino ratio and the Omega ratio remain high for both sub-periods.

Table 7: Summary sub-period Spearman rank correlations

In this table, Panel A displays the Spearman rank correlation coefficients of the performance measures over the two sub-periods, along with their respective t-statistics (p-values in parentheses). Panel B reports the Spearman rank correlation coefficients between the different measures per sub-period. Sub-period 1 is January 2002 to December 2006, while sub-period 2 covers January 2007 – December 2011.

Panel A: Sub-period Spearman rank correlation coefficients per performance measure

$r_{s,Sharpe}$	t-stat$_{Sharpe}$	$r_{s,Sortino}$	t-stat$_{Sortino}$	$r_{s,Omega}$	t-stat$_{Omega}$	$r_{s,Jensen}$	t-stat$_{Jensen}$
0.225	7.13 (0.00)	0.246	7.84 (0.00)	0.259	8.28 (0.00)	0.368	12.22 (0.00)

Panel B: Spearman rank correlation coefficients per sub-period

Sub-period 1

Performance Measure	Sharpe Ratio	Sortino Ratio	Omega Ratio	Jensen's alpha
Sharpe Ratio	1			
Sortino Ratio	0.985	1		
Omega Ratio	0.988	0.984	1	
Jensen's alpha	0.545	0.583	0.548	1

Sub-period 2

Performance Measure	Sharpe Ratio	Sortino Ratio	Omega Ratio	Jensen's alpha
Sharpe Ratio	1			
Sortino Ratio	0.999	1		
Omega Ratio	0.999	0.997	1	
Jensen's alpha	0.920	0.922	0.921	1

Note: *All values in this table are significant at the 1% confidence level.*

[15] Nguyen-Thi-Thanh (2008) studies funds over the period January 2000 to December 2005, while this paper studies funds over the period January 2002 to December 2011.

4. CONCLUDING REMARKS

In this paper I conduct hedge fund analysis on 955 hedge funds for the period January 2002 to December 2011 with the use of four of the most frequently applied performance measures: the Sharpe ratio, the Sortino ratio, the Omega ratio and Jensen's alpha.

The first objective of this study was to determine whether hedge funds outperform the market, as represented by the S&P 500. For this purpose, I computed the performance measures for both the hedge funds as well as the benchmark. Approximately 84% of the hedge funds scored better than the benchmark, regardless of the choice of the performance measure, indicating that hedge funds do indeed provide their investors with superior returns. Yet, a respectable amount of studies point out that the non-normality of the hedge fund returns lead to severe overestimations of the performance measures. Hence, they conjecture that the results are unreliable. However, this paper reveals that normality of the market returns is strongly rejected as well. Thus, it is reasonable to assume that the performance overestimation occurs in both camps. How this affects its performance relative to the hedge funds is not so clear-cut. Nonetheless, the results of the performance measure comparison cannot be so easily discarded.

The second objective was to investigate whether the choice of the performance measure has an impact on *relative* hedge fund performance. To this end, a Spearman rank correlation coefficients test was performed. Extremely high rank correlations were found between the Sharpe ratio, Sortino ratio and Omega ratio. In other words, the three measures produce virtually identical fund rankings. Surprisingly, they exhibit somewhat lower rank correlations with Jensen's alpha than usually seen in literature. However, it is shown that the rank correlations of Jensen's alpha are exceedingly sensitive to the choice of the sample period, which is one possible explanation for the deviating findings. Nevertheless, alpha provides a different insight into fund rankings for the sub-period 2002 to 2005. With regards to investment categories, the performance measures generally provide different rankings, although some common preferences are observed. Furthermore, better performing strategies characterize with relatively high standard deviations, supporting the economic theory that higher returns imply higher volatility.

The third and final objective was to determine the presence or absence of performance persistency. To this extent, the data was divided into two sub-periods of equal horizon. Both the contingency table-based test and the Spearman rank correlations test find significant persistency. Hence, one could conjecture that some predictive power is present with regards to a fund's future performance. It is interesting to note that the sub-period analysis revealed that Jensen's alpha provides a different fund ranking as opposed to the Sharpe-, Sortino- and Omega ratio during the first sub-period, while delivering nearly identical rankings for the second sub-period.

SUGGESTIONS FOR FURTHER RESEARCH

This section is dedicated to some suggestions for further research. The recommendations also reveal improvements of the approaches used in this paper.

