

Hedge Fund Performance Using Scaled Sharpe And Treynor Measures

Francois van Dyk, UNISA, South Africa
Gary van Vuuren, North-West University, South Africa
André Heymans, North-West University, South Africa

ABSTRACT

The Sharpe ratio is widely used as a performance measure for traditional (i.e., long only) investment funds, but because it is based on mean-variance theory, it only considers the first two moments of a return distribution. It is, therefore, not suited for evaluating funds characterised by complex, asymmetric, highly-skewed return distributions such as hedge funds. It is also susceptible to manipulation and estimation error. These drawbacks have demonstrated the need for new and additional fund performance metrics. The monthly returns of 184 international long/short (equity) hedge funds from four geographical investment mandates were examined over an 11-year period.

This study contributes to recent research on alternative performance measures to the Sharpe ratio and specifically assesses whether a scaled-version of the classic Sharpe ratio should augment the use of the Sharpe ratio when evaluating hedge fund risk and in the investment decision-making process. A scaled Treynor ratio is also compared to the traditional Treynor ratio. The classic and scaled versions of the Sharpe and Treynor ratios were estimated on a 36-month rolling basis to ascertain whether the scaled ratios do indeed provide useful additional information to investors to that provided solely by the classic, non-scaled ratios.

Keywords: Hedge Funds; Risk Management; Sharpe Ratio; Treynor Ratio; Scaled Performance Measure

1. INTRODUCTION

In 1949 Alfred Jones started an investment partnership that is regarded as the first hedge fund, although wealthy individuals and institutional investors have been interested in hedge funds or ‘private investment vehicles’ since around the 1920s (Jaeger, 2003). By 1968 there was an estimated 140 live hedge funds while by 1984, the number had dropped to 68 (Lhabitant, 2002). The mid-1980s saw a revival of hedge funds that is commonly ascribed to the publicity surrounding Julian Robertson’s Tiger Fund (Agarwal & Naik, 2002) and, to a lesser extent, its offshore sibling, the Jaguar Fund (Connor & Woo, 2003). During this time, hedge funds became admired for their profitability¹ and since the explosive growth in the hedge fund market during the early 1990s, interest in hedge funds and their activities by regulators, investors and money managers has been ever increasing. The interest in hedge funds was further helped along owing to some headline-making news and extravagant hedge fund phenomena, such as the collapse of Long Term Capital Management (LTCM)² in the late 1990s, the loss of US\$2bn in 1998 by George Soros’ Quantum Fund during the Russian debt crisis, Amaranth Advisors³ in 2006, and the Madoff Ponzi scheme⁴ in late 2008. More recent reasoning behind the heightened interest

¹ A 1986 article in Institutional Investor magazine noted that since its inception in 1980, Tiger Fund had a 43% average annual return (Agarwal & Naik, 2002; Connor & Woo, 2003).

² LTCM is a large US-based hedge fund that nearly caused the collapse of the global financial system in 1998 due to high-risk arbitrage bond trading strategies. The fund was highly leveraged when Russia defaulted on its debt causing a flight to quality. The fund suffered massive losses and was ultimately bailed out with the assistance of the Federal Reserve Bank and a consortium of banks.

³ To date, Amaranth Advisors marked the most significant loss of value for a hedge fund. The hedge fund attracted assets under management of US\$9bn where after faulty risk models and non-rebounding gas prices resulted in failure for the funds’ energy trading strategy as it lost US\$6bn on natural gas futures in 2006. Amaranth was also charged with the attempted manipulation of natural gas futures prices. Refer to Till (2007) for further details.

in hedge funds can be explained by the poor performance exhibited by traditional asset investments (Almeida & Garcia, 2012).

During the 1990s, global investment in hedge funds increased from US\$50bn in 1990 to US\$2.2tn in early 2007 (Barclayhedge, 2014a). Over the period 2003 to 2007, the hedge fund industry posted its most significant gains, in terms of performance and asset flows, where after the financial crisis growth reduced significantly. Industry growth reversed, declining to US\$1.4tn by April 2009 due to substantial investor redemptions and performance-based declines (Eurekahedge, 2012). In 2012 the hedge fund industry suffered US\$3.8tn of new outflows (Eurekahedge, 2013), although during 2013 recovery for the industry was significant as hedge funds attracted net asset flows of US\$124.7bn during the first 11 months and also realised their best year of performance-based gains since 2010⁵ (Eurekahedge, 2014b). Short bias strategy funds ended 2013 27.15% in the red, thereby surpassing the previous year’s record loss of 24.12% (Barclayhedge, 2014b). According to Deutsche Bank’s 12th annual Alternative Investor Survey, hedge fund assets under management (AUM) are expected to reach US\$3tn by the end of 2014 (Deutsche Bank, 2014). Approximately 80% of respondents to the survey also stated that hedge funds performed as expected or better in 2013,⁶ while almost half of institutional investors increased their hedge fund allocation in 2013, and that 57% planned an allocation increase in 2014 (Deutsche Bank, 2014). Figure 1 presents the AUM for the hedge fund industry for 1997 to 2013.

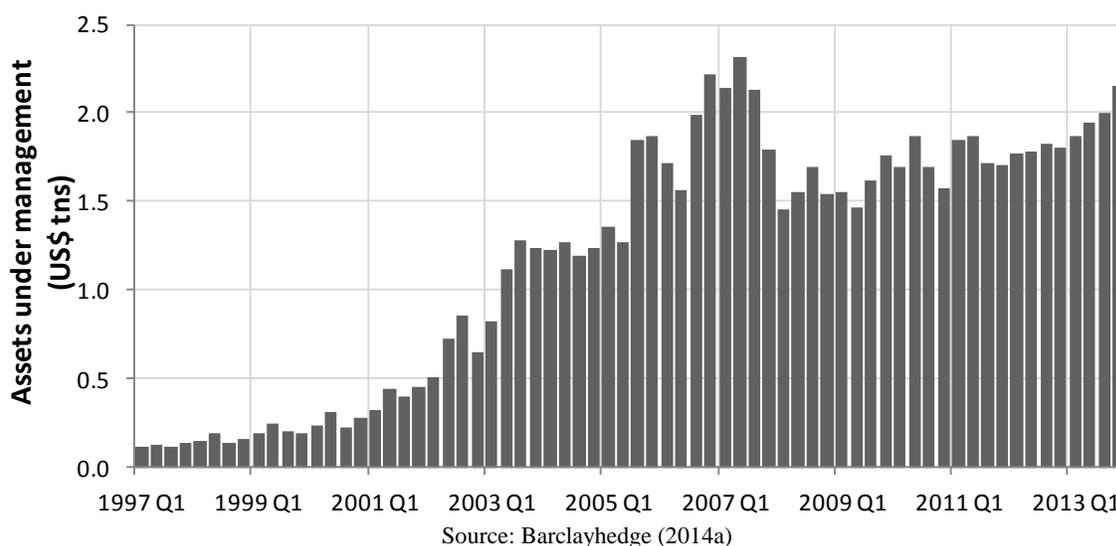


Figure 1: Hedge Funds’ Assets Under Management (US\$tn), Quarterly Since 1997

The recent (2007-9) financial crisis’ impact on hedge funds, and their performance compared to more traditional asset classes and benchmarks, also make for noteworthy reading. The average annual hedge fund return between 2002 and 2012 was 6.3% (TheCityUK, 2012) compared to 5.7% for U.S. bonds,⁷ 7.8% for global bonds⁸, and 6.0% for the S&P500. The 2013 comparison notes that the Barclay Hedge Fund Index gained 11.21% (Barclayhedge, 2014b) compared to returns of 29.6% for the S&P500 (CNBC, 2013) and -2.1% for U.S. bonds (Financial Times, 2014). In 2008 the hedge fund industry posted its worst annual performance since 1990 (-20%). In 2011 fund liquidations also rose to 775 - an increase of 4% from 743 in 2010. Even though the total number of funds rose to 9,523 in 2011 and further to 10 100 at the end of 2012 (TheCityUK, 2013), this number still (2014) fails to

⁴ Considered the largest financial scandal in modern times with losses estimated at US\$85bn, Madoff Securities LLC provided investors with modest, yet steady, returns and claimed to be generating these returns by trading in S&P 500 index options employing an index arbitrage strategy. Madoff Securities did, however, commit fraud through a Ponzi scheme structure.

⁵ Long/short equities strategies accounted for almost half of the gains in 2013 (Eurekahedge, 2014b).

⁶ According to the Deutsche Bank Alternative Investor Survey, allocations to hedge funds returned a weighted average of 9.3% in 2013. Equity long/short and event-driven funds also proved the most sought-after strategies (Deutsche Bank, 2014).

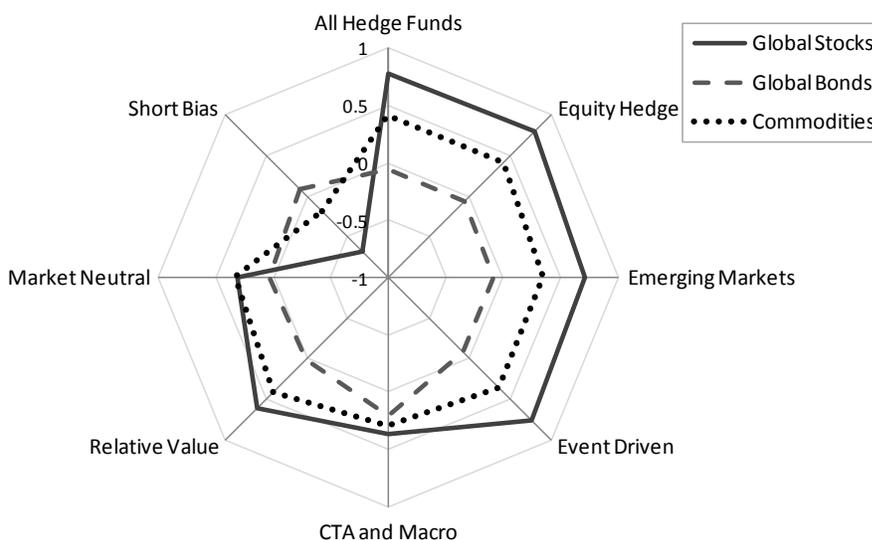
⁷ U.S. bonds as measured by the Barclays U.S. Aggregate Bond Index

⁸ Global bonds as measured by the JP Morgan Global Government Bond Index (unhedged)

eclipse the pre-crisis peak of 10,096 at the end of 2007 (Clarke, 2012). In terms of the industry’s asset size, 2008 saw AUM decline 27% to US\$1.4tn (Roxburgh *et al.*, 2009) and then even further in March 2009 to US\$1.29tn (Eurekahedge, 2010), reflecting both asset withdrawals and investment losses.

Investor withdrawals subsequent to the financial crisis added to poor performance, as it became evident that hedge funds had not “hedged” at all. This has resulted in a high attrition rate⁹ (Liang, 1999) which, over time, has also increased significantly. Only 91% of funds that were alive in 1996 were still alive in 1999, while this declined to 59.5% in 2001 (Kat & Amin, 2001). In addition, Kaiser and Haberfelner (2012) found that since the financial crisis, the attrition rate for hedge funds has nearly doubled. In the ruthless world of fund performance, the reporting of monthly returns can exacerbate investor outflows, halt them, reverse them, or increase them – depending on the reported figures. A strong incentive to exaggerate or misrepresent fund performance therefore exists, as not only does stronger performance bolster capital inflows, but it also reinforces a fund’s existence and increases manager incentive fees (see Goetzmann *et al.*, 2007; Bollen & Pool, 2009; Agarwal *et al.*, 2011; Feng, 2011). As investors also pay high fees - typically in the vicinity of a 2% management fee and a 20% performance fee - performance evaluation and an accurate performance evaluation methodology are of critical importance to investors (Lopez de Prado, 2013).

Hedge funds are often seen as a way of improving portfolio performance. For both hedge funds and investors, performance measurement is an integral part of investment analysis and risk assessment. It is, however, also the case that investors are enticed to invest in hedge funds for the influential motive that the returns of these funds appear uncorrelated with the broader market. Hedge funds are generally characterised by low correlations with traditional asset classes and hence put forward potentially attractive diversification benefits for asset portfolios (Fung & Hsieh, 1997; Liang, 1999; Kat & Lu, 2002; KPMG, 2012). Figure 2 presents the correlation between various hedge fund strategies and main asset classes for the period 1994 to 2011.

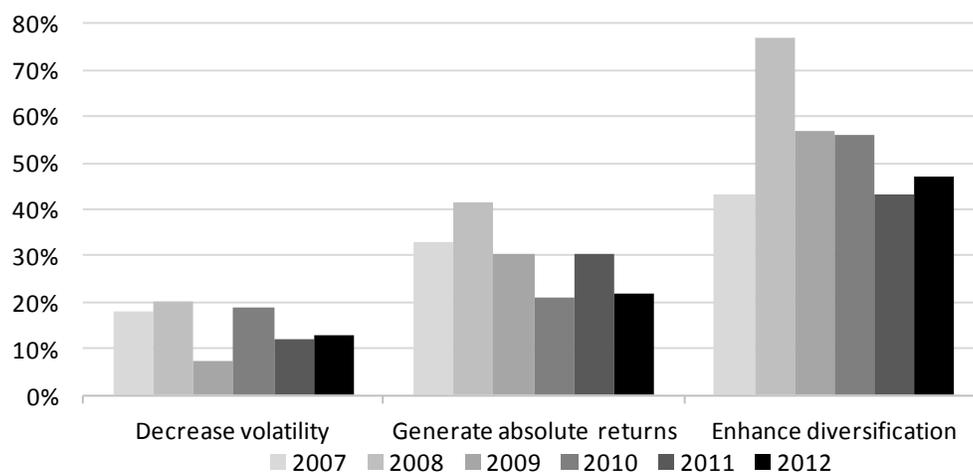


Source: KPMG (2012). Global Stocks = MSCI World Total Return Index, Global Bonds = JP Morgan Global Aggregate Bond Total Return Index, Commodities = S&P GSCI Commodity Total Return Index
Hedge fund performance using HFR equal-weighted index and strategy indices

Figure 2: Correlations Between Hedge Funds And Main Asset Classes (January 1994 – December 2011)

Survey results from SEI Knowledge Partnership (SEI, 2007; 2009-2013) also show that institutional investors are less concerned with achieving absolute returns than they are with obtaining differentiated, non-correlated returns (see Figure 3). Figure 3 also points to the heightened investor demand for the diversification benefit hedge funds offered during the recent financial crisis period.

⁹ Liquidation rate of funds



Source: SEI (2007; 2009-2013). Diversification category includes diversification and non-correlation with other asset classes.

Figure 3: Primary Objective Of Institutional Investors When Investing In Hedge Funds

As these alternative investments, which are hedge funds, embrace a variety of diverse strategies, styles and securities, specifically designed risk assessment techniques and measures are necessitated. Regardless of the potential diversification benefit being offered, these funds remain highly risky investments as stellar returns cannot be obtained without significant risk (Botha, 2007). Malkiel and Saha (2005) also state that although being outstanding diversifiers, hedge funds are risky due to the cross-sectional variation and the range of individual hedge fund returns being far greater than those of traditional asset classes. Hedge fund investors thus take on considerable risk in selecting a poorly performing or failing fund.

Although most comparisons of hedge fund returns concentrate exclusively on total return values, comparing funds with different expected returns and risks in this manner is meaningless. The arrangement of risk and return into a risk-adjusted number is one of the primary responsibilities of performance measurement (Lhabitant, 2004). According to Eling and Schuhmacher (2006), financial analysts, and often individual investors, rely on risk-adjusted return - i.e. performance measures in order to select among available investment funds, and since the seminal work of Jensen (1968), Treynor (1965) and Sharpe (1966), performance measures have been the focus of much attention from both practitioners and researchers. These measures are mostly used by researchers to evaluate market efficiency while practitioners use them in at least two instances: (i) to evaluate past performance (in the hope that the measure is a reliable indicator of future performance) and (ii) to measure performance and compare the results of one fund to its competitors or those of a representative market of benchmark (Nguyen-Thi-Thanh, 2010). Nguyen-Thi-Thanh (2010) argues that in the literature on portfolio performance evaluation, two kinds of portfolio performance measures come to light. The first kind evaluates the fund managers' skills¹⁰; i.e. their timing and selectivity ability, and includes measures such as Treynor, Jensen and other multi-factor models. The second kind includes measures, such as the Sharpe ratio, which relates to measures that lead to complete fund ranking. The latter type is primarily used in the first, or screening, phase to create a short list of the best performing funds on which further detailed quantitative and or qualitative analysis will be applied before the investment decision is made. To warrant that performance measures are not easily gamed by unskilled managers and also that investors do not pay manager for strategies that they themselves can easily replicate, Chen and Knez (1996) propose that a performance measure should (i) be fit for purpose; i.e., be reasonably useable; (ii) be scalable; (iii) be continuous; and (iv) exhibit monotonicity.¹¹

Evidence indicates that fund managers are not fully using the performance measurement techniques proposed by the literature. The results of a 2008 survey by Amenc *et al.*, (2008) indicate that the majority of survey

¹⁰ A rich literature has developed on methodologies that test for fund manager skills. These techniques can be classified into two main approaches - (i) returns-based performance evaluation and (ii) portfolio holdings-based performance evaluation (Wermers, 2011).

¹¹ The assignment of higher measures for more skilled managers and lower measures for less-skilled ones (Chen & Knez, 1996)

respondents do not use sophisticated approaches and that a large gap exists between practices and academic models. The survey results highlight that the Sharpe ratio (80%) and the Information ratio (80%) are the most widely used performance evaluation measures among asset managers. Amongst hedge funds, the Sharpe ratio is the metric of choice and also the most commonly used measure of risk-adjusted performance (Lhabitant, 2004; Opdyke, 2007; Schmid & Schmidt, 2007). Proposed by Sharpe as the “reward-to-variability” ratio as a mutual fund comparison tool (see, Sharpe, 1966, 1975, and 1994), the ratio is both conceptually simple and rich in meaning, providing investors with an objective, quantitative measure of performance. It enjoys widespread use and various interpretations, but it also has its drawbacks. Being unsuitable for dealing with asymmetric return distribution are, among others, a drawback of volatility measures (Lhabitant, 2004; Almeida & Garcia, 2012). Academic criticism of the classic capital asset pricing model (CAPM) performance measure is not new and a number of authors have pointed out the shortcomings of using both the Sharpe ratio for performance evaluation and the mean-variance framework for portfolio construction when the underlying returns distributions are highly non-symmetric. According to Almeida and Garcia (2012), the key is to risk-adjust hedge fund payoffs in a manner that accounts for the asymmetry (tail risk exposures) created by the dynamic strategies hedge funds pursue. A suitable risk-adjusted performance measure for hedge funds will therefore not only be based on returns’ means and volatilities, which are not adequate given the deviations from normality exhibited by hedge fund returns, but also on higher-order moments of the hedge fund returns distribution. Similar reasoning brought Leland (1999:30) to the conclusion that is additionally described as a daunting task - “any risk measure in this world must capture an infinite number of moments of the return distribution”.

This brings forward the aim of this study of evaluating whether scaled (risk-adjusted) performance measures, in the form of scaled Sharpe and Treynor¹² ratios, should augment the use of the classical or traditional Sharpe and Treynor ratios when evaluating hedge fund risk and, consequently, in the investment decision-making process. The rationale behind this is that the scaled performance measures provide a more suitable evaluation of hedge fund risk-adjusted performance since the traditional Sharpe and Treynor ratios are ill-suited to hedge funds.

The analysis is built upon data sourced from the Eureka hedge database. It contains data from 184 ‘live’ hedge funds which have a developed market focus from four geographical investment mandates. The analysis covers the years 2000 through 2011, which is advantageous for three reasons. First, the results do not suffer from survivorship and backfilling biases to the same extent that plagues a greater amount of the older hedge fund research.¹³ Second, unlike many other studies that are limited to analysis that only include bull markets,¹⁴ the chosen time period contains bull and bear markets, allowing fund analysis in different market conditions. Third, the chosen time period contains a critical event - the 2007-2009 global financial crisis - which is considered in additional detail during analysis and in sub-periods.

The methodology used in this study is based on the ratio scaling methodology by Gatfaoui (2012) while this study also builds upon and differentiates itself from the prior research in the following manners:

- A (36-month) rolling (geometric) analysis period is used compared to the static (month-by-month) methodology of Gatfaoui (2012).
- The data time-series include periods from *pre*, *during* and *post* the recent financial crisis, compared to the research data by Gatfaoui (2012) that only include the periods *pre* and *during* the crisis.
- The ratio analysis and comparisons are performed on ‘live’ individual hedge funds as well as market and hedge fund indices from four geographical investment mandates. The comparative ratio analysis by Gatfaoui (2012) focuses solely on various hedge fund strategy-applicable market indices.

¹² Reasons for the inclusion of the Treynor ratio (in this study) are: (i) the Treynor ratio is a commonly used performance measure, (ii) the Treynor ratio suffers from a similar drawback to the Sharpe ratio due to not accounting for higher-order moments of the return distribution, (iii) the addition of the Treynor ratio differentiates this study as numerous studies pertaining to the incorporation of higher-order moments into the Sharpe ratio have been conducted, and (iv) the addition of the Treynor ratio adds another element of analysis to the study.

¹³ Prior to 1994, most hedge fund data vendors (databases) did not cover dissolved hedge funds. Hedge fund data prior to 1994 are thus not very reliable. The unreliability of data prior to 1994 is discussed by Fung and Hsieh (2000), Liang (2000) and Li, and Kazemi (2007).

¹⁴ Capocci *et al.* (2005) found that the market phase may influence the results. Ding and Shawky (2007) stress the importance of considering different market cycles when analysing hedge fund performance. Also see Brown *et al.* (1999).

The remainder of this paper presents an existing literature overview of hedge fund performance measurement, alternative performance measures and the ill-suitedness of the Sharpe ratio as a hedge fund performance measure; introduces the scaled Sharpe and Treynor measures, as well as the data and methodology employed; presents the analysis and results; and concludes.

2. LITERATURE STUDY

2.1. Hedge Fund Performance Measurement

In the hedge fund industry, performance is of considerable importance as not only is investor returns based on fund performance, but hedge fund manager compensation is also tied to fund performance. As a result, performance *measurement* is an integral part of investment analysis and risk management. The literature on the topic is abundant and controversial.

Fund performance evaluation can be classified into two major approaches: (i) returns-based and (ii) portfolio holdings-based. Both approaches have been applied by researchers in simplistic as well as more sophisticated and innovative manners, and each approach has its advantages and disadvantages (Wermers, 2011). Returns-based approaches, for instance, rely on less information from fund managers and is therefore particularly useful where little information is disclosed, such as in hedge fund markets. Returns data are available on a more frequent basis, even where portfolio holdings are on hand. The returns-based performance approach is, however, the focus of this study.

The abundance of literature on performance measurement in the hedge fund industry stems from the fact that performance measurement is a key facet of the quantitative analysis required in the rigorous process of fund selection. Géhin (2006) describes (quantitative) fund selection as more than a challenging task on account of (i) the increasing number of funds, (ii) short fund track records, (iii) fund managers not having equal talent, and (iv) the hedge fund universe's opacity. The quantitative analysis of hedge funds consequently requires genuine expertise and must, moreover, be sophisticated. The controversial nature of the literature can arguably also be attributed to the numerous qualities of hedge funds, as these funds invest in a heterogeneous range of asset classes¹⁵, and that a broad range of strategies are covered that are, in turn, characterised by different risk and return profiles.¹⁶ The same reasons responsible for the abundance and controversial nature of the literature can arguably also be attributed as the reasons behind specific focus areas being especially prominent within the literature, for instance - the choice of performance measure in hedge fund performance evaluation, the role of the measure choice on performance evaluation¹⁷, and the consistency of these measures. Prior research on hedge fund performance rankings produced by common risk-adjusted performance measures also shows remarkable homogeneity¹⁸ and thus results in the same investment decision. Even though prior and current hedge fund performance studies have been criticised for the performance methods employed and conflicting conclusions, these studies contribute to a growing improvement in the understanding of alternative investments. Identifying a performance measure that can serve as a robust proxy for a number of other measures could thus significantly aid performance measurement by private and professional investors (Prokop, 2012).

Lastly, unlike traditional investments that invest only in traditional asset classes, hedge funds include options and derivative products. These sophisticated financial instruments create various further complications seeing that the commonly used performance measures, which were developed based on modern portfolio theory, were specifically designed for traditional asset classes or investments and, in particular, for equity investments.¹⁹ The key task of performance measurement, however, remains to condense risk and return into one useful risk-adjusted number (Lhabitant, 2004) that can thereafter be used to make sound investment decisions.

¹⁵ Examples of the financial assets that hedge funds invest in include equities, bonds, swaps, currencies, sophisticated derivative securities, convertible debt and mortgage-backed securities.

¹⁶ Hedge funds can, for example, employ directional and non-directional strategies. Directional strategies aim to benefit from market trends and include fund strategies such as macro, short-selling and emerging markets. Non-directional strategies have weak correlation with the related market and include strategies such as distressed securities, market neutral, convertible arbitrage and event driven.

¹⁷ See for example Eling & Schuhmacher (2006), Nguyen-Thi-Thanh (2007, 2010) and Prokop, (2012).

¹⁸ See for example Kooli *et al.* (2005), Nguyen-Thi-Thanh (2007) and Prokop, (2012).

¹⁹ Some factors have been included to suit other asset classes (Sharpe, 1992; Elton *et al.*, 1993).

2.2. Inadequacy Of Traditional Performance Measures

Risk-adjusted performance measures can be classified into one of two categories, namely ‘absolute’ or ‘relative’ performance measures. The former is considered such as no benchmarks are used in the calculation with the Sharpe and Treynor ratios being the most common measures within this category. Jensen’s alpha (Jensen, 1968) is an example of a relative risk-adjusted performance measure and, in contrast to absolute performance measures, employs a benchmark (Géhin, 2006).

The Sharpe ratio is one of the most commonly cited statistics in financial analysis and the metric of choice amongst hedge funds, particularly as a measure of risk-adjusted performance (Lo, 2002; Lhabitant 2004; Opdyke, 2007; Schmid & Schmidt, 2007; Koekebakker & Zakamouline, 2008). Also known as the risk-adjusted rate of return, it measures the relationship between the risk premium²⁰ and the standard deviation of the fund returns (Sharpe, 1966, 1975, 1992, 1994). Another popular indicator of fund performance is the reward to variability or Treynor ratio (Treynor, 1965) and is defined through the relation of the risk premium and systematic risk²¹ of the portfolio (beta).²² The Sharpe and Treynor ratios are similar in that they both divide the fund’s excess return by a numerical risk measure. The Sharpe ratio, however, employs total risk, which is appropriate when evaluating the risk return relationship of a poorly diversified portfolio, while the Treynor ratio uses systematic (market) risk, which is the relevant measure of risk when evaluating a fully diversified portfolio (Jagric *et al.*, 2007). For fully diversified portfolios, total and systematic (market) risk are equal and fund rankings based on total risk and systematic risk should be identical for a well-diversified portfolio.²³ Despite the widespread use of these measures, they do have some failings.

Parameters and statistics for both the Sharpe and Treynor ratios in expected returns, volatilities and beta²⁴ are non-observable quantities, and, as they must be estimated, these are fraught with estimation errors.

The Sharpe ratio’s statistical properties have been afforded only modest consideration, which is surprising given that the accuracy of the Sharpe ratio’s estimators rely on the statistical properties of returns and that these may be very different among portfolios, strategies, and over time (Lo, 2002). The performance of more volatile investment strategies is more difficult to determine compared to less volatile strategies (Lo, 2002). Since hedge funds are generally more volatile than more traditional investments (Ackermann *et al.*, 1999; Liang 1999), hedge fund Sharpe ratio estimates are likely to be less accurate. Several statistical tests that look into comparing Sharpe ratios between two portfolios have been proposed by Jobson and Korkie (1981), Gibbons *et al.* (1989), Lo (2002), and Memmel (2003). Conversely, the unavailability of multiple Sharpe ratio comparisons has led to the search and development of alternative approaches (e.g. Ackermann *et al.*, 1999; Maller & Turkington, 2002). It is nonetheless apparent that a more refined Sharpe ratio interpretation approach is necessary whilst information pertaining to the investment style or strategy, and also the market environment which produced the returns, should possibly be considered by such an approach. Additionally, it has been established that the Sharpe ratio is susceptible to manipulation (e.g. Spurgin, 2001; Goetzmann *et al.*, 2002, 2007).

The Treynor ratio also has its drawbacks. Firstly, the measure validity depends significantly on the hypothesis that the fund's beta is stationary.²⁵ The selection of the correct benchmark is also critical when employing the Treynor ratio (Eling, 2006; Ambrosio, 2007).

²⁰ Risk premium is defined as the additional expected return from holding a risky asset rather than a riskless asset – i.e., the difference between the expected return (on an investment) and the estimated risk-free return.

²¹ Systematic risk is also known as “market risk”, “undiversifiable risk”, or “volatility”.

²² The Treynor ratio, unlike the Sharpe ratio, should not be used on a stand-alone basis as beta is a measure of systematic (market) risk only. This is so, as Choosing a stand-alone investment portfolio on the basis of the Treynor ratio may be inclined to maximise excess return per unit of systematic (market) risk, but not excess return per unit of total risk, except if each investment is well diversified (Anson *et al.*, 2012).

²³ This is the case as the total risk is reduced (through diversification) to leave only systematic risk.

²⁴ Beta consists of variance and co-variance.

²⁵ i.e., that the fund manager does not adapt the portfolio’s weights according to future market variation expectations (Eling, 2006)

The assumption of normally distributed returns is widely considered the most significant drawback of both measures, as both are based on the mean-variance framework which employs the Capital Asset Pricing Model (CAPM) methodology. Strong assumptions underlie the CAPM, *e.g.* (i) returns are normally distributed and (ii) investors care only about the mean and variance of returns, so upside and downside risks are viewed with equal dislike (Leland, 1999). Hedge fund return distributions and their markedly non-normal characteristics have been extensively portrayed in the literature (see, *e.g.* Fung & Hsieh, 2001; Lo, 2001; Brooks & Kat, 2002; Malkiel & Saha, 2005). Brooks and Kat (2002) established that hedge fund indices show evidence of low skewness and high kurtosis while Eling (2006), Eling and Schumacher (2006), and Taleb (2007) found hedge fund return distributions to be negatively skew and to possess positive excess kurtosis.^{26 27}

Also, under the CAPM methodology, the appropriate measure of risk is represented by beta while the named CAPM assumptions rarely hold in practice. Even if the underlying assets' returns are normally distributed, the returns of portfolios that contain options on these assets, or use dynamic strategies will not be (Leland, 1999). Hedge funds generally employ dynamic investment strategies with accompanying dynamic risk exposures and these have important implications for investors who seek to manage the risk/reward trade-offs of their investments (Chan *et al.*, 2005). For this reason, hedge fund performance is often summarised with multiple statistics.²⁸ While beta is an adequate risk measure for static investments, there is no single measure capturing the risks of a dynamic investment strategy (Chan *et al.*, 2005). Linear performance measures can often not capture the dynamic trading strategies that several hedge funds pursue (Agarwal & Naik, 2004) whilst hedge funds make use of a range of trading strategies. Analysing all hedge funds using a singular performance measurement framework that does not consider the characteristics of the specific strategies is of limited value. Therefore, it is necessary for hedge fund style-specific performance measurement models or measures to capture the differences in management style (Fung & Hsieh, 2001, 2004; Agarwal & Naik, 2004). A large number of equity-orientated hedge fund strategies also bear significant (left-tail) risk that is ignored by the mean-variance framework²⁹ (Lhabitant, 2004).

Asymmetric distributions further influence the validity of volatility as a risk measure which, in turn, impacts the exactness of the Sharpe ratio. Volatility solely measures the dispersion of returns around their historical average and since positive and negative deviations (from the average) are penalised in an equivalent manner in the computation, the concept is only logical and legitimate for symmetrical distributions (Lhabitant, 2004). In reality, return distributions are neither normal nor symmetrically distributed, and so even when two investments have an identical mean and volatility, they may exhibit substantially different higher moments. This is especially true for strategies that entail dynamic trading, buying, and selling of options and active leverage management (Lhabitant, 2004) – all strategies regularly employed by hedge funds. The return distributions of such strategies are highly asymmetric and possess “fat tails”, which leads to volatility being a less-meaningful measure of risk. The relevance of the dispersion of returns around an average has also been queried from an investor's viewpoint, as most investors perceive risk as a failure to achieve a specific goal, such as a benchmark rate (Lhabitant, 2004). In such circumstances, risk is only considered as the downside of the return distribution and not the upside; the difference is not captured by volatility (Lhabitant, 2004). Also, investors are more adverse to negative deviations than to positive deviations of the same magnitude (Lhabitant, 2004).

2.3. Alternative Risk Performance Measures

Lhabitant (2004) gives the drawbacks of volatility as a measure of risk as the reason behind the search for alternative risk measures. The Sharpe ratio's denominator (volatility) is replaced by an alternative measure of risk in many alternative risk performance measures. For example, under the mean-downside deviation framework, Sortino and Price (1994), as well as Ziemba (2005), substitute standard deviation by downside-deviation. Other downside

²⁶ Hedge fund index returns and market benchmarks generally exhibit the same stylised facts; *i.e.*, negative skewness and positive excess kurtosis (Gatfaoui, 2012).

²⁷ Investors show a preference for high first (mean) and third (skewness) moments and low second (standard deviation) and fourth (kurtosis) moments (Scott & Horvath, 1980).

²⁸ *E.g.* mean, standard deviation, Sharpe ratio, market beta, Sortino ratio, maximum drawdown, etc. (Chan *et al.*, 2005).

²⁹ These left-tail risks originate from hedge fund strategies that exhibit payoffs resembling a short position in a put option on the market index (Lhabitant, 2004).

risk measures, such as the Calmar ratio³⁰ (CR), Sterling ratio³¹ and Burke ratio³², use drawdown³³ in the denominator to quantify risk.

Gregoriou and Gueyie (2003) propose a modified Sharpe ratio, under the mean-VaR framework, as an alternative measure specifically for hedge fund returns by employing a Modified VaR³⁴ (MVaR) in place of standard deviation as the denominator. Also, Dowd (2000) uses a VaR measure as a standard deviation replacement, whilst conditional VaR (CVaR)³⁵ can be used as well. In addition, the Stutzer index is another performance measure that is slightly different, yet still relevant. The Stutzer index is founded on the behavioural hypothesis that investors aim to minimise the probability that the excess returns over a given threshold will be negative (Stutzer, 2000).

Performance measures based on lower partial moments (LPMs) include the Omega ratio and the Kappa measure. The Omega ratio expresses the ratio of the gains to losses with respect to a chosen (return) threshold (Keating & Shadwick, 2002) and it implicitly adjusts for both skewness and kurtosis in the return distribution. An Omega ratio conversion, the Omega-Sharpe ratio, generates ranking statistics that are in similar form to the Sharpe ratio and identical to Omega rankings. The Kappa measures, as introduced by Kaplan and Knowles (2004), generalises the Sortino and Omega ratios. Also of importance is the Sortino ratio, which is a natural extension of the Sharpe and Omega-Sharpe ratios, that uses downside risk in the denominator (see Sortino & van der Meer, 1991).

Alternative performance measures' compatibility with utility functions has also led to familiar generalisations of the Sharpe ratio. The generalised Sharpe ratio (GSR) (Hodges, 1998) is an extension of the Sharpe ratio and delivers equivalent fund rankings to the traditional Sharpe ratio when returns are normally distributed and the utility function is exponential. The advantage of the GSR is that its range of applicability extends to any type of return distribution while its main drawbacks being its restriction to exponential utility functions and that it requires an expected utility maximisation. The Adjusted Sharpe ratio (ASR), a natural extension compatible with utility theory, uses a Taylor series expansion of an exponential utility function to account for return distributions' higher moments (see Koekebakker & Zakamouline, 2008). Pezier and White (2006) further suggest making use of the ASR which explicitly corrects for higher moments by including a penalty factor for negative skewness and excess kurtosis.

Several of these alternative performance measures, however, fall short of having firm theoretical foundations (considering the Sharpe ratio is based on the expected utility theory) and do not permit accurate ranking of portfolio performance given that ranking based on these measures depends significantly on the choice of threshold. Most of these measures also only consider downside risk while the upside potential is not accounted for. Performance measures with a VaR foundation also have a number of problematic failings (Wiesinger, 2010). For instance, VaR is criticised for not being a coherent risk measure, as far as non-normal distributions are concerned, as it does not conform to the requirements³⁶, specifically to that of the sub-additivity property, and thus does not support diversification. Although VaR remains a popular measure of risk, it is sensitive to the underlying parameters and the employed calculation method whilst also relying on the risk factors being normally distributed, making this measure flawed in a hedge fund context. The Conditional Value-at-Risk (CVaR)³⁷-based Sharpe ratio, called the Conditional Sharpe ratio (CSR), overcomes essential standard deviation defects by replacing the Sharpe ratio's denominator (i.e., standard deviation) with CVaR. Not only is CVaR a coherent risk measure (Pflug, 2000), but it is also considered a more consistent measure of risk than VaR and can be used in risk-return analysis similar to the Markowitz mean-variance approach (Rockafellar & Uryasev, 2000).

³⁰ The Calmar ratio (CR) is the quotient of the excess return over risk-free rate and the maximum loss (i.e., maximum drawdown) incurred in the relevant period (Young, 1991).

³¹ The Sterling ratio uses the average of a number of the smallest drawdowns, within a certain time period, to measure risk (Lhabitant, 2004).

³² The Burke ratio expresses risk as the square root of the sum of the squares of a certain number of the smallest drawdowns (see Burke, 1994).

³³ Drawdown is defined as "the decline in net asset value from the highest historical point" (Lhabitant, 2004:55), and thus describes the loss incurred over a certain period of time (Wiesinger, 2010).

³⁴ The standard VaR only considers mean and standard deviation while modified VaR considers both the means and the standard deviation as well skewness and (excess) kurtosis.

³⁵ Artzner *et al.* (1997) introduced Conditional VaR (CVaR) to remedy against the shortcoming that VaR does not make a statement about the loss if VaR is exceeded.

³⁶ See requirements (of a coherent risk measure) proposed by Artzner *et al.* (1997).

³⁷ CVaR is also called mean excess loss, mean shortfall, or tail VaR.

3. METHODOLOGY AND DATA

3.1. Scaled Sharpe And Treynor Ratios

Traditional risk-adjusted performance measures, such as the Sharpe and Treynor ratios (Treynor, 1965; Sharpe, 1966), are founded on a Gaussian return assumption and a mean-variance efficient state. Asset returns, and specifically hedge fund returns, however, often violate the Gaussian assumption (Fung & Hsieh, 1997; Lo, 2001; Eling, 2006; Taleb, 2007) and hedge fund strategies' returns are known to exhibit (persistent) patterns of skewness and kurtosis (Eling & Schuhmacher, 2006).³⁸ Employing classic performance measures for performance assessment is therefore a biased approach as these classic measures do not account for return distribution's higher moments. For instance, standard deviation, as used in the denominator of the classic Sharpe ratio as a proxy for risk, does not appreciate positive skewness, which is commonly considered an attractive feature for a rational investor (see for example, Kraus & Litzenberger, 1976 and Kane 1982), but on the contrary penalises for it. Concerns pertaining to comparability in risk assessment and asset performance valuation thus produce a need for robust and reliable performance measures, which account for higher returns distribution moments – skewness at least, and kurtosis when possible. The scaled Sharpe and Treynor ratios, as used in this study, are adjusted modifications of these well-known performance measures to account, to some extent, for skewness and kurtosis that describe the deviations from normality. Thus, the classic Sharpe and Treynor ratios are adjusted for asymmetries in both the upside and downside deviations from the mean asset returns by weighting the upside and downside deviation risks. This accounting for skewness and kurtosis generally alters hedge fund performance ranking.

Adjustments to classical performance measures, to account for return asymmetries, remains relevant and further contributes to the literature on performance evaluation that takes non-normality of return distributions into account.