First, this study uses the S&P 500 as a market proxy for all hedge funds. Theoretically, this implies that only hedge funds exclusively trading stocks from the S&P 500 should be considered. Given the dynamic trading strategies of hedge funds and the study objective, it is difficult to assign a proper market proxy to each individual hedge fund. However, it is possible to use different market proxies for each hedge fund investment category. This should improve the accuracy of the benchmark comparisons.

Second, with regards to hedge fund rankings, the deviating results of Jensen's alpha invite for further investigation. In particular, the study would focus on the period January 2002 to December 2005, and examine what the cause is of the differing hedge fund rankings.

Third, while this paper provides some insight into the stability of hedge fund performance over time, it does not provide extensive information about the individual hedge fund movement. A possible approach is suggested in (Nguyen-Thi-Thanh, 2008), who puts forth an ascendant hierarchical clustering algorithm to determine individual consistency.

Finally, the exploration of relatively new methods for hedge fund analysis, such as data envelopment analysis (DEA) or a Bayesian and bootstrap approach (Kosowski, Naik, & Teo, 2005) should provide additional information on hedge fund performance.

REFERENCES

Agarwal, & Naik. (2000). *Multi-Period Performance Persistence Analysis of Hedge Funds.* Journal of Finance and Quantitative Analysis, Vol.35, No.3.

Agarwal, & Naik. (2004). *Risks and Portfolio Decisions Involving Hedge Funds.* Review of Financial Studies, Vol. 17, 2004, pp. 63-98.

Agarwal, Daniel, & Naik. (2004). *Flows, Performance and Managerial Incentives in Hedge Funds.* EFA 2003 Annual Conference Paper No. 501 .

Amin, & Kat. (2001). *Hedge Fund Performance 1990 - 2000. Do the 'Money Machines' Really Add Value?*

Brooks, & Kat. (2001). *"The Statistical Properties of Hedge Fund Index Returns and their Implications for Investors.* Journal of Alternative Investments, vol. 5: 26-44.

Brown, Fraser, & Liang. (2008). *Hedge Fund Due Dilligence: A Source of Alpha in a Hedge Fund Portfolio Strategy.*

Brown, Goetzmann, & Ibbotson. (1999). *Offshore Hedge Funds: Survival & Performance 1989-1995.* The Journal of Business, Vol.72, No.1.

Brown, Goetzmann, & Ibbotson. (1997). *Offshore Hedge Funds; Survival and Performance 1989-1995.*

Edwards, & Caglayan. (2001). *Hedge Fund Performance and Manager Skill.* Journal of Futures Markets, Vol. 21,11, pp.1003-1028.

Edwards, & Caglayan. (2000). *Hedge Funds and Commodity Fund Investments in Bull and Bear Markets.*

Eling, & Schuhmacher. (2007). *Does the Choice of Performance Measure Influence the Evaluation of Hedge Funds.*

Fung, & Hsieh. (1997a). *Empirical Characteristics of Dynamic Trading Strategies: The Case of Hedge Funds.* The Review of Financial Studies, Vol. 10, No. 2.

Fung, & Hsieh. (2000). *Performance characteristics of hedge funds and CTA funds: Natural versus spurious biases.* Journal of Financial and Quantitative Analysis 35, 291-307.

Fung, & Hsieh. (2001). *The Risk in Hedge Fund Strategies: Theory and Evidence from Trend Followers.* Review of Financial Studies, vol. 14: 313-341.

Gajda. (2009). *Hedge Funds Performance Evaluation.*

Gregoriou. (2002). *Hedge fund survival lifetimes.* Journal of Asset Management (2002) 3, 237–252.

Gregoriou, & Zhu. (2005). *Evaluating Hedge Fund and CTA Performance.* John Wiley & Sons, Inc.

Kaplan, & Knowles. (2004). *A Generalized Downside Risk-Adjusted Performance Measure.*

Kat, & Miffre. (2006). *The Impact of Non-Normality Risks and Tactical Trading on Hedge Fund Alphas.*

Kat, & Palaro. (2007). *FundCreator-Based Evaluation of Hedge Fund Performance.*

Kat, & Palaro. (2006). *Superstars or Average Joes? A Replication-Based Performance Evaluation of 1917 Individual Hedge funds.*

Keatwick, & Shadon. (2002). *A Universal Performance Measure.*

Kosowski, Naik, & Teo. (2005). *Do Hedge Funds Deliver Alpha? A Bayesian and bootstrap analysis.*

Liang. (2001). *Hedge Funds Performance: 1990-1999.* Financial Analyst Journal.

Liang. (2000). *Hedge funds: the living and the dead.* Journal of Financial and Quantitative Analysis 35, 309-327.

Liang. (1999). *On the Performance of Hedgefunds.* Financial Analysts Journal Vol. 55, No. 4, pp. 72-85.

Murthi, Choi, & Desai. (1997). *Efficiency of Mutual Funds and Portfolio Performance Measurement: A Non-Parametric Approach.* European Journal of Operation Research 98, 408-418.

Nalpas, Vanhems, & Simar. (2011). *Hedge Fund Selection: a Robust Directional Application.*

Nguyen-Thi-Thanh. (2008). *On the Consistency of Performance Measures for Hedge Funds.*

Schneeweis, Kazemi, & Martin. (2001). *Understanding Hedge Fund Performance; Research Results and Rules of Thumb for the Institutional Investor.*

Sharpe. (1966). *Mutual Fund Performance.* Journal of Business 39 (1), 119-138.

Sortino, & Price. (1994). *Performance Measurement in a Downside Risk Framework.*

APPENDIX A

A brief overview of the different hedge fund strategies[16].

Long/Short Equity: Involves the buying of long equities that are undervalued and the selling of short equities that are overvalued.

Managed Futures: The fund managers use futures contracts as part of their overall investment strategy.

Fund of Funds: This comprises of hedge funds that invest in other hedge funds.

Equity Market Neutral: This strategy exploits price differences in stocks by being long and short in stocks within the market.

Multi-Strategy: Uses various investment strategies to realize short- and long term gains.

Event Driven: Managers employing this strategy thrive on events, or 'special situations'. They take significant positions when these events occur. Examples these situations are mergers, takeovers and distressed stocks.

Global Macro: Hedge funds using this strategy make extensive use of leverage and derivates, whenever there is a perceived profit opportunity. They are responsible for most media attention.

Emerging Markets: This strategy is known for its high volatility. It entails investing in equity or debt of emerging markets, which tend to have higher inflation and volatile growth.

Fixed Income Arbitrage: This category profits from small arbitrage opportunities in interest rate securities.

Dedicated Short Bias: Hedge funds with this investment objective attempt to profit when the market declines, by holding investments that are biased to the short side.

Convertible Arbitrage: This investment strategy involves a long position on a convertible security and a short position in its converting common stock. It attempts to exploit pricing errors made during the conversion.

Options Strategy: This strategy uses a combined option position to profit from movements of the underlying stock.

[16] Loosely based on information from www.investopedia.com.

APPENDIX B

Table 8: S&P 500 Performance measures and descriptive statistics

This table reports the mean of the returns r_i, the net-of-risk free rate returns $r_i - r_f$ and of the returns standard deviations σ_i of the S&P 500. Columns four through six present its performance evaluation according to the Sharpe ratio, Sortino ratio and Omega ratio for the period January 2002 to December 2011.

Mean					
(r_i) [%]	$(r_i - r_f)$ [%]	σ_i [%]	Sharpe Ratio*	Sortino Ratio*	Omega Ratio*
0.347	0.194	4.597	0.042 (84.6%)	0.058 (84.2%)	1.119 (84.8%)

Note: *The percentage of hedge funds that had a higher performance measure is given in parentheses.*

Table 9: S&P 500 Summary statistics of alpha

This table reports the results of the alpha regression for the S&P 500. The first column shows the estimated Jensen's alpha, while the second and third column report its standard error and t-statistic (with *p*-value in parentheses). Columns four and five display skewness and kurtosis of the residuals. Columns six through eight show the p-values of tests for normality (Jarque-Bera), heteroskedasticity (Breusch-Pagan) and first order autocorrelation (Ljung-Box). All tests are conducted on the fund residuals.

alpha	Std. Err.	t-stat.	Skewness	Kurtosis	Jarque-Bera	Breusch-Pagan	Ljung-Box
0.163*	0.016	10.262 (0.00)	1.568	23.946	0.000	0.000	0.000

*: *significant at the 1% confidence level*

APPENDIX C

Table 10: Summary statistics of the Sharpe ratio, Sortino ratio and Omega ratio per hedge fund category.