3.2. Data

A total of 26 496 monthly returns, net of management and performance fees,³⁹ from 184 'live' individual⁴⁰ hedge funds between January 2000 and December 2011 were used. These monthly fund returns were sourced from a EurekaHedge database data extract and funds with an incomplete monthly return history for the chosen period were not considered. As hedge funds universally report performance figures on a monthly frequency, this basis was used as it is also compatible with investors' month-end, holding-period return. Hedge fund databases can potentially suffer from several biases that may have a significant impact on performance measurement. The data do not suffer from the most common biases of the variety - survivorship, backfilling, or sampling - while selection bias cannot be dealt with as it would call for access to returns from hedge funds that decide not to report.

Summary statistics, in monthly percentages, for the hedge fund returns, as well as some other apposite information, is presented in Table 1. The *t*-statistics indicate that the mean returns are significantly different from 0 at the 5% significance level of all funds. Moreover, 29 out of the 184 funds (15.8%) show evidence of normal distributions at the 5% significance level, using the Jarque-Bera (JB) test, while the leftover 155 funds (84.2%) exhibit non-normal distributions.

³⁸ It is well known that hedge fund return distributions' deviations from normality are statistically significant Zakamouline (2011) and that hedge fund return distributions are negatively skewed with positive excess kurtosis (Eling, 2006; Eling & Schumacher, 2006; Taleb, 2007). According to Black (2006), skewness and kurtosis also reflect the event and liquidity risks taken on by hedge funds while Brooks and Kat (2002) highlight the high Sharpe ratios in the presence of negative skewness and positive excess kurtosis.

³⁹ Raw returns usually produce upward biased performance measures since fees tend to positively skew related performance measures. According to Wermers (2010), this drawback advocates the use of net-of-fees returns in that net returns represent a real performance proxy for hedge funds.

⁴⁰ Meaning not fund of funds, which are funds holding a portfolio of other investment funds, or commodity trading advisors (CTA), but funds that invest directly in securities.

Table 1: Summary Statistics For Long/Short Equity Hedge Funds

	All Funds	North America	Europe	Asia	Global
No. Of Funds	184	85	38	15	46
Sample Size	26 496	12 240	5 472	2 160	6 624
Mean Age (Years)	15.8	16.5	14.3	14.4	16.1
Mean Size (US\$m)	188	143	145	87	346
Return Statistics					
μ	0.66	0.76	0.55	0.34	0.66
$t (\mu = 0)$	22.48	16.14	11.49	3.92	10.64
σ	4.8	5.2	3.5	4.0	5.1
Median	0.6	5.2	0.6	4.0	0.6
Min	-56.7	-56.7	-20.0	-22.4	-54.7
Max	76.2	76.2	29.6	19.2	39.8
Skewness	0.75	1.14	0.49	-0.15	0.05
Kurtosis	18.4	22.3	10.0	4.9	9.6
ρ_1	0.29	0.21	0.74	0.43	0.21
ρ_2	0.03	0.15	0.59	0.31	0.23
ρ_3	0.02	0.01	0.55	0.29	0.21
p-value of LB-Q	0.00	0.01	0.00	0.00	0.01

The overall significance of the first k autocorrelation coefficients is measured by the Ljung-Box Q-statistic and is asymptotically χ_k^2 under the null hypothesis of no autocorrelation.

All of the funds included are categorised as long/short equity (strategy) funds. This strategy of fund was favoured as this particular strategy is the largest among hedge funds, comprising 35% of the industry (Brown, *et al.*, 2009). More recent figures, as at the end of November 2013, confirm that the long/short strategy is the most sought after as this strategy attracted US\$78bn of the US\$1.99tr that make up the total assets in the hedge fund industry (Eurekahedge, 2013). All funds are mandated only in highly liquid markets as funds mandated in developing markets were omitted from the sample - this ensured that funds are equity funds holding liquid instruments. Consequently, it can be assumed that all securities held have readily available prices and that no subjective valuations are required. This practice also minimises the stale price bias within the data sample (Géhin, 2006). As an analytic indication of liquidity, the first-order return autocorrelation (ρ_1) of all but two geographical areas are ≤ 0.30 (Getmansky *et al.*, 2004). The near zero levels of autocorrelation, for liquid securities such as equity funds, are also consistent with those found by Bisias *et al.* (2012).

An informational breakdown of the representative geographical mandates of the funds, as well as the relevant risk-free rate proxies accordingly used, are presented in Table 2. Data on the risk-free rates were sourced from the Federal Reserve Bank of St. Louis (FRED) and Bloomberg.

Table 2: Breakdown Of Geographical Mandates Of Funds & Risk-Free Rate Proxies

Geographical Mandate	# Funds	Risk-Free Rate Proxy
North America*	85 (46%)	10-year Treasury bond rate (US)
Europe	38 (21%)	10-year Treasury bond rate (Germany)
Asia	15 (8%)	10-year Treasury bond rate (Japan)
Global	46 (25%)	JPMorgan Global Government Bond Index

*Includes one Canadian fund (RFR = 10-year Treasury bond rate (Canada)).

As a proxy for the European geographical areas risk-free rate, the use of the German 10-year Treasury bond rate is generally accepted⁴¹ (Damodaran, 2008), although a number of alternative options exist.

⁴¹ Part of the logic for this practice being commonly accepted is that Germany is the largest issuer of bonds in the European geographical area.

Hedge funds are commonly weighed against passive benchmark⁴² indices,⁴³ even though hedge funds (particularly long/short strategy funds) are absolute investments. The data on the passive market benchmark indices were sourced from Bloomberg, whereas hedge fund benchmark indices were sourced from Eurekahedge, Hedge Fund Research (HFR), and Barclahedge. Table 3 exhibits the market and hedge fund benchmark indices used.

Table 3: Market And Hedge Fund Benchmark Indices

Benchmark Market Indices	Region Specific	
S&P500, S&P TSX*	North America	
DAX	Europe	
Nikkei 225	Asia	
MSCI World Index	Global	
Benchmark Hedge Fund Indices	Region Specific	Style Specific
Eurekahedge North America Long/short Equities Index	North America	Long/short Equity
Barclayhedge European Equities Index	Europe	Equities
Eurekahedge Asian Hedge Fund Index	Asia	-
Hedge Fund Research (HFR)(X) Global HF Index	Global	-

*The S&P TSX was used for the sole Canadian fund that forms part of the North American regional mandate.

The summary return statistics for the market and hedge fund benchmark indices for the period January 2000 to December 2011 are presented in Table 4. Table statistics are drawn from the monthly returns with the monthly means and standard deviations in percentages.

Table 4: Summary Statistics For Market And Hedge Fund Benchmark Indices

	S&P500	DAX	S&P TSX	Nikkei 225	Global Index ⁺	L/S HF Index [*]
Sample size	144	144	144	144	144	144
μ	0.004	0.12	0.35	0.39	0.28	0.76
$t(\mu = 0)$	0.01	0.21	0.92	0.81	0.06	3.78
σ	4.71	6.72	4.55	5.80	4.90	2.4
Median	0.60	0.73	1.01	0.13	1.17	0.99
Min	-16.9	-25.4	-16.9	-23.8	-25.48	-6.5
Max	10.8	21.4	11.2	12.9	14.06	10.6
Skewness	-0.43	-0.52	-0.86	-0.53	-1.42	0.01
Kurtosis	3.66	4.88	4.58	3.89	5.16	4.86
ρ_1	0.13	0.07	0.22	0.12	0.31	0.20
ρ_2	-0.07	-0.06	0.07	0.06	0.03	0.04
ρ_3	0.12	0.10	0.06	0.11	0.19	0.04
p-value of LB-Q	0.10	0.39	0.01	0.15	0.00	0.01

⁺ Global index = MSCI World Index.

^{*} L/S HF Index = Eurekahedge North America long/short Equities Index.

Both hedge fund and market indices exhibit non-normal distributions using the Jarque-Bera test at the 5% significance level.

3.3. Methodology

A 36-month rolling (window) period, beginning in January 2000, was used to estimate the relevant statistics and ratios. Monthly returns and risk-free rates were transformed to a geometric annualised basis using the 36-month rolling period.

⁴² Lhabitant (2004:116) defines the term benchmark as “an independent rate of return (or hurdle rate) forming an objective test of the effective implementation of an investment strategy”.

⁴³ Incipient hedge fund performance was not compared relative to a benchmark. According to Lhabitant (2004), hedge fund managers are hired for their skills and they should be allowed to venture wherever their value-creating instincts lead them, without considering benchmarks. Thus, hedge fund portfolios should aim to produce positive absolute returns rather than to outperform a particular benchmark.

The annualised Sharpe and Treynor ratios were calculated from monthly returns that are not independently and identically distributed (IID). According to Lo (2002) a computation bias arises when annual Sharpe ratios are computed from monthly means and standard deviation by multiplying by the square root time; in this case, $\sqrt{12}$, as monthly returns data are annualised. Lo (2002) continues that the method of computing annualised Sharpe ratios by multiplying by the square root of time is more suitable when returns are IID, but when returns are non-IID, an alternative procedure that considers serial correlation (of returns) must be used. It is also well established that hedge fund returns exhibit significant first-order auto-correlation (see Books and Kat, 2002) and this first-order auto-correlation introduces a serial dependence that, by itself, explains why returns are both non-identically distributed and non-normal. Thus, the IID normal assumption is not supported by hedge fund returns data and, although the assumption is often used, it can be described as “a convenient leap of faith that simplifies the math involved” (Bailey & Lopez de Prado, 2013). Also, the IID normal assumption is often said to be justified on a sufficiently large sample under Central Limit Theorems (CLTs) – this is false, as CLTs require either independence or at least weak dependence, and normality is also not evident over time in the presence of dependence. Although the measure proposed by Lo (2002), known as the $\eta(q)$ SR or annualised autocorrelation adjusted Sharpe ratio, is founded and its use advocated, this study does not employ it as the focus is fully on the scaling methodology concerning the named risk-adjusted performance measures that account for higher (returns distribution) moments. Purely for the purpose of illustrating the impact of the Lo (2002) annualised autocorrelation adjusted Sharpe ratio methodology, a selection of comparative summary statistics, using the 184 long/short equity hedge funds, is conveyed in Table 5. Note that the summary statistics in Table 5 are based on annualised geometric returns over a 36-month rolling period with the sole aim of presenting a statistical comparison between the annualised Sharpe ratio computation methods. For further details pertaining to the adjustment for non-IID returns, refer to Lo (2002).

Table 5: Comparative Sharpe Ratio Summary Statistics (All Figures Annualised)

	Sharpe Ratio	SC-adjusted Sharpe Ratio
Sample Size	20 056*	20 056
μ	0.38	0.41
σ	0.85	0.95
Median	0.26	0.25
Min	-2.1	-3.8
Max	3.5	5.1
Skewness	0.49	0.73
Kurtosis	2.86	3.92

*184 funds × 109 (144-35) monthly returns

Using the 36-month rolling method, monthly time-rolling annualised Sharpe and Treynor ratios were estimated in both traditional or classic and scaled forms for each fund and relevant market and hedge fund indices. Equation (1) was used to estimate the traditional or classic Sharpe ratio (Sharpe, 1966, 1975, 1992, 1994; van Vuuren *et al.*, 2003):

$$SR = \frac{r_p - r_f}{\sigma_p \sqrt{t}} \tag{1}$$

where r_p is the cumulative portfolio return measured over t months, r_f is the cumulative risk-free rate of return measured over the same period, and σ_p is the portfolio volatility (risk) measured over t months using the conventional standard deviation formula; namely:

$$\sigma_p = \frac{1}{T-1} \sum_{t=1}^T (r_t - \mu)^2 \tag{2}$$

where r_t is the portfolio return, measured at t -intervals over the full period under investigation, T , and μ is the average portfolio return over the full period. The scaled Sharpe ratio (SSR) was calculated using (Gatfaoui, 2012):

$$SSR = W_- \left(\frac{r_{p-} - r_f}{\sigma_{p-}} \right) + W_+ \left(\frac{r_{p+} - r_f}{\sigma_{p+}} \right) \tag{3}$$

where SSR is a skew-specific adjusted risk premium (SSRP) with S_- and S_+ being left-skew specific (LSSARP) and right-skew specific (RSSARP) adjusted risk measures respectively, thus effectively downside and upside Sharpe ratios, and r_{p-} and r_{p+} are monthly returns below and above the monthly arithmetic average return for the rolling 36-month period respectively.⁴⁴ Similarly, σ_{p-} and σ_{p+} represent the standard deviation of the returns as identified as either below or above the monthly arithmetic average return for the rolling 36-month period, $W_- = n_-/n$ and $W_+ = n_+/n$ are weights based on the monthly returns and the monthly arithmetic average return within the corresponding 36-month rolling period, and r_f is the risk-free rate. Upon the completion of classifying returns into either the upside or downside based on the 36-month rolling arithmetic average return, both upside and downside returns and standard deviations were estimated in a geometric annualised fashion using a 36-month rolling period.

The traditional Treynor ratio was estimated by using (Treynor, 1965):

$$TR = \frac{r_{p,t} - r_{f,t}}{\beta_p} \tag{4}$$

where r_p is the annualised portfolio return measured over t months, r_f is the annualised risk-free rate of return measured over the same period and β_p is the beta (systematic risk) of the portfolio using the conventional beta formula; namely:

$$\beta = \frac{cov(r_p, r_m)}{\sigma_m^2} = \frac{cov(r_p - r_f, r_m - r_f)}{\sigma_m^2} \tag{5}$$

The scaled Treynor ratio (STR) was calculated using (Gatfaoui, 2012):

$$\beta = \frac{cov(r_p, r_m)}{\sigma_m^2} = \frac{cov(r_p - r_f, r_m - r_f)}{\sigma_m^2} \tag{6}$$

with $\beta^* = \frac{W_{m-} \sigma_{pm-} + W_{m+} \sigma_{pm+}}{W_{m-} \sigma_{m-}^2 + W_{m+} \sigma_{m+}^2}$

where

$$\begin{aligned} \sigma_{m-}^2 &= \sum_{t=1, m_t \leq \mu_m}^n (m_t - \mu_m)^2 / n_{m-}, \\ \sigma_{m+}^2 &= \sum_{t=1, m_t > \mu_m}^n (m_t + \mu_m)^2 / n_{m+}, \\ \sigma_{pm-} &= \sum_{t=1, m_t \leq \mu_m}^n (p_t - r_f)(m_t - \mu_m) / n_{m-} \text{ and} \\ \sigma_{pm+} &= \sum_{t=1, m_t > \mu_m}^n (p_t - r_f)(m_t - \mu_m) / n_{m+}. \end{aligned}$$

$W_{m-} = n_{m-}/n$ and $W_{m+} = n_{m+}/n$ are weights, based on the monthly returns and the monthly arithmetic average return within the corresponding 36-month rolling period. Similar to the scaled Sharpe ratio, the monthly scaled Treynor ratio estimations are geometrically annualised, although portions of the estimation procedure are carried out using monthly returns and monthly arithmetic averages. Figure 4 presents a comparative illustration of the traditional and scaled versions of both the Sharpe and Treynor ratios.

⁴⁴ Returns equal to the (36-month rolling) monthly arithmetic average return are classified as r_{p-} .

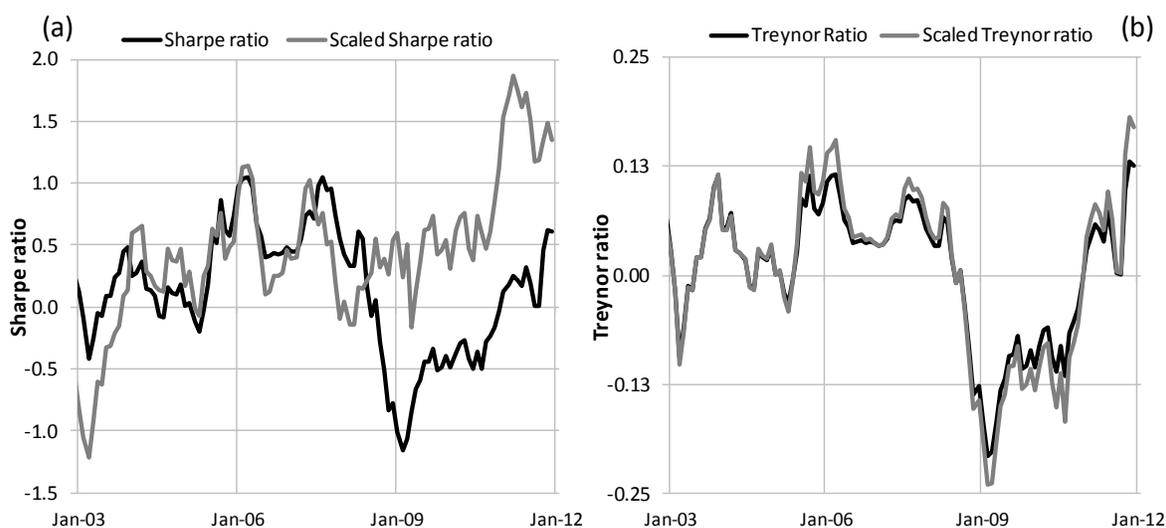


Figure 4: Comparative Illustration Of Traditional vs. Scaled Versions Of (a) Sharpe Ratio And (b) Treynor Ratio For Fund #109, A North American Fund

The subsequent section presents analysis and results by first highlighting how ill-suited the Sharpe ratio is for use within a hedge fund context due to the non-normality of hedge fund returns. The section will also explore comparative fund rankings between classic or traditional risk-adjusted measures and scaled versions of these measures that account for higher moments of the hedge fund returns distribution. To conclude, the section will present some comparative selective statistics over different economic phases.

4. ANALYSIS AND RESULTS

4.1. Inappropriateness Of The Sharpe Ratio (Non-Normal Returns)

Higher moment estimates of the returns data are presented in Table 6 which indicates that funds from all the geographical mandated areas exhibit, mostly positive, excess skewness (> 0.50), with the exception of globally mandated funds. Asian funds exhibit negative skewness. Table 6 also shows that the fund returns from all geographical areas are severely leptokurtic.

Table 6: Hedge Fund Higher Moment Estimates

	All Funds	North America	Europe	Asia	Global
Skewness	0.75	1.14	0.49	-0.15	0.05
S.E. Skewness (SES)	0.18	0.27	0.40	0.63	0.36
Kurtosis	18.40	22.29	10.01	4.87	9.58
S.E. Kurtosis (SEK)	0.36	0.53	0.79	1.26	1.44

According to the Jarque-Bera (JB) test, only 29 out of the 184 funds (15.8%) exhibit normal distributions at the 5% significance level, whereas the remaining 155 funds (84.3%) show evidence of having non-normal returns distributions. Figure 5 depicts the returns distribution’s state of normality for both the relevant market indices (Figure 5a) and the funds (Figure 5b) through time. Figure 5a and Figure 5b are both constructed using 36 months of rolling monthly data, whereas the thresholds for distribution normality at the 1% and 5% significance levels are represented by the two horizontal dotted-lines. Jarque-Bera (JB) test statistical values below these thresholds are indicative of normal distributions at the relevant level of significance.

Vertical lines are also used to partition Figures 5a and 5b into three periods or phases. Each of these three periods corresponds to a specific stage relating to the 2007 financial crisis: (1) *pre-crisis*, (2) *during* the crisis, and (3) *post-crisis* (i.e., after the height of the crisis). According to Figure 5a, some of the market indices pass the (rolling) goodness of fit test for normal return distributions, by means of the JB-test statistic, at either or both the 1%

and 5% significance levels (represented by the horizontal dotted-lines). The instances where some market indices do pass as normal distributions, however, only occur in limited cases and for short and limited time spans. Figure 5b shows that funds from all regional mandates are non-normal for the full-time period under investigation with the exception of Asian funds that exhibit return distribution normality, but only for November 2006 – this is, however, fairly insignificant considering the Asian funds are, on average, only deemed normal for 1 out of 109 rolling months. Also evident from Figure 5b is the rapid and elaborate increase (further) away from normality during 2008, along with the high non-normality for North American and European funds. By also comparing the average normality of funds for a specific regional mandate to its relevant market index, it is apparent that trends, trend changes and the magnitude of change do, for the most part, not coincide, while at certain times rather odd comparative behaviour is observed.