This table reports summary statistics of the Sharpe ratio, Sortino ratio and Omega ratio per investment objective. Columns one to three show the means of the returns r_i, the net-of-risk free rate returns $r_i - r_f$ and the returns standard deviations σ_i of the 955 hedge funds per investment category. Columns four through nine present the mean and standard errors of the respective performance measures. The sample period is January 2002 to December 2011.

Category	Nr. of funds	Mean (r_i) [%]	Mean $(r_i - r_f)$ [%]	Mean σ_i [%]	Avg. Sharpe ratio*	Std. Err. Sharpe ratio	Avg. Sortino ratio*	Std. Err. Sortino ratio	Avg. Omega ratio	Std. Err. Omega ratio
Long/Short Equity	241	0.617	0.464	3.954	0.129 (9)	0.111	0.212 (9)	0.210	1.489 (7)	0.652
Managed Futures	92	0.853	0.700	4.896	0.337 (3)	0.632	0.337 (6)	0.672	1.329 (9)	0.989
Fund of Funds	344	0.472	0.319	1.979	0.270 (4)	0.508	0.422 (2)	1.198	1.391 (8)	0.878
Equity Market Neutral	18	0.596	0.443	2.803	0.171 (8)	0.117	0.293 (8)	0.248	1.780 (4)	0.761
Multi-Strategy	59	0.801	0.648	2.924	0.460 (1)	0.590	1.046 (1)	2.072	1.937 (2)	1.553
Event Driven	55	0.685	0.532	2.781	0.234 (5)	0.139	0.404 (3)	0.308	2.148 (1)	0.968
Global Macro	22	0.833	0.675	3.575	0.191 (7)	0.094	0.340 (5)	0.207	1.734 (5)	0.457
Emerging Markets	54	1.204	1.051	5.975	0.196 (6)	0.094	0.315 (7)	0.188	1.793 (3)	0.553
Other	70	0.743	0.590	2.831	0.406 (2)	0.656	0.374 (4)	0.653	1.713 (6)	1.133

*: *The rank of each investment category is provided in parentheses.*

APPENDIX D

Table 11: Summary statistics Jensen's alpha per hedge fund category.

This table reports the results from the single-factor model, in which the fund returns in excess of the risk-free rate are regressed on the excess market returns obtained through the CAPM. Columns one and two report the mean and average standard error of the monthly alpha, respectively, per investment category. Columns three and four display skewness and kurtosis of the hedge fund residuals, while column five reveals the percentage of funds for which alpha is significant at the 10% confidence level. Columns six through nine respectively report the percentage of funds for which the null hypothesis of normality (Jarque-Bera), homoskedasticity (Breusch-Pagan) or no first-order autocorrelation (Ljung-Box) is rejected. All tests are conducted on fund residuals and maintain a 10% significance level. The sample period is January 2002 to December 2011.

		Mean					Test of normality	Test of heteroskedasticity	Test of autocorrelation
Category	Nr. of funds	alpha [%]	Std. err. alpha [%]	Skewness	Kurtosis	% of funds with alpha $p < 0.1$	% of funds with Jarque-Bera $p < 0.1$	% of funds with Breusch-Pagan $p < 0.1$	% of funds with Ljung Box $p < 0.1$
Long/Short Equity	241	0.45 (7)	0.29	0.11	5.53	48.13	63.49	44.81	23.24
Managed Futures	92	0.70 (2)	0.44	0.24	4.88	38.04	44.57	17.39	19.57
Fund of Funds	344	0.31 (9)	0.15	-0.58	6.68	50.29	85.17	60.47	54.07
Equity Market Neutral	18	0.44 (8)	0.23	-0.20	6.77	50.00	77.78	50.00	38.89
Multi-Strategy	59	0.64 (4)	0.24	-0.11	8.96	69.49	89.83	61.02	57.63
Event Driven	55	0.52 (6)	0.20	-0.07	7.40	80.00	94.55	54.55	58.18
Global Macro	22	0.67 (3)	0.30	0.23	4.70	68.18	77.27	40.91	13.64
Emerging Markets	54	1.03 (1)	0.44	-0.24	6.43	83.33	66.67	48.15	48.18
Other	70	0.53 (5)	0.23	0.45	11.67	64.29	91.43	58.71	82.86
All Categories	955	0.49	0.25	-0.21	6.71	54.76	75.71	50.58	43.98