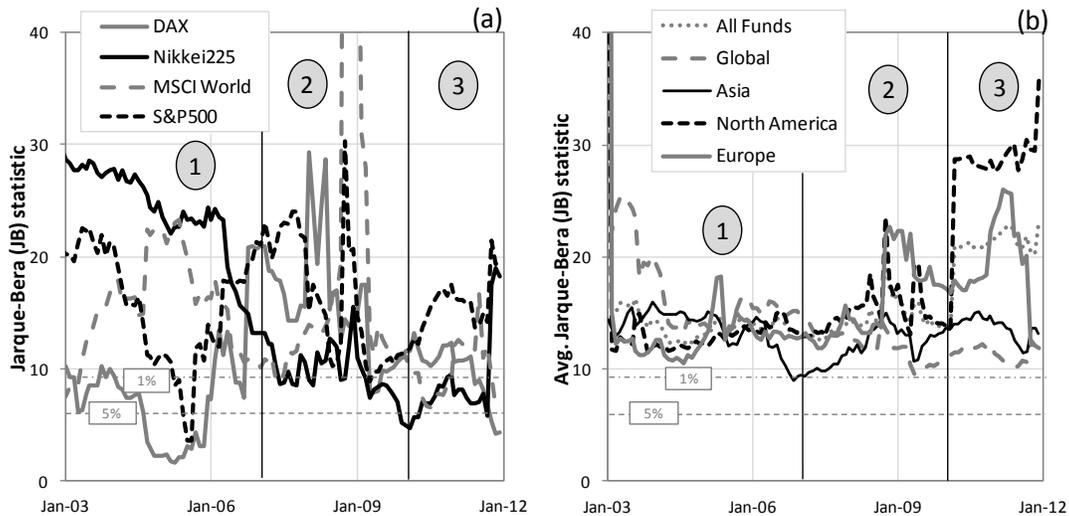


Figure 5: (a) Rolling JB-Test Statistic Of Relevant Market Indices And (b) Average Rolling JB-Test Statistic For All Funds And Also For Hedge Funds Per Geographical Mandate, Over Time

Figure 6a shows the skewness and Figure 6b the kurtosis of the funds grouped per geographic region.

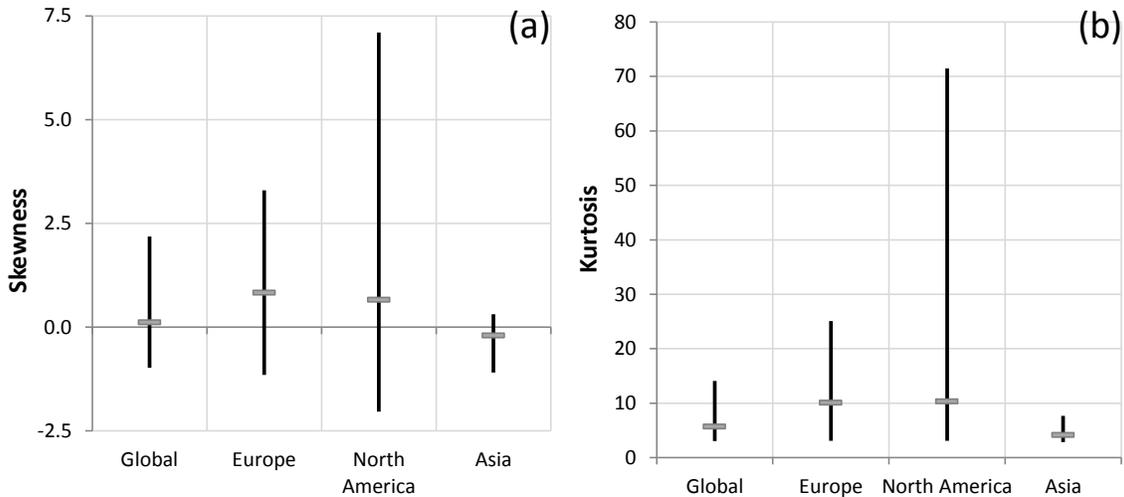


Figure 6: (a) Skewness Of Individual Funds Per Region And (b) Kurtosis Of Individual Funds Per Region

Figures 5 and 6, collectively with Table 6, confirm that most of the return distributions of these hedge funds are not ideally suited for Sharpe ratio application. The 15.8% (29 of 184) of funds that show evidence of

normal distributions, as per the JB-test, might be possible exceptions. However, this will require investors to test each fund for normality before applying the Sharpe ratio, which is far from ideal. To further reveal how ill-suited these funds’ return distributions are to Sharpe ratio application, not only at a point-in-time but also through time, the rolling skewness and kurtosis are presented in Figure 7. Using the 36-month rolling period, Figure 9 shows the average skewness (Figure 7a) and kurtosis (Figure 7b). Figure 7 is also partitioned into three periods by way of vertical lines – each period again representing a specific period relating to the 2007 financial crisis, consistent with those declared earlier (see Figure 5).

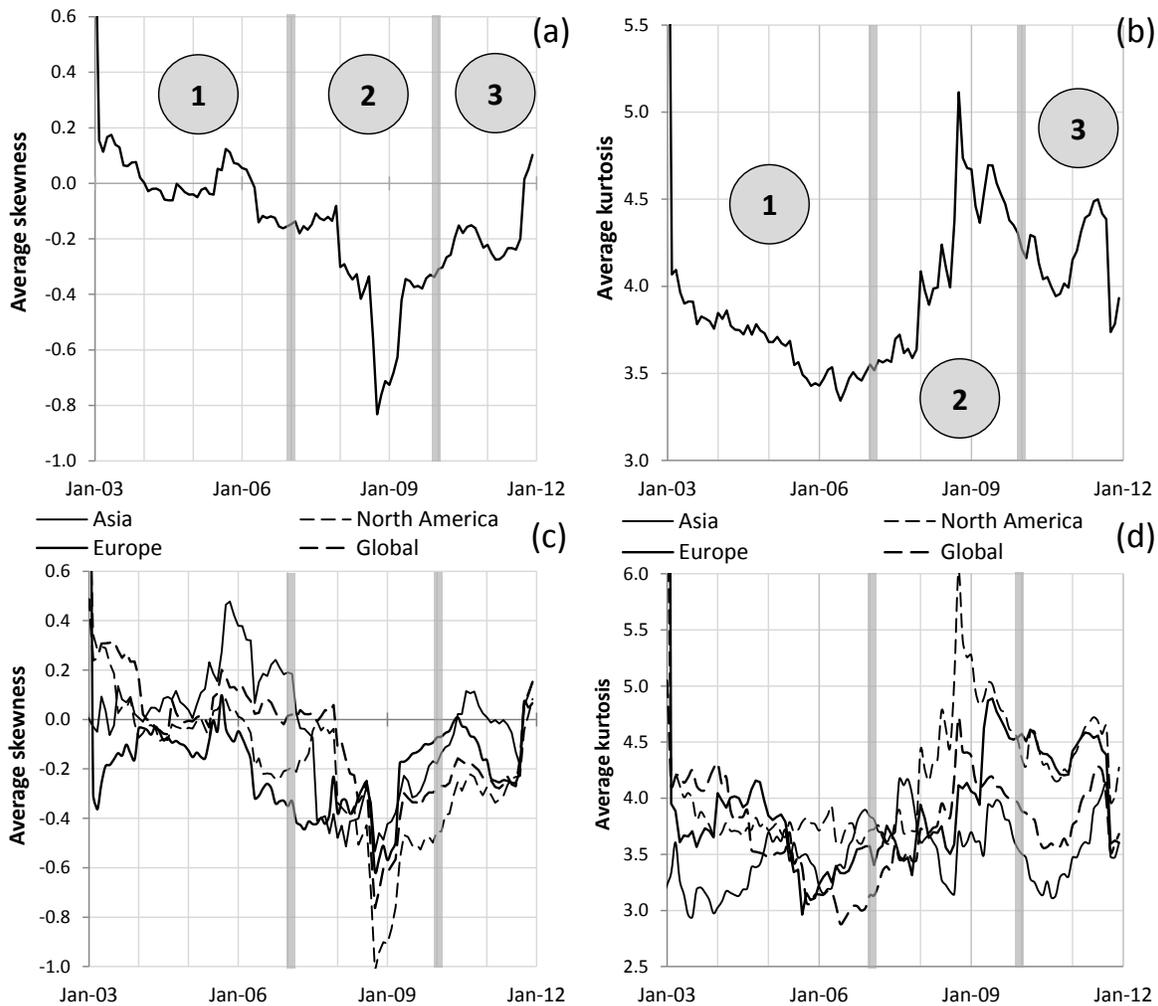
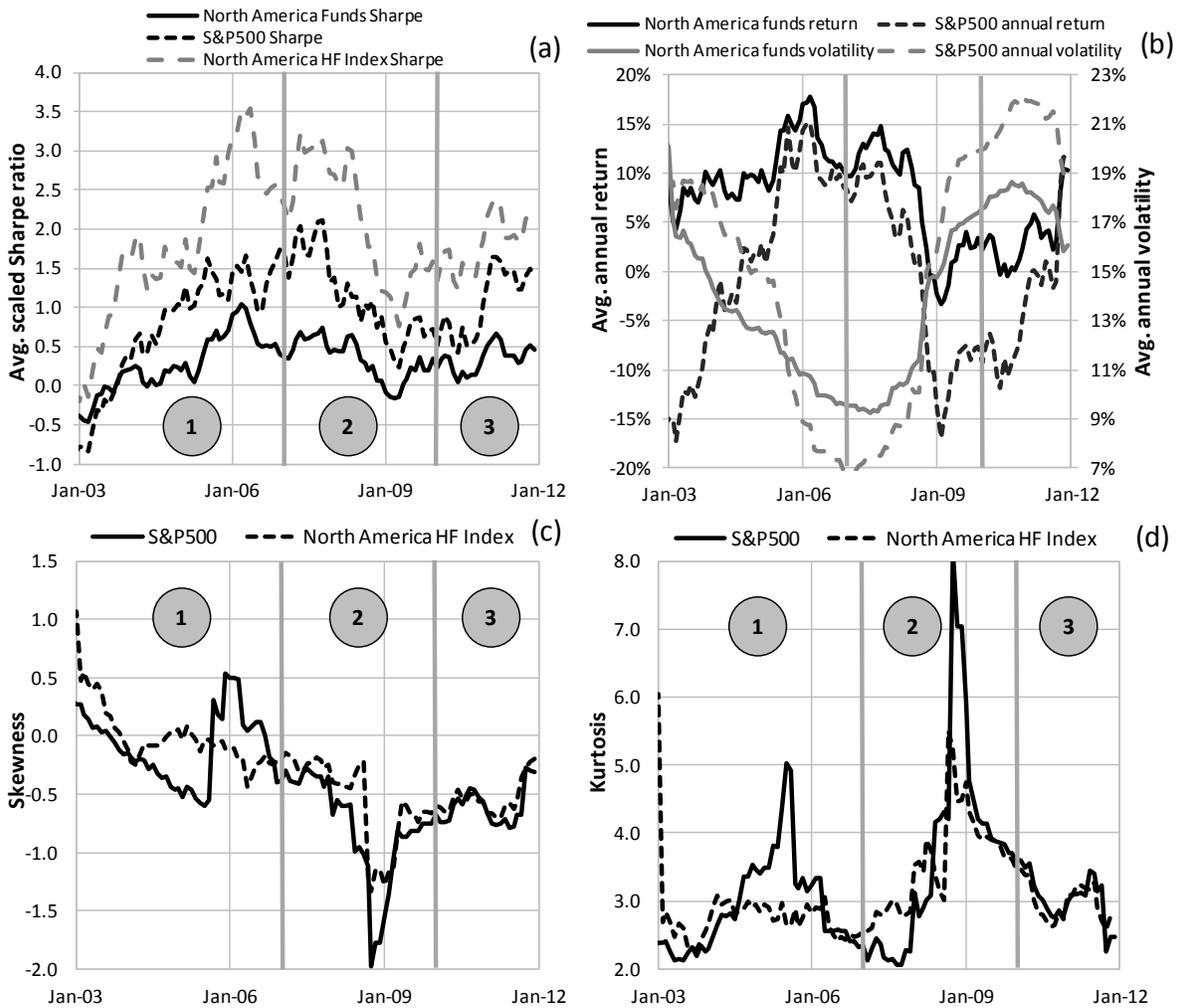


Figure 7: Average Values, Through Time, For (a) Skewness – All Funds, (b) Kurtosis – All Funds, (c) Skewness – Funds Per Region, And (d) Kurtosis – Funds Per Region

Figure 7 shows that during the 2007 crisis period, the average skewness turned considerably negative, whereas average kurtosis, which was at high level prior, reached extreme levels. Figure 7 can thus be added to Figures 5 and 6 and Table 6, thereby strengthening the case that the (traditional) Sharpe ratio is not adequately compatible with the return distributions of these hedge funds, as these distributions exhibit non-normal characteristics.

When considering a specific geographic region - for example, North America as presented in Figure 8 - the relevant statistics also indicate to the non-normality of returns for North American funds as well as the North American market index (S&P500) and the North American hedge fund index. Figure 8 was constructed using the rolling period analysis method and statistics are presented on an annual basis.



* Hedge fund index in Figure 10 = Eurekahedge North America Long/Short Equities Hedge Fund Index.

Figure 8: (a) Average Scaled Sharpe Ratio: North America Funds vs. S&P500 vs. North America HF Index, (b) Average Annual Return And Volatility: North America Funds vs. S&P500, (c) Skewness: S&P500 vs. North America HF Index, And (d) Kurtosis: S&P500 vs. North America HF Index

The higher moments of the hedge fund benchmarks, as depicted in panels (c) and (d) of Figure 8, also indicate the inappropriateness (of these return distributions) for the use of the Sharpe ratio. Panels (c) and (d) also indicate the altered behaviour for these higher moments of the return distribution around the time period of the recent financial crisis. The financial crisis also impacted the returns of these funds along with their volatility (Figure 8b). Figure 8b shows the decline in average returns and the increase in average volatility for both these mandated funds and the S&P500 during the crisis time period. Figure 8a presents the average scaled Sharpe ratios, specifically for the funds with North America mandates, along with the scaled Sharpe ratios for relevant benchmarks. The average scaled Sharpe ratio for the funds with North America mandates are relatively lower compared to those of both the market and hedge fund indices. Figure 8a also shows that the funds and benchmarks follow a similar trend over time and that during the crisis period, a decline in the trend is obvious.

4.2. Comparative Performance Measurement: Traditional vs. Scaled

This section presents comparative rankings of the sample of hedge funds using both the traditional and scaled Sharpe and Treynor ratios at different points of economic activity, seeing that investors frequently use rankings to differentiate between potential fund investments from less promising fund investments. The emphasis is on comparing the rankings of the traditional measure to those of the scaled measure within each type of measure (Sharpe and Treynor).

The (36-month) rolling Sharpe and Treynor ratios are again used, as described earlier, and three points-in-time were selected in accordance with the identified phases. Phase 1 (*pre-crisis*) is represented by December 2006, phase 2 (*during*) by December 2009 and phase 3 (*post*) by December 2011. Points-in-time are used since a static point produces an easier and more stable method to work with rankings and also as static point-in-time methods are most commonly used in practice when considering the ranking of funds. Owing to space constraints the top and bottom 25 funds in the sample are identified at December 2009 (i.e. *during* the crisis period) according to either the traditional Sharpe or Treynor ratio, and then ranked backwards and forwards in time within the full fund data sample of 184 funds.

4.2.1. Traditional vs. Scaled Sharpe Ratio Rankings

The comparative traditional and scaled Sharpe values, as well as rankings for the top and bottom 25 funds for the three economic phases, are presented in Figure 9 which indicates the shift in fund performance during the crisis period as opposed to prior, as a division is apparent between strong (best) and weak (worst) performing funds. Although the funds are scattered fairly randomly during the period prior to the crisis, both the traditional and scaled Sharpe ratios respectively value (Figure 9a) and rank (Figure 9d), the best performing funds a bit higher than the worst performing funds. Phase 3 shows that the scaled Sharpe ratio both value (Figure 9c) and rank (Figure 9e) a large number of the worst performing funds higher (better) than the traditional Sharpe ratio - this is more obvious for the ranking than the valuation. During the pre- and post-crisis periods the best performing funds are distinguishable from the worst performing funds, again, more in rank than in value. Discrepancies (in terms of risk-adjusted values) between funds are smaller for the periods prior and after the crisis compared to the period during the crisis - these discrepancies are not only observable between all funds, but even more stressed between the best and worst performing funds.

In the phase 2 period, a clear distinction is apparent between the best and worst performing funds while also from Figure 9b (representing the phase during the crisis), a large contingent of the best performing funds cluster around values between 0.5 and 1 for both the traditional and scaled Sharpe ratios – just rewarding investors with an equal return for the amount of risk taken on-board. Hence, even the top performing funds did not deliver exceptional risk-adjusted performance compared to performance expectations during normal economic conditions. Still, during this period, this level of performance would be classified as exceptional – and thus the rationale that these funds are the top funds during phase 2.

During the crisis period, (Figure 9b and 9e) both good and bad performing funds were identified valued. For this period, there is also no clear indication of any relationship between the valuing and or ranking rationale of the traditional and scaled Sharpe ratios. Period 2 does show that the traditional Sharpe ratio tends to value funds somewhat higher compared to the scaled Sharpe ratio. This lower risk-adjusted valuation by the scaled Sharpe ratio, in comparison to the traditional Sharpe ratio, makes perfect sense as the scaled ratio accounts for the increased risk due to the higher levels of skewness and kurtosis that characterised the period (see Figures 5b and 7) and also Figure 10 as a further example. Some discrepancies do, however, exist between how the traditional and scaled Sharpe ratios rank and value funds.

It is moreover apparent that some of the top funds during the crisis performed badly prior to the crisis and *vice versa*. This suggests that investors would possibly not have selected these funds prior to the crisis due to mediocre or weak risk-adjusted performance, and yet these funds performed the best during the crisis. Compare, for example, the positioning of the indicated fund (fund #167 as the larger datum point) in Figures 9a, 9b and 9c. The performance of this particular fund (fund #167) deteriorated in phase 3 to an even lower (risk-adjusted) level than it recorded in phase 1.

The comparative fund rankings based on the traditional and scaled Sharpe ratios for phases 1 to 3 are presented in Figures 9d, 9e and 9f, respectively. The numbers next to the data points are the (traditional Sharpe, scaled Sharpe) rank coordinates.

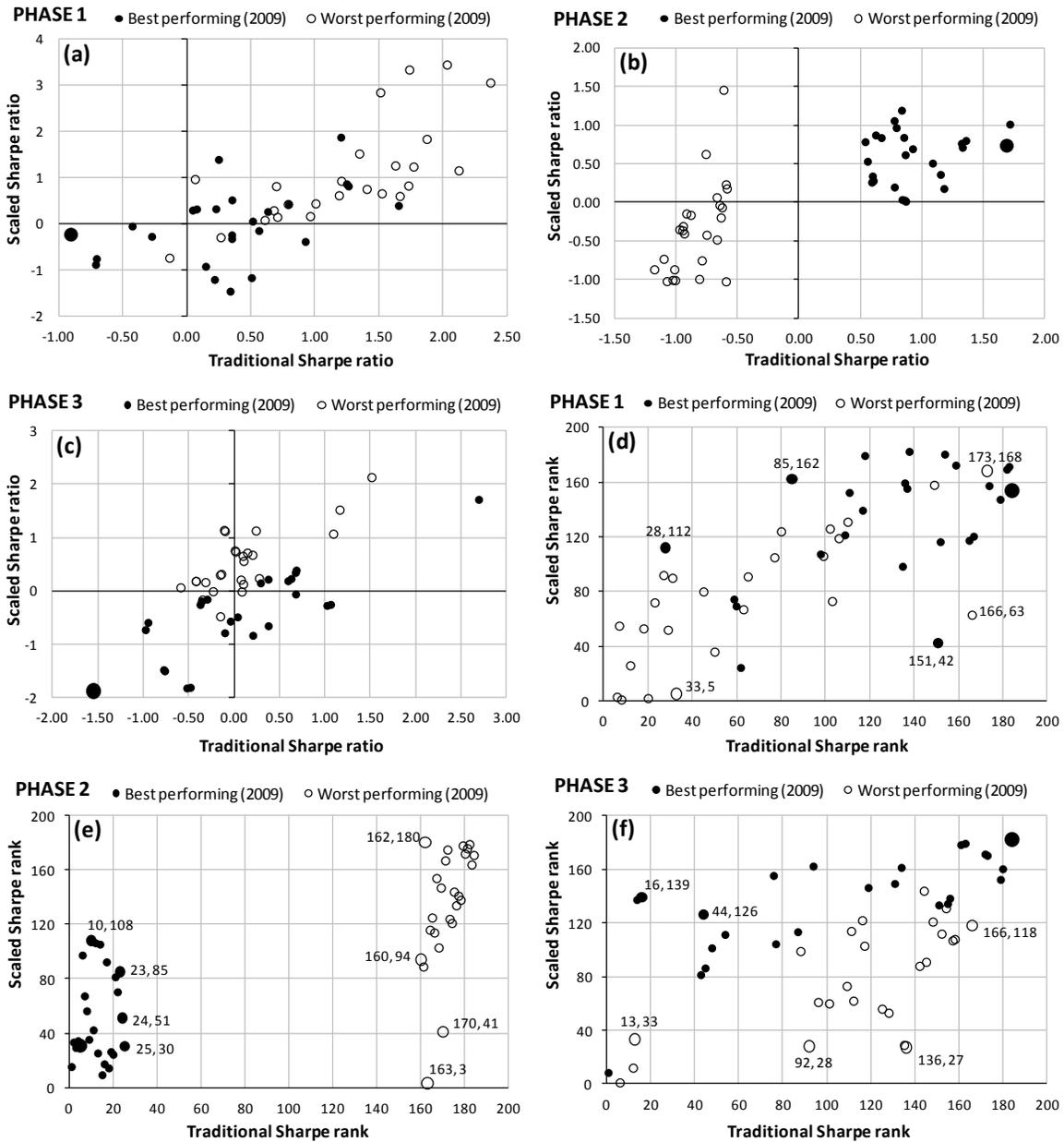


Figure 9: Traditional Vs. Scaled Sharpe Ratio Values For The Top And Bottom 25 Funds In The Sample For (a) Phase 1, (b) Phase 2, And (c) Phase 3. Traditional vs. Scaled Sharpe Rank For The Top And Bottom 25 Funds In The Sample For (d) Phase 1, (e) Phase 2, And (f) Phase 3

In the period prior to the financial crisis, a wider discrepancy between the traditional and scaled Sharpe ratios existed than compared to the period during the crisis (this phenomenon is, to some extent, reinitiated in the post-crisis phase). From Figure 9e, it is clear, although now according to fund rankings, that there was a clear distinction between the best and worst performing funds during the crisis period (phase 2). During the crisis period (phase 2), the traditional and scaled Sharpe ratios also generally ranked the funds similarly – meaning that both

measures ranked the best performing funds higher than the worst performing funds. This is evidenced by the majority of the best performing funds being located in the southwest region and the bulk of the worst performing funds in the northeast region of Figure 9e. A small number of exceptions (see funds in southeast region of Figure 9e) are observed – fund #34 is ranked as low - as 163rd by the traditional Sharpe ratio - while the scaled Sharpe ratio ranks this same fund as high as 3rd. There are also instances in which the two ratios do not rank coherently, as fund #144 is, for example, ranked 10th by the traditional Sharpe ratio but a distant 108th by the scaled Sharpe ratio. Figure 10 shows the characteristics, particularly the skewness (Figure 10b) and kurtosis (Figure 10c), of fund #144 that contribute to this particular fund being both valued, on a risk-adjusted basis, and ranked much lower (worse) by the scaled Sharpe ratio compared to the traditional Sharpe ratio in phase 2, as shown in Figure 10d.

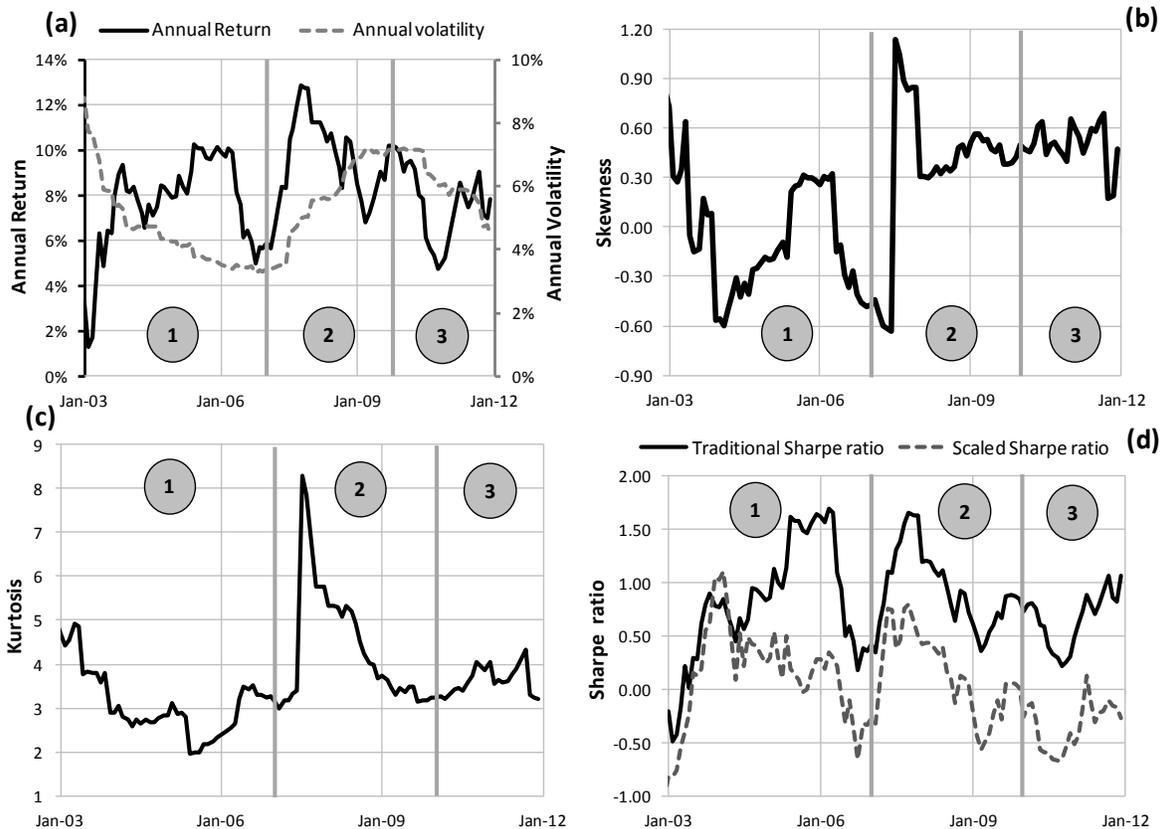


Figure 10: Fund #144 (a) Annual Return vs. Annual Volatility, (b) Skewness, (c) Kurtosis, And (d) Traditional vs. Scaled Sharpe Ratio

This analysis of fund #144, along with Figures 9e and 10, present a clear and logical case that the scaled Sharpe ratio incorporates the risk contained in the return distribution’s higher moments (skewness and kurtosis), which are crucial to hedge fund performance analysis.

The comparative rankings reiterate the point made earlier that selecting a highly ranked fund prior to the crisis resulted in a weak (low) rank for the same fund during the crisis period – as some of the worst ranked funds during the crisis are ranked rather high in the period prior to the crisis. As an example of this phenomenon, see and compare the ranking position of the indicated fund (fund #167 indicated by larger datum point) in Figures 9d, 9e and 9f.

In conclusion, the scaled Sharpe ratio is arguably an improvement on the traditional Sharpe ratio as far as non-normal return distributions are concerned, while it can be considered as a measure that should augment the use of the traditional Sharpe ratio.

4.2.2 *Traditional vs. Scaled Treynor Ratio Rankings*

Figure 11 presents the comparative traditional and scaled Treynor values and rankings for the top and bottom 25 funds across the three economic phases. Although the Treynor and Sharpe ratios are not directly comparable⁴⁵, a number of similar observations are apparent. For instance, similar to the Sharpe ratio analysis (see Figure 9), Figure 11 also signifies a shift in fund performance during the crisis period as opposed to prior, while from Figure 11e, a marked division between strong (best) and weak (worst) performing funds is apparent.

Another observation similar to the Sharpe ratio case is that during the pre-crisis period, both the traditional and scaled Treynor ratios respectively value and rank the best performing funds higher than the worst performing funds. Interestingly, the scaled Treynor ratio both value (Figure 11c) and rank (Figure 11e) a number of the worst performing funds lower (worse) than the traditional Treynor ratio – this phenomenon is more noticeable for the ranking. When considering traditional Treynor ratio's valuation, the pre- and post-crisis periods show a degree of clustering among the best performing funds, while the valuation by the scaled Treynor ratio tends to be more dispersed and, also during phase 1, lower than that of the traditional measure. Phase 2 (the period during the crisis) shows a wider dispersion of funds, in terms of valuation for both measures (Figure 11b), while it is also observed that both measures value a large number of funds below 0 – the latter is more evident from the perspective of the traditional Treynor ratio. Thus, during phase 2, the scaled Treynor ratio values the best performing funds slightly higher than the traditional measure. Figures 11a, 11b and 11c show the discrepancies, in terms of risk-adjusted values, to be smaller between funds prior and after the crisis compared to during the crisis period – these value differences are more apparent when comparing the best and worst performing funds.

Phase 2 delivers a clearer distinction between the best and worst performing funds. A large group of the best performing funds are valued just above 0 while a large group of funds are valued between 0 and -2, by both the traditional and scaled Treynor ratios. The former is indicative that during the crisis period, most of the best performing funds rewarded investors with an amount just exceeding the parallel amount of reward for the amount of accepted risk. So, according to Treynor ratios, even the top performing funds did not deliver exceptional performance compared to performance expectations during normal economic conditions – during this exacting period, however, this level of performance would be classified as exceptional and therefore why these funds are the top funds during phase 2. In addition, it is interesting to note that during phase 3, the majority of funds, especially the best performing funds, delivered even lower Treynor ratios than during phase 2, according to both Treynor measures.

No pattern is evident between the Treynor valuation of the traditional or scaled measures and some valuation and ranking discrepancies between the two Treynor measures do exist.

Some of the top performing funds during the crisis performed poorly prior to the crisis and *vice versa*. This suggests that these funds would not have been the choice of investors prior to the crisis due to mediocre or weak risk-adjusted performance; however, these performed the best during the crisis. The larger datum point of fund #167 is, once again, an example of this and provides an opportunity to compare the positioning of the fund through the three phases (see Figures 11a, 11b and 11c).

Respectively, 11d, 11e and 11f present the comparative fund rankings derived from the traditional and scaled Treynor ratios for phases 1 through 3. The numbers next to the data points are the (traditional Treynor, scaled Treynor) rank coordinates.

⁴⁵ Or stated differently, accounting for different types of risk – total vs. systematic (market)

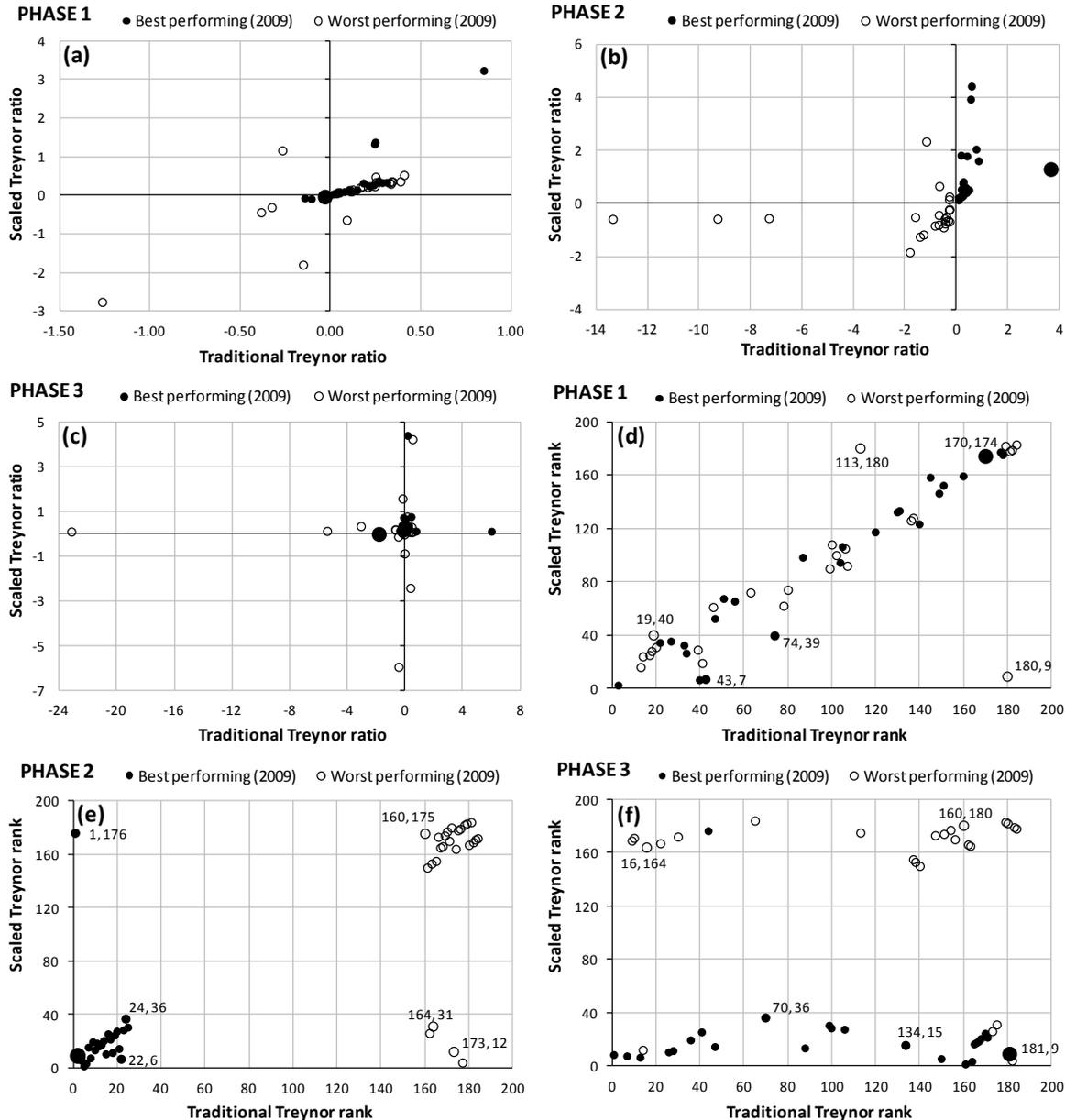


Figure 11: Traditional Treynor Ratio Vs. Scaled Treynor Ratio Values For The Top And Bottom 25 Funds In The Sample For (a) Phase 1, (b) Phase 2, And (c) Phase 3. Traditional Treynor Ratio Vs. Scaled Treynor Ratio Rank For The Top And Bottom 25 Funds In The Sample For (d) Phase 1, (e) Phase 2, And (f) Phase 3

In terms of ranking, the period prior to the crisis exhibits a random, though somewhat analogous, ranking between the traditional and scaled Treynor measures, as the fund rankings are located in straight line along the diagonal. Moving from phase 1 to phase 2, a noticeable shift in performance is clear and a very clear distinction between the best and worst ranked funds exist during the crisis period (see Figure 11e). Also evident is that during the crisis period, the traditional and scaled Treynor ratios generally ranked the funds similarly – meaning that both measures predominantly ranked the best performing funds higher (better) than the worst performing funds. This is, as with the Sharpe ratio analysis, evidenced by the bulk of the best performing funds being located in the southwest region of Figure 11e, whilst the majority of the worst performing funds are located in the northeast region. A number of exceptions are, however, observed – see funds in the northwest and southeast regions of Figure 11e. Fund #148 is an example of this exceptional phenomenon and is ranked 164th by the traditional Treynor ratio while being

ranked as high as 31st by the scaled Treynor ratio. Very interesting is that in the period after the crisis (phase 3), the best and worst performing funds remain significantly distinct from each other - the worst performing funds are located in a horizontally spread manner in the northern region of Figure 11e while best performing funds are to be found in a similar horizontal spread, but in the southern region. The difference between this fund distinction during phase 3, from that of phase 2, is that the traditional and scaled Treynor ratios do not rank the funds similarly as either high (best) or low (worst). During this phase - the period after the crisis - there seems to exist an exceptionally random ranking of both the best and worst performing funds by both measures with no particular ranking pattern or relationship. The only observation that is obvious is that the scaled Treynor ratio almost wholly ranks the best performing funds high (best) and the worst performing funds as low (worst).

As with the Sharpe ratio analysis, the comparative Treynor rankings repeat the point that selecting a highly ranked fund prior to the crisis resulted in a weak (low) rank for the same fund during the financial crisis period – as some of the worst ranked funds during the crisis are ranked rather highly in the period prior to the crisis. See and compare the ranking position of the indicated fund (#167 indicated by the larger datum point) in Figures 11d, 11e and 11f as an example of this phenomenon. Refer to Appendix 4, Figures 4A and 4B which detail the comparative traditional and scaled betas, over time, for the Treynor ratio.

4.3 Selective Statistics Over Different Economic Conditions

This section presents some selective summary performance statistics focusing on returns and scaled Sharpe and Treynor ratios for both the hedge funds and the relevant market benchmarks. Refer to the Appendix for selective summary statistics concerning both the traditional (i.e., unscaled) and scaled versions of the Sharpe and Treynor ratios with the following breakdown: Appendix 1A shows the summary statistics for the traditional Sharpe and Treynor ratios for all hedge funds and Appendix 1B shows the traditional Sharpe and Treynor ratios for hedge fund grouped geographically. Appendix 2A details the statistics for traditional Sharpe and Treynor ratios for the relevant market indices. Appendix 3A contains the summary statistics for returns, as well as scaled Sharpe and Treynor ratios, for the regional hedge fund indices, while Appendix 3B presents the traditional Sharpe and Treynor ratios for these hedge fund indices. The statistics are partitioned into three phases to highlight the altering characteristics of the funds and the relevant benchmarks throughout different economic periods. The three phases represent the periods *prior*, *during*, and *post* the 2007 financial crisis. January 2002 until December 2006 constitutes phase 1, January 2007 until December 2009 phase 2, and January 2010 until December 2011 phase 3. The rolling annual calculation methodology based on 36 months, as discussed earlier, is employed in this section.

Figure 12 exhibits the changing characteristics during the three economic phases through the average annual returns and standard deviation for all funds. Panel (a) of Figure 12 shows the average scaled Sharpe ratio, and panel (b) the average scaled Treynor ratio, for all funds in this study. The summary statistics for all funds per phase is conveyed in Table 7.

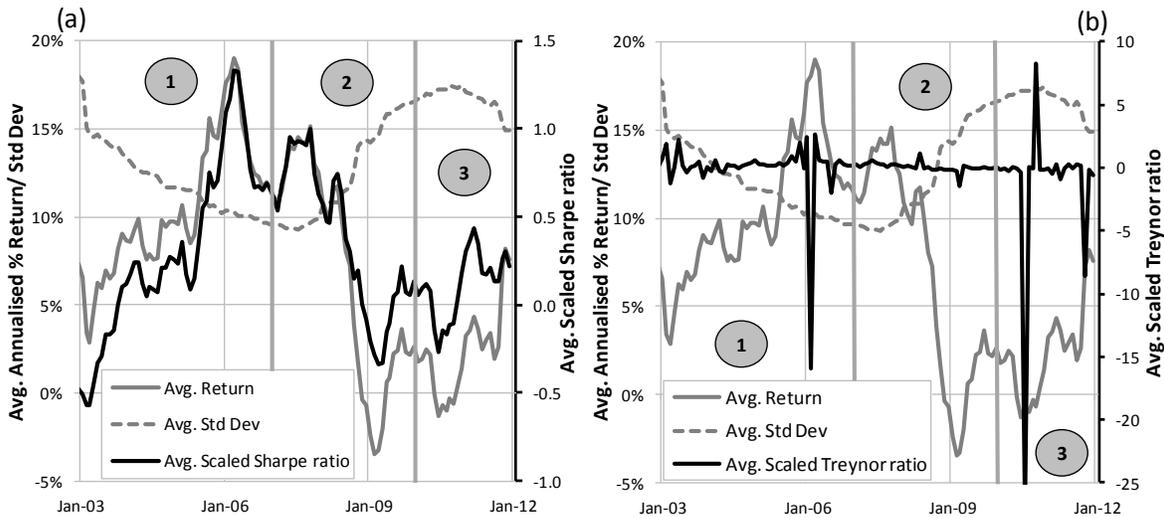


Figure 12: Average Annual Return And Standard Deviation And (a) Scaled Sharpe Ratio And (a) Scaled Treynor Ratio, For All Hedge Funds

The impact of the 2007 financial crisis is seen in Figure 12 through a decrease in the average scaled Sharpe ratio of all funds during phase 2. Figure 12 also shows the decrease in the average return across all funds in tandem with a near simultaneous increase in volatility during the crisis period. Of further significance is that mostly during phase 2, the period *during* the 2007 financial crisis, the average scaled Sharpe ratio of all funds reduces to below zero which implies that a risk-less asset would have performed better, on average, during this time compared to the analysed funds sample.

The scaled Treynor ratio, as presented in Figure 12b, does not provide any information of extraordinary surprise as the scaled Treynor ratio is relatively constant at the same level throughout, apart from the few observed extreme values. The results in Figure 12 are reflected in the summary statistics in Table 7, which indicates a similar trend between the average scaled Sharpe and Treynor ratios of a slight increase from phase 1 to phase 2 and thereafter a (relative larger) decrease moving from phase 2 to phase 3. Table 7 shows that the average return of all funds decreases over time while the standard deviation of returns and the scaled Sharpe ratio also reduce over time, indicating a diminishing performance spectrum between funds, on average.

Table 7: Summary Statistics For All Hedge Funds Per Phase

	Phase 1	Phase 2	Phase 3	Phase 1	Phase 2	Phase 3	Phase 1	Phase 2	Phase 3
	Return Statistics			Scaled Sharpe Ratio			Scaled Treynor Ratio		
<i>n</i>	9016	6624	4416	9016	6624	4416	9016	6624	4416
μ	10.41%	6.86%	2.39%	0.31	0.56	0.11	0.01	0.06	-1.25
σ	10.93%	10.20%	8.38%	1.06	1.00	0.82	33.16	4.22	81.70
Median	9.59%	7.33%	2.21%	0.27	0.50	0.15	0.14	0.04	-0.05
Min	-44.96%	-48.39%	-36.57%	-2.80	-2.31	-2.55	-2954.2	-260.49	-4925.3
Max	59.50%	42.39%	74.39%	5.33	4.53	2.60	474.84	115.97	1545.35

Figure 13 presents the average scaled Sharpe ratios of both funds and their relevant market benchmarks, per region. From this figure, it is apparent that funds and also market benchmarks from the included regions behaved similarly across the three phases. From Figure 13a, it is interesting that funds from most regional mandates exhibit significant scaled Sharpe ratio increases and decreases at similar times. None of the regional funds indicate significant better performance than any other time during or post the financial crisis. Asian funds, however, performed better, on average, shortly prior to the crisis, while European funds also enjoyed a short period of superior risk-adjusted performance during the crisis period (phase 2) (Figure 13a). Figure 13b shows that the market benchmark for global mandated funds (MSCI World Index) performed better than other included market

benchmarks for a period prior to the crisis and also during the crisis, while the market benchmark for Asian funds (Nikkei225) generally underperformed other market benchmarks, especially for the period during the crisis.

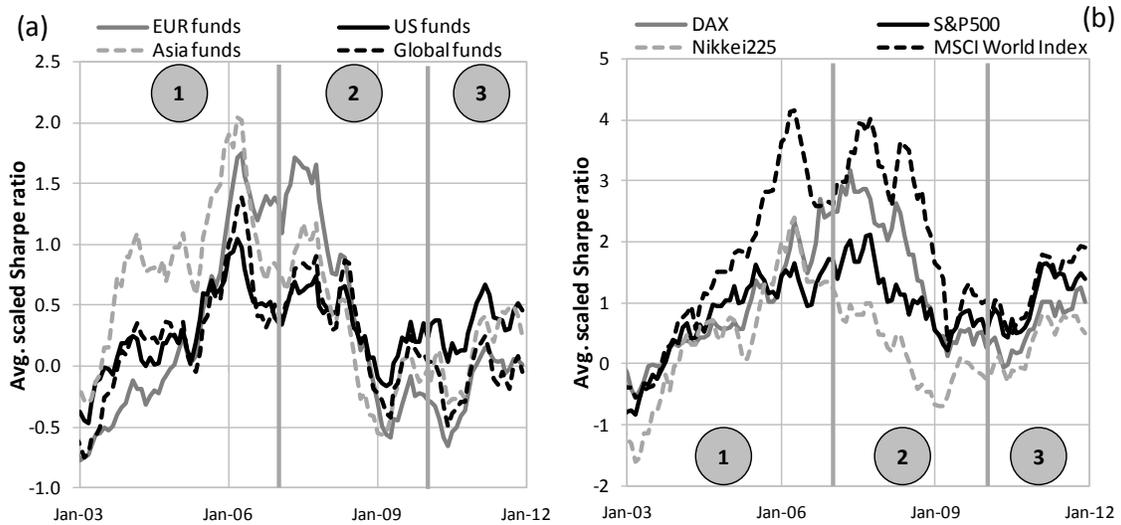


Figure 13: (a) Average Fund Scaled Sharpe Ratios Per Region And (b) Average Market Index Scaled Sharpe Ratios, Through Time

The average scaled Treynor ratios of both funds and their relevant market benchmarks, per region, are presented in Figure 14. From Figure 14a, it is once again apparent that funds from all the regions behaved in a similar manner across the three phases. Similar to Figure 13a, Figure 14a indicates that funds from most regional mandated areas exhibit risk-adjusted increases and decreases at similar times. These points of significant scaled Treynor ratio increases and decreases are thus very similar to those of the scaled Sharpe ratio. Interestingly the general pattern through time, as exhibited by the scaled Sharpe and Treynor ratios, are also very similar. Similar to the average scaled Sharpe ratios, none of the regional funds indicate significant better performance than any other during or post the financial crisis. Asian funds, however, show slightly lower risk-adjusted performance during the post phase (phase 3) (Figure 14a). Figure 14b shows a number of scaled Treynor ratio extreme values at various points through time from different regional market benchmarks with relative stable results at a constant level otherwise. The period prior to the crisis - phase 1 - indicates more scaled Treynor ratio volatility, mostly from the European funds' market benchmark (DAX).

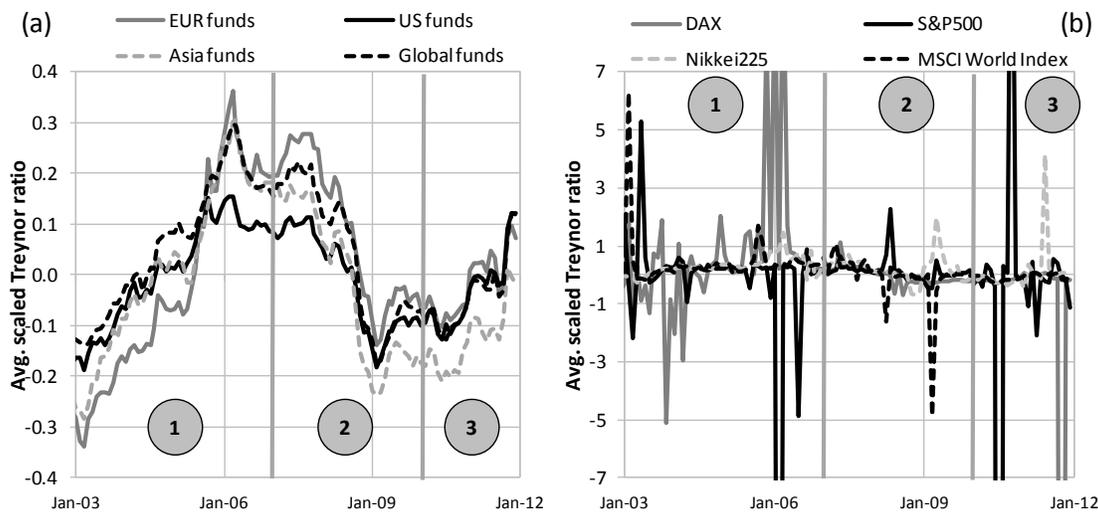


Figure 14: (a) Average Fund Scaled Treynor Ratios Per Region And (b) Average Market Index Scaled Treynor Ratios, Through Time

To facilitate comparisons, the summary statistics in terms of fund returns, scaled Sharpe and Treynor ratios, grouped by regional mandates, are presented in Table 8. Table 9 presents the corresponding summary statistics for the relevant regional market benchmarks.

Table 8: Summary Statistics For Regionally Grouped Hedge Funds Per Phase

	Phase 1	Phase 2	Phase 3	Phase 1	Phase 2	Phase 3	Phase 1	Phase 2	Phase 3
	Return Statistics			Scaled Sharpe Ratio			Scaled Treynor Ratio		
North American Hedge Funds									
<i>n</i>	4165	3060	2040	4165	3060	2040	4165	3060	2040
μ	10.52%	6.90%	3.45%	0.27	0.34	0.34	-0.58	0.14	-1.87
σ	11.16%	10.15%	9.11%	0.96	0.80	0.85	46.96	3.27	115.18
Median	9.61%	7.61%	2.42%	0.25	0.34	0.38	0.12	0.03	-0.03
Min	-44.96%	-48.39%	-36.57%	-2.80	-2.31	-2.44	-2954.2	-49.87	-4952.3
Max	59.50%	36.52%	68.29%	3.76	3.75	2.60	443.46	115.97	1545.35
European Hedge Funds									
<i>n</i>	1862	1368	912	1862	1368	912	1862	1368	912
μ	8.30%	7.23%	1.52%	0.27	0.54	-0.17	0.74	-0.03	-1.83
σ	9.49%	9.20%	6.68%	1.28	1.22	0.63	17.26	1.33	51.49
Median	7.42%	7.17%	1.98%	0.05	0.33	-0.13	0.15	0.02	-0.08
Min	-23.33%	-16.56%	-18.82%	-2.52	-1.80	-1.83	-240.90	-24.25	-1555.1
Max	42.97%	37.65%	32.60%	5.33	4.53	2.13	474.84	30.70	4.21
Asian Hedge Funds									
<i>n</i>	735	540	360	735	540	360	735	540	360
μ	11.32%	3.73%	-1.11%	0.85	0.28	0.13	0.31	0.22	0.09
σ	11.07%	11.58%	7.15%	0.87	0.99	0.54	0.94	2.15	3.28
Median	10.31%	3.46%	-0.53%	0.90	0.08	0.14	0.23	0.05	-0.08
Min	-16.67%	-22.70%	-18.10%	-1.36	-1.28	-1.05	-6.18	-7.52	-5.49
Max	43.61%	42.39%	14.73%	3.79	3.49	1.43	13.88	41.35	60.70
Global Hedge Funds									
<i>n</i>	2254	1656	1104	2254	1656	1104	2254	1656	1104
μ	11.66%	7.51%	2.29%	0.25	0.32	-0.09	0.40	-0.06	-0.07
σ	11.31%	10.42%	8.22%	1.05	0.94	0.86	8.77	6.95	0.58
Median	11.47%	7.67%	2.83%	0.25	0.29	0.08	0.16	0.07	-0.05
Min	-29.69%	-33.59%	-24.35%	-2.66	-2.05	-2.55	-80.26	-260.49	-11.60
Max	52.70%	38.64%	74.64%	3.39	3.34	2.15	400.97	28.48	5.55

From Table 8, the mean returns decline moving through phase 1 to phase 3. The mean of the scaled Sharpe and Treynor ratios for all regions shows no distinct or constant pattern when moving between phases. Highest and lowest scaled Sharpe ratio levels occur during phase 1, while the scaled Treynor ratio statistics show no evident pattern.

The standard deviation of both returns and scaled Sharpe ratios are also at their highest levels, for most regions, during phase 1. Mean Asian hedge fund returns did not increase into positive territory from phase 2 to phase 3 as the funds from other regional mandates did (see μ and median for returns in Table 8). This phenomenon is again echoed for the Asian market, as represented by the Nikkei 225 in Table 9. Comparing the mean returns for the hedge funds, per region, with their relevant market benchmark indicates that although these funds did not perform very well in absolute terms, they did outperform their respective markets in phase 3 - this was not the case during the crisis (phase 2). During phase 1, all the funds outperformed their respective market benchmarks in terms of return performance. The mean scaled Sharpe ratios of particularly phase 3 of Tables 8 and 9 highlight that, at times, it could have served investors better to hold riskless assets rather than investments in these funds, or even a basket of the market index.

Table 1B in Appendix 1B details comparative summary statistics to Table 8 with an altered focus of traditional Sharpe and Treynor ratios. Table 2A in Appendix 2A provides comparative summary statistics to Table 9, but detailing the traditional Sharpe and Treynor ratios for the relevant market benchmarks.

Table 9: Summary Statistics For Market Indices Per Phase

	Phase 1	Phase 2	Phase 3	Phase 1	Phase 2	Phase 3	Phase 1	Phase 2	Phase 3
	Return Statistics			Scaled Sharpe Ratio			Scaled Treynor Ratio		
<i>n</i>	49	36	24	49	36	24	49	36	24
US Market Index – S&P500									
μ	-1.54%	9.14%	-1.80%	0.72	1.14	1.05	0.01	-0.01	-0.04
σ	19.17%	13.29%	5.63%	0.74	0.53	0.43	0.10	0.10	0.07
Median	-6.02%	10.91%	-3.62%	0.97	1.05	1.13	0.02	0.02	-0.04
Min	-31.68%	-12.80%	-9.35%	-0.85	0.23	0.42	-0.19	-0.18	-0.13
Max	35.05%	26.51%	9.25%	1.73	2.12	1.65	0.16	0.11	0.12
European Market Index – DAX									
μ	0.86%	9.14%	-1.80%	0.81	1.66	0.63	-0.02	0.09	-0.02
σ	9.75%	13.29%	5.63%	0.83	1.01	0.40	0.20	0.14	0.06
Median	1.82%	10.91%	-3.62%	0.58	1.92	0.68	-0.07	0.11	-0.04
Min	-17.28%	-12.80%	-9.35%	-0.54	0.12	-0.07	-0.34	-0.14	-0.10
Max	15.14%	26.51%	9.25%	2.70	3.17	1.26	0.36	0.28	0.10
Asian Market Index – Nikkei 225									
μ	1.77%	-1.07%	-11.29%	0.48	0.21	0.33	0.02	-0.01	-0.13
σ	15.69%	14.41%	5.90%	1.04	0.57	0.40	0.17	0.15	0.06
Median	1.29%	3.18%	-12.40%	0.61	0.16	0.45	0.01	0.03	-0.13
Min	-26.81%	-22.41%	-19.73%	-1.61	-0.70	-0.29	-0.28	-0.24	-0.21
Max	29.24%	17.25%	1.57%	2.40	1.24	0.80	0.31	0.18	0.02
Global Market Index – MSCI World									
μ	6.42%	4.95%	-3.46%	1.56	2.49	1.27	0.06	0.05	-0.04
σ	12.43%	12.88%	6.37%	1.41	1.12	0.49	0.13	0.13	0.07
Median	8.08%	10.02%	-4.66%	1.52	2.85	1.35	0.08	0.10	-0.05
Min	-13.36%	-16.59%	-11.89%	-0.54	0.36	0.48	-0.14	-0.18	-0.13
Max	27.41%	20.99%	12.29%	4.15	4.03	1.93	0.30	0.22	0.12

5. SUMMARY AND CONCLUSION

Scaled versions of the traditional or classic Sharpe and Treynor ratios were used to augment the traditional measures as performance measure in a hedge fund context. ‘Live’, individual, long/short equity hedge funds, sourced from the EurekaHedge database and spanning geographical mandates that included North America, Europe, Asia and global, were used, spanning the period January 2000 to December 2011. These scaled risk-adjusted performance measures account for higher-order moments (skewness and kurtosis) of return distributions seeing that the traditional versions of these measures do not. The latter is a widely known shortcoming of the traditional measures and results in these measures being unfit for use with asymmetric returns, which is a characteristic of hedge funds. In methodological terms, this study used a 36-month rolling (window) period to estimate the relevant statistics and ratios.

As the use of most traditional performance measures, like the Sharpe ratio, are not appropriate considering non-normal returns, the normality of the funds were statistically investigated. To discover the state of normality of the hedge funds within the funds sample, a normality test or analysis was performed using the Jarque-Bera (JB) test, while a number of other relevant statistics were also used to gauge and present the state of normality. The majority (84.3%) of the funds, within the sample, showed evidence of having non-normal return distributions at the 5% significance level. These goodness-of-fit tests for normal return distributions, using the JB test, were performed and presented for the entire time-series and partitioned into phases corresponding the 2007 financial crisis; namely, pre, during and after. This partitioning practice was employed throughout the study and aids in viewing how a particular characteristic or phenomenon evolved during the different economic phases that surround the 2007 financial crisis. Results indicate that return distributions move further from normal(ity) during the crisis period (December 2006-December 2009) and that skewness turned rapidly and significantly negative during the financial crisis, while extreme levels of excessive positive kurtosis were evident for both funds and regional market indices.

Results were thereafter presented in two sections. Firstly, comparative traditional versus scaled Sharpe and Treynor values and rankings were assembled using the 36-month rolling method at different points of economic activity (*pre-, during, and post* crisis) to gauge how these measures value and rank funds over changing economic conditions. The top and bottom 25 funds in the data sample were identified at a point during the crisis period, according to the traditional Sharpe and Treynor ratios, respectively, and thereafter ranked both backwards and forwards in time within the full data sample.

Sharpe Ratio Findings - Noticeable shifts in fund performance during the 2007 financial crisis were observed. A clear distinction between strong (best) and weak (worst) performing funds is apparent while discrepancies, in terms of risk-adjusted values, between funds are smaller for the periods prior and after the crisis compared to during the crisis.

During the crisis period, a large contingent of the top performing funds was also found to cluster around values between 0.5 and 1, for both the traditional and scaled Sharpe ratios. Also, the traditional Sharpe ratio had the propensity to value funds somewhat higher compared to the scaled Share ratio during the crisis period. This lower risk-adjusted valuation by the scaled measure is shown to be realistic and validated as the scaled measure accounts for the increased risk due to the higher levels of skewness and kurtosis that characterised this particular period. Moreover, it is found that some of the top funds during the crisis performed poorly prior to the crisis, suggesting that investors would possibly not have selected these funds prior to the crisis due their mediocre or weak (Sharpe) risk-adjusted performance.

Treynor Ratio Findings - Similar to the Sharpe ratio, the Treynor ratio valuation and ranking analysis also signified a shift in fund performance during the crisis period, as opposed to prior, while a marked division is evident between the strong (best) and weak (worst) performing funds. The risk-adjusted value discrepancies between funds are wider during the crisis period compared to prior or after, for both measures. A number of the worst performing funds are valued and ranked lower (worse) by the scaled measure than the traditional measure during the crisis. Also, during the crisis period, a large group of the top performing funds are valued just above 0, while a large group of funds are valued between 0 and -2 by both the traditional and scaled measures.

During the crisis, the traditional and scaled Treynor measures did generally rank the funds similarly – meaning that both measures predominantly ranked the best performing funds higher (better) than the worst performing funds. In the after-crisis period, the best and worst performing funds remain significantly distinct from each other albeit that during this period the traditional and scaled Treynor measures do not rank the funds similarly as either high (best) or low (worst). An exceptionally random ranking of both the best and worst performing funds by both measures, with no particular ranking pattern or relationship seem to exist. The only observation that is obvious is that the scaled Treynor ratio almost wholly ranks the best performing funds high (best) and the worst performing funds as low (worst). Similar to the Sharpe ratio, it is again found that some of the top funds during the crisis performed disappointingly prior to the crisis, suggesting that investors would possibly not have selected these funds prior to the financial crisis due to their mediocre or weak (Treynor) risk-adjusted performance.

Secondly, to highlight the changing characteristics of hedge funds and their respective market benchmarks over the varying economic conditions around the 2007 financial crisis, a selective statistical analysis of returns and the scaled Sharpe and Treynor ratios were conducted. Results show a decrease in the average return of all funds over time. The same is true for the standard deviation of returns and the scaled Sharpe ratio, indicating a diminishing performance spectrum between funds, on average. Moving through the periods, the average Sharpe and Treynor ratios increased slightly during the crisis period and then decrease significantly more in the period thereafter.

Negative scaled Sharpe and Treynor ratios are also more prevailing during the period after the crisis. Comparing the mean returns for the hedge funds per region, with their relevant market benchmark, indicates that although these funds did not perform very well in absolute terms, they did outperform their respective markets after the crisis - this was not the case during the crisis. Prior to the crisis, all the funds outperformed their respective market benchmarks in terms of return performance. The mean scaled Sharpe ratios, of particularly after the crisis, emphasise that, at times, it could have served investors better to hold riskless assets rather than investments in these funds or even a basket of the market index.

The need to accurately distinguish between good and poor quality fund returns has not diminished, and, in actual fact, is ever increasing. More sophisticated risk-adjusted performance measures are therefore required to augment classical or traditional performance measures. Higher moments of the return distributions must also be accounted for if accurate and trustworthy fund comparisons (in terms of risk-adjusted returns) are desired, and even more so if the popularity of hedge fund investing continues to grow. The scaled Sharpe and Treynor ratios are arguable improvements on the traditional Sharpe and Treynor ratios as far as non-normal return distributions are concerned and can be considered as measures that should augment the use of the traditional Sharpe and Treynor ratios.

Future research considerations may be aimed at exploring and comparing the classical or traditional Sharpe and Treynor ratios to scaled versions, which account for higher-order moments, using similar scaling methodology albeit with the added element of allowing for the serial correlation of returns that are non-IID by using the annualised autocorrelation adjusted Sharpe ratio methodology as proposed by Lo (2002). A similar study that includes both the traditional and scaled performance measures, along with a third (objective) measure that accounts for higher moments, such as the Omega ratio, is another consideration.

ACKNOWLEDGEMENTS

This work is based on research financially supported by the National Research Foundation (NRF), along with the University of South Africa. Any opinion, finding, and conclusion or recommendation expressed in this material is that of the author(s); the NRF and the University of South Africa do not accept any liability in this regard.

The authors thank the Workwell Research Unit at the North-West University, Potchefstroom, South Africa, for their financial assistance in procuring the data.

AUTHOR INFORMATION

Dr. Francois van Dyk is a Senior Lecturer in the Department of Finance, Risk Management and Banking, at UNISA, Pretoria, South Africa. He began his career in South Africa as a risk analyst specialising in Basel II at FirstRand Bank Ltd. He furthered his career as a senior consultant at a niche international risk consultancy. He recently obtained his Ph.D, which this study forms part of, in risk management at the North-West University. His Ph.D focused on novel, quantitative risk measures within a hedge fund context. He has further expertise and interest in investment and portfolio risk and holds PRM and CHP designations while currently pursuing his CFA charter.. Email: vdykf@unisa.ac.za.

Dr. Gary van Vuuren began his career with a Masters in astrophysics and a Ph.D. in nuclear physics. He transferred to quantitative finance and, after a spell at Goldman Sachs in London, obtained a Masters in market risk and a Ph.D. in credit risk. He then worked as a risk manager for South African retail banks and asset managers before moving to London and working in retail and investment banks. He settled on quantitative risk assessment and management in financial institutions for Fitch Ratings where he remains employed. He is an accredited GARP Financial Risk Manager. Email: vvgary@hotmail.com.

Dr. André Heymans completed his Ph.D. in finance in 2007 then moved to London where he was employed by BNY MELLON until mid-2008. He then moved to South Africa to fill the position of Head of Research and Development in the trading room at an agricultural trading firm (Free State Maize). André moved back to academia in April 2009 where he currently holds the position Program Head of Risk Management.

REFERENCES

1. Ackermann, C., Mcenally, R. and Ravenscraft, D. 1999. The performance of hedge funds: Risk, returns and Incentives. *Journal of Finance*, 54(3):833-874.
2. Agarwal, V. and Naik, N.Y. 2002. Introduction to hedge funds. *Gestion Alternative*, Juillet 2002, AFG-ASFFI, France.
3. Agarwal, V. and Naik, N.Y. 2004. Risk and portfolio decisions involving hedge funds. *Review of*

- Financial Studies*, 17(1):63-98.
4. Agarwal, V., Daniel, N.D. and Naik, N.Y. 2011. Do hedge funds manage their reported returns? *The Review of Financial Studies*, 24(10):3281-3320.
 5. Almeida, C. and GARCIA, R. 2012. Robust assessment of hedge fund performance through nonparametric discounting. EDHEC-Risk Institute. [Online] <http://www.edhec-risk.com/features/RISKArticle.2012-09-20.2507> [Accessed: 1 Dec. 2013].
 6. Ambrosio, F.J. 2007. An evaluation of risk metrics. [Online] <https://personal.vanguard.com/pdf/flgerm.pdf> [Accessed: 17 Mar. 2013].
 7. Amenc, N., Goltz, F., Le Sound, V. and Martellini, L. 2008. EDHEC European investment practices survey 2008. EDHEC-Risk Institute, January.
 8. Anson, M.P., Chambers, D.R., Black, K.H. and Kazemi, H. 2012. *CAIA Level I: An introduction to core topics in alternative investments*. 2nd ed. Hoboken, NJ: Wiley.
 9. Artzner, P., Delbaen, F., Eber, J.M. and Heath, D. 1997. Thinking coherently. *Risk Magazine*, 10(11):68-71.
 10. Bailey, D.H and Lopez De Prado, M. 2013. Drawdown-based stop-outs and the ‘triple penance’ rule. Working paper. [Online] SSRN: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2201302 [Accessed: 10 Jan. 2014].
 11. Barclayhedge. 2014a. [Online] http://www.barclayhedge.com/research/indices/ghs/mum/Hedge_Fund.html [Accessed: 29 Jun. 2013].
 12. Barclayhedge. 2014b. Press release. 15 January. [Online] http://www.barclayhedge.com/research/press_releases/PR_Jan_15_2014.html [Accessed: 24 Feb. 2014].
 13. Bisias, D., Flood, M., Lo, A.W. and Valavanis, S. 2012. A survey of systemic risk analytics. *Annual Review of Financial Economics*, 4(1):255-296.
 14. Black, K.H. 2006. Improving the hedge fund risk exposures by hedging equity market volatility, or how the VIX strategy ate my kurtosis. *Journal of Trading*, 1(2):6-15.
 15. Bollen, N.P.B. and Pool, V.K. 2009. Do hedge fund managers misreport returns? Evidence from the pooled distribution. *Journal of Finance*, 64(5):2257-2288.
 16. Botha, M. 2007. A comparison of South African hedge fund risk measures. *South African Journal of Economics*, 75(3):459-477, September.
 17. Brooks, C. and Kat, H. 2002. The statistical properties of hedge fund index returns and their implications for investors. *Journal of Alternative Investments*, 5(2):26-44.
 18. Brown, S.J., Goetzmann, W.N. and Ibbotson, R.G. 1999. Offshore hedge funds: Survival and performance 1989-95. *Journal of Business*, 72(1):91-117.
 19. Burke, G. 1994. A sharper Sharpe ratio. *Futures*, 23(3):56.
 20. Capocci, D., Corhay, A. and Hubner, G. 2005. Hedge fund performance and persistence in bull and bear markets. *European Journal of Finance*, 11(5):361-392.
 21. Chan, N., Getmansky, M., Haas, S.M. and Lo, A.W. 2005. Systemic risk and hedge funds. *Optimization*, 920(617):235-338.
 22. Chen, Z. and Knez, P.J. 1996. Portfolio performance measurement: theory and applications. *Review of Financial Studies*, 9(2):511-556.
 23. Clarke, W. 2012. Hedge fund launches still below their pre-crisis peak. Investment Europe. 15 March. [Online] <http://www.investmenteurope.net/investment-europe/news/2159709/hedge-fund-launches-pre-crisis-peak> [Accessed: 3 Mar. 2013].
 24. CNBC. 2013. Wall street closes 2013 at records; best year in 16 for S&P, 18 for Dow. 31 December. [Online] <http://www.cnbc.com/id/101303244> [Accessed: 24 Feb. 2014].
 25. Connor, G. and Woo, M. 2003. An introduction to Hedge Funds. Financial Markets Group, London School of Economics, Discussion Paper. [Online] <http://eprints.lse.ac.uk/24675/1/dp477.pdf> [Accessed: 6 Apr. 2014].
 26. Deutsche Bank. 2014. Deutsche Bank releases 2014 alternative investment survey highlighting key trends shaping the hedge fund industry. Press release, 18 February. [Online] https://www.db.com/medien/en/content/4666_4819.htm [Accessed: 24 Feb. 2014].
 27. Ding, B. and Shawky, H.A. 2007. The performance of hedge fund strategies and the asymmetry of return distributions. *European Financial Management*, 13(2):309-331.
 28. Dowd, K. 2000. Adjusting for risk: An improved Sharpe ratio. *International Review of Economics and*

- Finance*, 9(3):209-222.
29. Eling, M. 2006. Autocorrelation, bias and fat tails: Are hedge funds really attractive investments? *Derivatives Use, Trading Regulation*, 12(1):28-47.
 30. Eling, M. and Schuhmacher, F. 2006. Does the choice of performance measure influence the evaluation of hedge funds? *Journal of Banking and Finance*, 31(9):2632-2647.
 31. Elton, E., Gruber, M., Das, S. and Hlavka, M. 1993. Efficiency with costly information: A reinterpretation of evidence from managed portfolios. *Review of Financial Studies*, 6(1):1-22.
 32. Eurekahedge. 2010. 2010 Key Trends in Global Hedge Funds. Hedge Fund Monthly. August. [Online] http://www.eurekahedge.com/news/10_aug_EH_Aug-10_GHF_Key_Trends_ABR.asp [Accessed: 23 Feb. 2013].
 33. Eurekahedge. 2012. 2012 Key trends in Global Hedge Funds. Hedge Fund Monthly. July. [Online] http://www.eurekahedge.com/news/Eurekahedge_Jul_2012_Global_Hedge_Funds_Key_Trends.asp [Accessed: 19 Feb. 2013].
 34. Eurekahedge. 2013. The Eurekahedge report. December. [Online] http://www.eurekahedge.com/news/13_Dec_Eurekahedge_Report_online.asp [Accessed: 6 Jan. 2014].
 35. Eurekahedge. 2014a. The Eurekahedge report. February. [Online] http://www.eurekahedge.com/news/14_Feb_Eurekahedge_Report_online.asp [Accessed: 24 Feb. 2014].
 36. Eurekahedge. 2014b. 2013 Overview: Key trends in global hedge funds. [Online] http://www.eurekahedge.com/news/Eurekahedge_Jan_2014_Global_Hedge_Funds_Key_Trends.asp [Accessed: 24 Feb. 2014].
 37. Feng, S. 2011. *Three essays on hedge fund fee structure, return smoothing and gross performance*. Amherst: University of Massachusetts. (Dissertation – D.Phil). [Online] http://scholarworks.umass.edu/open_access_dissertations/461/ [Accessed: 3 Mar. 2013].
 38. Financial Times. 2014. Bond investors braced for new year shock. 2 January. [Online] <http://www.ft.com/cms/s/0/76ec3d96-6812-11e3-a905-00144feabdc0.html#axzz2uI0Wy1Ew> [Accessed: 24 Feb. 2014].
 39. Fung, W. and Hsieh, D.A. 1997. Empirical characteristics of dynamic trading strategies: the case of hedge funds. *Review of Financial Studies*, 10(2):275-302.
 40. Fung, W. and Hsieh, D.A. 2000. Performance characteristics of hedge funds and commodity funds: Natural vs. spurious biases. *Journal of Financial and Quantitative Analysis*, 35(3):291-307.
 41. Fung, W. and Hsieh, D.A. 2001. The risk in hedge funds strategies: theory and evidence from trend followers. *Review of Financial Studies*, 14(2):313-341.
 42. Fung, W. and Hsieh, D.A. 2004. Hedge fund benchmarks: A risk based approach. *Financial Analysts Journal*, 60(5): 65-80.
 43. Gatfaoui, H. 2012. A correction for classic performance measures. *Chinese Business Review*, 1(1):1-28.
 44. Géhin, W. 2006. The Challenge of Hedge Fund Performance Measurement: a Toolbox rather than a Pandora's Box. EDHEC Risk and Asset Management Research Centre. [Online] http://www.edhec-risk.com/edhec_publications/RISKReview.2006-1117.4344/attachments/position%20paper%20the%20challenge%20of%20hedge%20fund070111.pdf [Accessed: 9 Jan. 2014].
 45. Getmansky, M., Lo, A. and Makarov, I. 2004. An econometric analysis of serial correlation and illiquidity in hedge fund returns. *Journal of Financial Economics*, 74(3):529-609.
 46. Gibbons, M., Ross, S. and Shanken, J. 1989. A test of the efficiency of a given portfolio. *Econometrica*, 57(5):1121-1152.
 47. Goetzmann, W., Ingersoll, J. and Spiegel, M. 2007. Portfolio performance manipulation and manipulation-proof performance measures. *Review of Financial Studies*, 20(5):1503-1546.
 48. Goetzmann, W., Ingersoll, J., Spiegel, M. and Welch, I. 2002. Sharpening Sharpe ratios. Working paper, NBER.
 49. Gregoriou, G.N. and Geuyie, J.P. 2003. Risk-adjusted performance of funds of hedge funds using a modified Sharpe ratio. *Journal of Wealth Management*, 6(3):77-83.
 50. Hodges, S. 1998. A generalization of the Sharpe ratio and its applications to valuation bounds and risk measures. Working paper, Financial Options Research Centre. [Online] <http://www2.warwick.ac.uk/fac/soc/wbs/subjects/finance/research/wpaperseries/1998/98-88.pdf> [Accessed: 29 Jun. 2013].
 51. Jaeger, R.A. 2003. *All about hedge funds: The easy way to get started*. New York: McGraw-Hill.
 52. Jagric, T., Podobnik, B., Strasek, S. and Jagric, V. 2007. Risk-adjusted performance of mutual funds:

- Some tests. *South-Eastern Europe Journal of Economics*, 2(1):233-244.
53. Jensen, M. 1968. The performance of mutual funds in the period 1945-1964. *Journal of Finance*, 23(2):389-146.
54. Jobson, J.D. and Korkie, B.M. 1981. Performance hypothesis testing with the Sharpe and Treynor measures. *Journal of Finance*, 36(4):889-908, September.
55. Kaiser, D. and Haberfelner, F. 2012. Hedge fund biases after the financial crisis. *Managerial Finance*, 38(1):27-43.
56. Kane, A. 1982. Skewness preference and portfolio choice. *Journal of Financial and Quantitative Analysis*, 17(1):15-25.
57. Kaplan, P.D. and Knowles, J.A. 2004. Kappa: A generalized downside risk-adjusted performance measures. *Journal of Performance Measurement*, 8(3):42-54.
58. Kat, H.M. and Amin, G.S. 2001. Welcome to the dark side: Hedge fund attrition and survivorship bias over the period 1994-2001. Working paper. [Online] SSRN: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=293828 [Accessed: 1 Mar. 2013].
59. Kat, H. and Lu, S. 2002. An excursion into the statistical properties of hedge fund returns. Working paper. [Online] http://www.eurekahedge.com/news/attachments/stats_hfr.pdf [Accessed: 5 Apr. 2014].
60. Koekebakker, S. and Zakamouline, V. 2008. Generalized Sharpe ratios and portfolio performance evaluation. Working paper. [Online] SSRN: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1028715 [Accessed: 12 Dec. 2012].
61. Kooli, M., Morin, F. and Sedzro, K. 2005. Evaluation des Mesures de Performance des Hedge Funds. Paper presented at the Annual Conference of the French Association of Finance, June.
62. KPMG. 2012. The value of the hedge fund industry to investors, markets, and the broader economy. [Online] <https://www.kpmg.com/KY/en/Documents/the-value-of-the-hedge-fund-industry-part1.pdf> [Accessed: 23 Jul. 2013].
63. Kraus, A. and Litzenberger, R.H. 1976. Skewness preference and the valuation of risk assets. *Journal of Finance*, 31(4):1085-1100.
64. Leland, H.E. 1999. Beyond mean-variance: Risk and performance measurement in a nonsymmetrical world. *Financial Analysts Journal*, 55(1):27-36, Jan/Feb.
65. Lhabitant, F. 2002. *Hedge Funds: Myths and Limits*. London: Wiley.
66. Lhabitant, F. 2004. *Hedge funds: Quantitative insight*. Hoboken, NJ: Wiley.
67. Liang, B. 1999. On the performance of hedge funds. *Financial Analysts Journal*, 55(4):72-85.
68. Liang, B. 2000. Hedge funds: The living and the dead. *Journal of Financial and Quantitative Analysis*, 35(3):333-370.
69. Li, Y. and Kazemi, H. 2007. Conditional properties of hedge funds: Evidence from daily returns. *European Financial Management*, 13(2):211-238.
70. Lo, A.W. 2001. Risk Management for Hedge Funds: introduction and overview. *Financial Analysts Journal*, 57(6):16-33.
71. Lo, A.W. 2002. The statistics of Sharpe Ratios. *Financial Analysts Journal*, 58(4):36-52.
72. Lopez De Prado, M. 2013. How long does it take to recover from a drawdown? Working paper. [Online] SSRN: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2254668 [Accessed: 11 Oct. 2013].
73. Malkiel, B.G. and Saha, A. 2005. Hedge funds: Risk and return. *Financial Analysts Journal*, 61(6):80-88.
74. Maller, R.A. and Turkington, D.A. 2002. Mathematical methods of operations research. *Mathematical Methods of Operations Research*, 56(3):501-511.
75. Memmel, C. 2003. Performance hypothesis testing with the Sharpe Ratio. *Finance Letters*, 1(1):21-23.
76. Nguyen-Thi-Thanh, H. 2010. On the consistency of performance measures for hedge funds. *Journal of Performance Measurement*, 14(2):1-16.
77. Nguyen-Thi-Thanh, H. 2007. Assessing hedge fund performance: Does the choice of measures matter? HAL – CCSD. [Online] <http://halshs.archives-ouvertes.fr/halshs-00184814/en/> [Accessed: 16 Jan. 2014].
78. Opdyke, J.D. 2007. Comparing Sharpe Ratios: So where are the p-values? *Journal of Asset Management*, 8(5):308-336.
79. Pezier, J. and WHITE, A. 2006. The relative merits of investable hedge fund indices and of funds of hedge funds in optimal passive portfolios. ICMA Centre Discussion papers in Finance. [Online] <http://www.icmacentre.ac.uk/pdf/discussion/DP2006-10.pdf> [Accessed: 5 Feb. 2014].
80. Pflug, G. 2000. *Some remarks on the Value-at-Risk and the Conditional Value-at-Risk*. *Probabilistic*

81. *Constrained Optimization: Methodology and Applications*. Dordrecht: Kluwer Academic Publishers.
81. Prokop, J. 2012. Further evidence on the role of ratio choice in hedge fund performance evaluation. *Journal of Finance and Investment Analysis*, 1(3):181-195.
82. Rockafellar, R.T. and Uryasev, S. 2000. Optimization of Conditional Value-at-Risk. *Journal of Risk*, 2(3):21-41.
83. Roxburgh, C., Lund, S., Lippert, M., White, O.L. and Zhao, Y. 2009. The new power brokers: How oil, Asia, hedge funds, and private equity are faring in the financial crisis. McKinsey & Company Report. July. [Online] http://www.mckinsey.com/insights/economic_studies/how_the_new_power_brokers_are_faring_in_financial_crisis [Accessed: 23 Feb. 2013].
84. Schmid, F. and SCHMIDT, R. 2007. Statistical Inference for Sharpe's Ratio. Working paper. [Online] http://wifo4.bwl.unimannheim.de/fileadmin/files/hws10/Schmid_Schmidt_Statistical_Inference.pdf [Accessed: 27 Feb. 2013].
85. Scott, R.C. and HORVATH, P.A. 1980. On the direction of preferences for moments of higher order than the variance. *Journal of Finance*, 35(4):915-919, September.
86. SEI. 2007. Five critical challenges for hedge funds. Annual Global Survey. [Online] <http://www.seic.com/docs/IMS/SEI-FiveCriticalChallengesUS.pdf> [Accessed: 15 Jul. 2013].
87. SEI. 2009. Hedge funds under the microscope: Examining institutional commitment in challenging times. [Online] <http://www.seic.com/docs/IMS/SEI-HFs-Under-the-MicroscopeUK.pdf> [Accessed: 3 Dec. 2013].
88. SEI. 2010. The era of the investor: New rules of institutional hedge fund investing. [Online] <http://www.seic.com/docs/IMS/SEI-EraOfTheInvestor-UK.pdf> [Accessed: 2 Dec. 2013].
89. SEI. 2011. Institutional hedge fund investing comes of age: A new perspective on the road ahead. [Online] http://www.seic.com/docs/IMS/SEI_2011HedgeFundStudy_UK.pdf [Accessed: 5 Dec. 2013].
90. SEI. 2012. The shifting hedge fund landscape: Institutions put fund managers to the test. [Online] http://www.seic.com/docs/IMS/IMS_0112_SEI_ShiftingLandscapePt1_US.pdf [Accessed: 8 Aug. 2013].
91. SEI. 2013. 6 Ways hedge funds need to adapt now. 6th Annual Global Survey. [Online] http://www.seic.com/docs/IMS/SEI-HF-Paper-6-Ways-to-Adapt_US.pdf [Accessed: 13 Jul. 2013].
92. Shadwick, W.F. and Keating, C. 2002. A universal performance measure. *Journal of Performance Measurement*, 6(3):59-87.
93. Sharpe, W.F. 1966. Mutual fund performance. *The Journal of Business*, 39(1):119-138, January.
94. Sharpe, W.F. 1975. Adjusting for risk in portfolio performance measurement. *Journal of Portfolio management*, 1(2):29-34, Winter.
95. Sharpe, W.F. 1992. Asset allocation: Management style and performance measurement. *Journal of Portfolio Management*, 18(2):7-19, Winter.
96. Sharpe, W.F. 1994. The Sharpe Ratio. *Journal of Portfolio Management*, 21(1):49-58, Fall.
97. Sortino, F.A. and Price, L.N. 1994. Performance measurement in a downside risk framework. *Journal of Investing*, 3(3):59-65.
98. Sortino, F. and Van Der Meer, R. 1991. Downside risk. *Journal of Portfolio Management*, 17(4):27-31, Summer.
99. Spurgin, R.B. 2001. How to game your Sharpe ratio. *Journal of Alternative Investments*, 4(3):38-46.
100. Stutzer, M. 2000. A portfolio performance index. *Financial Analysts Journal*, 56(3):52-61, May/June.
101. Taleb, N. 2007. *The black swan: The impact of the highly improbable*. New York: Random House.
102. TheCityUK. 2012. Global hedge funds see \$70bn net inflow of funds in 2011 as launches outpace liquidations for second year running. Hedge Fund Report 2012. [Online] <http://www.thecityuk.com/media/press-releases/global-hedge-funds-see-70bn-net-inflow-of-funds-in-2011-as-launches-outpace-liquidations-for-second-year-running/> [Accessed: 2 Mar. 2013].
103. TheCityUK. 2013. Hedge funds 2013. Research report, 17 May. [Online] <http://www.thecityuk.com/research/our-work/reports-list/hedge-funds-2012/> [Accessed: 23 Jul. 2013].
104. Till, H. 2007. The Amaranth collapse: what happened and what have we learned thus far? EDHEC Risk and Asset Management Research Centre. [Online] http://facultyresearch.edhec.com/servlet/com.univ.collaboratif.util.LectureFichiergw?ID_FICHIER=1328885972396 [Accessed: 15 Jul. 2013].
105. Treynor, J. 1965. How to rate management of investment funds. *Harvard Business Review*, 43(1):63-75, January-February.
106. Van Vuuren, G., Botha, M. and Styger, P. 2003. Tricky business: Measuring risk and return. GARP Risk Review, 15, Nov/Dec. [Online] http://w.eraider.com/images/articles/Issue15_1.pdf [Accessed: 10 Dec.

- 2012].
107. Wermers, R. 2010. Mutual fund performance: An empirical decomposition into stock-picking talent, style, transaction costs, and expenses. *Journal of Finance*, 55(4):1655-1703.
108. Wermers, R. 2011. Performance measurement of mutual funds, hedge funds, and institutional accounts. *Annual Review of Financial Economics*, 3(1):537-574.
109. Wiesinger, A. 2010. *Risk-adjusted performance measurement: State of the art*. St. Gallen: University of St. Gallen. (Bachelor Thesis – BBA). [Online] <http://www.iorcf.unisg.ch/Forschung/~media/0D77AF25603347E795C77DDA1DD83676.ashx> [Accessed: 6 Jun. 2013].
110. Young, T. 1991. Calmar ratio: A smoother tool. *Futures*, 20(11):40-41.
111. Zakamouline, V. 2011. The performance measure you choose influences the evaluation of hedge funds. *Journal of Performance Measurement*, 15(3):48-64.
112. Ziemba, W. 2005. The symmetric downside-risk Sharpe ratio. *Journal of Portfolio Management*, 32(1):108-122.

APPENDIX 1A

Table 1A: Traditional Sharpe And Treynor Ratio Summary Statistics For All Hedge Funds, Per Phase

	Phase 1	Phase 2	Phase 3	Phase 1	Phase 2	Phase 3
	Traditional Sharpe Ratio			Traditional Treynor Ratio		
<i>n</i>	9016	6624	4416	9016	6624	4416
μ	0.60	0.37	-0.07	0.07	0.01	-0.31
σ	0.91	0.83	0.50	31.23	1.82	14.54
Median	0.56	0.33	-0.11	0.14	0.04	-0.04
Min	-2.13	-1.74	-1.54	-2376.49	-130.04	-888.29
Max	3.51	3.31	2.86	923.90	25.00	170.83

APPENDIX 1B

Table 1B: Traditional Sharpe And Treynor Ratio Summary Statistics For Hedge Funds, Grouped Geographically, Per Phase

	Phase 1	Phase 2	Phase 3	Phase 1	Phase 2	Phase 3
	Traditional Sharpe Ratio			Traditional Treynor Ratio		
<i>n</i>	49	36	24	49	36	24
North American Hedge Funds						
μ	0.56	0.37	0.004	-0.65	0.03	-0.33
σ	0.86	0.82	0.52	40.48	0.88	20.42
Median	0.49	0.33	-0.07	0.11	0.04	-0.02
Min	-1.45	-1.74	-1.54	-2376.49	-10.98	-888.29
Max	3.23	3.31	2.86	49.05	22.89	170.83
European Hedge Funds						
μ	0.55	0.46	-0.12	1.58	-0.11	-0.65
σ	1.03	0.90	0.51	32.45	3.73	9.54
Median	0.53	0.44	-0.12	0.22	0.03	-0.06
Min	-2.13	-1.63	-1.35	-174.97	-130.04	-275.18
Max	3.51	2.69	1.51	923.90	25.01	15.88
Asian Hedge Funds						
μ	0.79	0.25	-0.13	0.22	0.004	-0.03
σ	0.91	0.90	0.46	0.27	0.62	-0.23
Median	0.78	0.23	-0.11	0.23	0.04	-0.05
Min	-1.41	-1.40	-1.26	-0.57	-10.60	-0.56
Max	3.20	2.41	0.90	0.99	1.41	0.86
Global Hedge Funds						
μ	0.62	0.35	-0.16	0.11	0.07	-0.10
σ	0.88	0.77	0.44	1.56	0.46	0.32
Median	0.60	0.30	-0.19	0.15	0.07	-0.07
Min	-1.70	-1.45	-1.30	-60.62	-2.66	-3.10
Max	3.43	2.35	1.99	17.28	10.51	0.85

APPENDIX 2A

Table 2A: Traditional Sharpe And Treynor Ratio Summary Statistics For Market Indices Per Phase

	Phase 1	Phase 2	Phase 3	Phase 1	Phase 2	Phase 3
	Traditional Sharpe Ratio			Traditional Treynor Ratio		
<i>n</i>	49	36	24	49	36	24
	US Market Index – S&P500					
μ	0.27	0.23	-0.16	-0.02	-0.01	-0.04
σ	0.82	0.82	0.34	0.20	0.10	0.07
Median	0.10	0.17	-0.18	-0.07	0.02	-0.04
Min	-1.01	-1.03	-0.58	-0.34	-0.18	-0.13
Max	1.69	1.46	0.60	0.36	0.11	0.12
	European Market Index – DAX					
μ	0.18	0.83	-0.08	-0.02	0.09	-0.02
σ	0.97	1.08	0.24	0.20	0.14	0.06
Median	-0.22	0.67	-0.16	-0.07	0.11	-0.04
Min	-1.15	-0.65	-0.40	-0.34	-0.14	-0.10
Max	1.97	2.42	0.38	0.36	0.28	0.10
	Asian Market Index – Nikkei 225					
μ	0.20	0.10	-0.47	0.02	-0.01	-0.13
σ	0.93	0.81	0.23	0.17	0.15	0.06
Median	0.07	0.18	-0.48	0.01	0.03	-0.13
Min	-1.33	-1.03	-0.76	-0.28	-0.24	-0.21
Max	1.99	1.22	0.07	0.31	0.18	0.02
	Global Market Index - MSCI World					
μ	0.65	0.68	-0.14	0.06	0.05	-0.04
σ	1.16	1.05	0.30	0.13	0.13	0.07
Median	0.55	0.81	-0.20	0.08	0.10	-0.05
Min	-0.97	-0.79	-0.50	-0.14	-0.18	-0.13
Max	2.98	2.30	0.62	0.30	0.22	0.12

APPENDIX 3A

Table 3A: Summary Statistics For Regionally Grouped Hedge Fund Indices Per Phase

	Phase 1	Phase 2	Phase 3	Phase 1	Phase 2	Phase 3	Phase 1	Phase 2	Phase 3
	Return Statistics			Scaled Sharpe Ratio			Scaled Treynor Ratio		
<i>n</i>	49	36	24	49	36	24	49	36	24
North American Hedge Fund Index*									
μ	0.10	0.08	0.06	1.80	2.11	1.81	-0.001	-0.02	-0.09
σ	0.02	0.04	0.02	1.00	0.76	0.34	0.34	0.22	0.19
Median	0.10	0.09	0.06	1.67	2.13	1.84	0.06	0.04	-0.10
Min	0.05	-0.002	0.03	-0.19	0.76	1.24	-0.66	-0.44	-0.32
Max	0.16	0.14	0.11	3.55	3.23	2.39	0.63	0.26	0.34
European Hedge Fund Index[#]									
μ	0.08	0.08	0.04	1.95	2.92	0.66	-1.15	0.28	-0.09
σ	0.03	0.04	0.01	1.29	1.33	0.25	3.39	0.48	0.31
Median	0.07	0.09	0.04	1.48	3.30	0.63	-1.24	0.44	-0.20
Min	0.02	0.01	0.01	-0.36	0.75	0.19	-14.56	-0.56	-0.50
Max	0.15	0.14	0.06	4.76	4.96	1.12	6.79	0.95	0.60
Asian Hedge Fund Index⁺									
μ	0.12	0.10	0.05	2.01	2.39	0.72	-0.04	-0.03	-0.42
σ	0.04	0.06	0.03	0.61	0.89	0.28	0.70	0.54	0.20
Median	0.13	0.11	0.05	1.99	2.32	0.73	0.07	0.11	-0.43
Min	0.03	0.01	0.01	0.68	0.96	0.27	-1.52	-0.89	-0.67
Max	0.20	0.19	0.10	3.50	4.09	1.22	1.05	0.69	0.06
Global Hedge Fund Index[‡]									
μ	0.07	0.02	-0.03	1.17	1.51	1.27	0.10	0.23	-0.10
σ	0.02	0.05	0.02	0.40	0.62	0.33	0.85	0.52	0.29
Median	0.06	0.04	-0.03	1.21	1.62	1.34	0.33	0.32	-0.17
Min	0.04	-0.05	-0.06	0.25	0.39	0.69	-1.85	-0.54	-0.42
Max	0.11	0.07	0.03	1.73	2.61	1.75	1.07	1.10	0.66

* North American hedge fund index: Eureka hedge North America long/short equities hedge fund index.

[#] European hedge fund index: Barclay hedge European equities index.

⁺ Asian hedge fund index: Eureka hedge Asian hedge fund index.

[‡] Global hedge fund index: HFR(X) global hedge fund index.

APPENDIX 3B

Table 3B: Traditional Sharpe And Treynor Ratio Summary Statistics For Regionally Grouped Hedge Fund Indices Per Phase

	Phase 1	Phase 2	Phase 3	Phase 1	Phase 2	Phase 3
	Traditional Sharpe Ratio			Traditional Treynor Ratio		
<i>n</i>	49	36	24	49	36	24
	North America Hedge Fund Index*					
μ	1.59	1.27	0.62	-0.02	-0.04	-0.08
σ	0.54	0.88	0.27	0.25	0.20	0.16
Median	1.51	1.54	0.58	0.04	0.04	-0.10
Min	0.66	-0.07	0.26	-0.51	-0.45	-0.29
Max	2.67	2.63	1.20	0.37	0.21	0.26
	European Hedge Fund Index[‡]					
μ	1.80	1.40	0.46	-0.73	0.24	-0.09
σ	0.59	0.91	0.21	1.84	0.42	0.26
Median	1.69	1.45	0.49	-0.63	0.38	-0.17
Min	0.55	0.07	0.07	-4.84	-0.53	-0.44
Max	3.18	2.81	0.82	2.46	0.76	0.42
	Asian Hedge Fund Index⁺					
μ	2.14	1.49	0.48	-0.02	0.01	-0.37
σ	0.72	1.13	0.29	0.65	0.45	0.17
Median	2.30	1.36	0.45	0.05	0.08	-0.40
Min	0.52	0.04	0.09	-1.34	-0.63	-0.58
Max	3.51	3.41	1.16	0.98	0.62	0.05
	Global Hedge Fund Index[#]					
μ	1.79	0.45	-0.33	0.10	0.10	-0.11
σ	0.52	0.85	0.31	0.64	0.30	0.23
Median	1.72	0.75	-0.41	0.37	0.20	-0.16
Min	1.00	-0.68	-0.69	-1.47	-0.41	-0.37
Max	2.87	1.66	0.50	0.80	0.47	0.47

* North American hedge fund index: Eureka hedge North America long/short equities hedge fund index.

‡ European hedge fund index: Barclay hedge European equities index.

+ Asian hedge fund index: Eureka hedge Asian hedge fund index.

Global hedge fund index: HFR(X) global hedge fund index.

APPENDIX 4

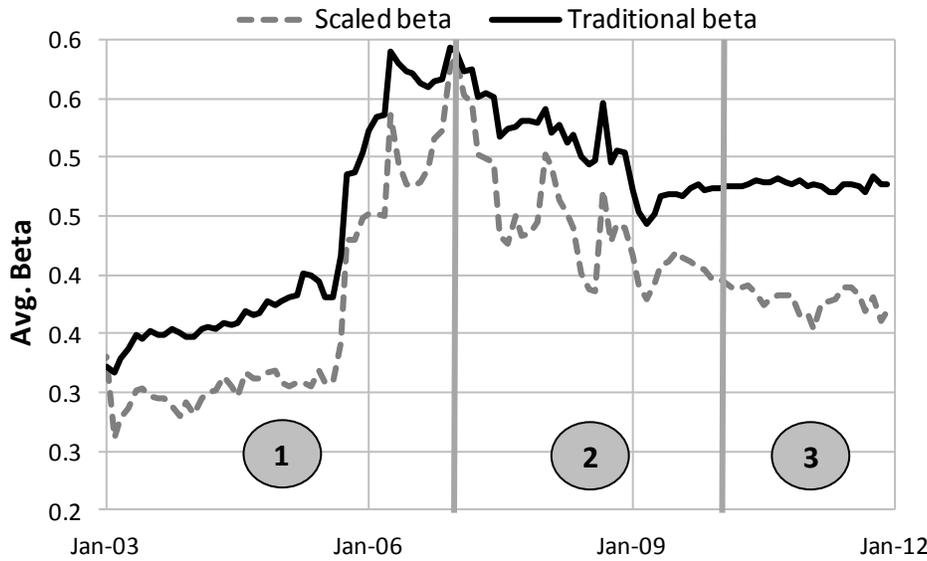


Figure 4A: Average Traditional vs. Scaled Beta As In Treynor Ratios, Over Time - All Hedge Funds

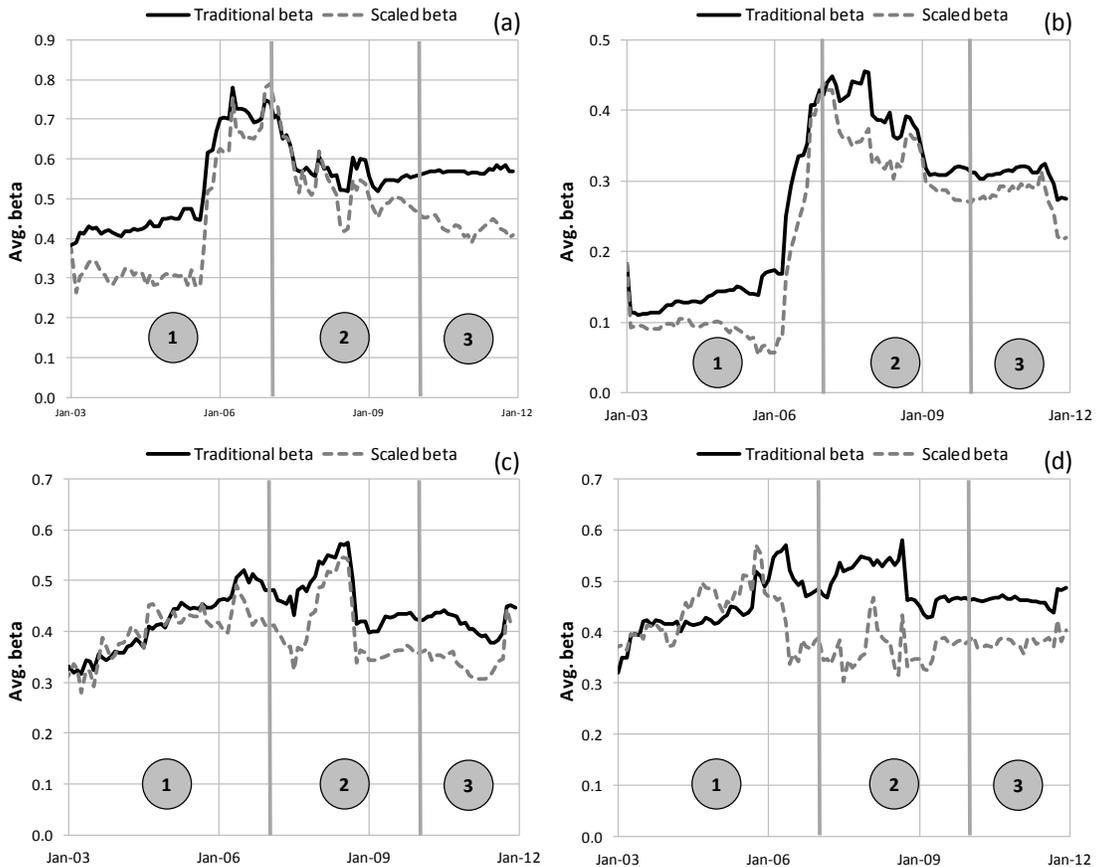


Figure 4B: Average Traditional vs. Scaled Beta Used As In Treynor Ratios – For Hedge Funds Per Geographical Mandate, Over Time: (a) North America, (b) Europe, (c) Asia, And (d) Global