# ALTERNATIVE INVESTMENT RESEARCH CENTRE WORKING PAPER SERIES

**Working Paper # 0030**

# SUPERSTARS OR AVERAGE JOES?

## A Replication-Based Performance Evaluation Of 1917 Individual Hedge Funds

**Harry M. Kat**

**Helder P. Palaro**

**Alternative Investment Research Centre**
**Cass Business School, City University**
**106 Bunhill Row, London, EC2Y 8TZ**
**United Kingdom**
**Tel. +44.(0)20.70408677**
**E-mail: harry@airc.info**
**Website: www.cass.city.ac.uk/airc**

# SUPERSTARS OR AVERAGE JOES?

## A Replication-Based Performance Evaluation Of 1917 Individual Hedge Funds

**Harry M. Kat***

**Helder P. Palaro****

This version: February 6, 2006

Please address all correspondence to:

Harry M. Kat
Professor of Risk Management and
Director Alternative Investment Research Centre
Cass Business School, City University
106 Bunhill Row, London, EC2Y 8TZ
United Kingdom
Tel. +44.(0)20.70408677
E-mail: harry@harrykat.com

*Professor of Risk Management and Director Alternative Investment Research Centre, Cass Business School, City University, London. **PhD Student, Cass Business School, City University, London. The authors like to thank CSI and Rudi Cabral for allowing us to access the CSI futures database.

# SUPERSTARS OR AVERAGE JOES?

## A Replication-Based Performance Evaluation Of 1917 Individual Hedge Funds


## Abstract

In this paper we use the hedge fund return replication technique recently introduced by Kat and Palaro (2005) to evaluate the net-of-fee performance of 1917 individual hedge funds. Comparing fund returns with the returns on dynamic futures trading strategies with the same risk and dependence characteristics, we find that no more than 17.7% of the hedge funds in our sample beat the benchmark. In other words, the majority of hedge funds have not provided their investors with returns, which they could not have generated themselves by mechanically trading S&P 500, T-bond and Eurodollar futures. Over time, we observe a substantial deterioration in overall hedge fund performance. In addition, we find a tendency for the performance of successful funds to deteriorate over time. This supports the hypothesis that increased assets under management endanger future performance.



**Keywords**: Hedge funds, performance evaluation, return replication.

**JEL Classification**: G10, G13, G23.

## 1. Introduction

With the first hedge fund dating back to 1949, hedge funds have been around for quite some time now. Academic research into hedge funds, however, only took off towards the end of the 1990s, when sufficient data became available. Since then, and inspired by the strong growth of the hedge fund industry worldwide, a large number of research papers and articles have provided insight in many different aspects of hedge funds[1]. One question largely remains unanswered though. Do hedge funds provide their investors with superior returns? In other words, do hedge funds provide their investors with returns, which they could not have obtained otherwise?

According to the hedge fund industry itself, the answer to the above question is of course affirmative, although with the somewhat disappointing recent performance of hedge funds, this point is put forward less often and less forcefully than it used to. Nowadays, most emphasis is on the diversification properties of hedge funds. Various academic studies have attempted to shed light on the issue of hedge fund return superiority. Most of these apply traditional performance measures, such as the Sharpe ratio or factor model based alphas, to hedge fund returns obtained from one or more of the main hedge fund databases. The conclusion is typically that hedge fund returns are indeed superior. From other studies, however, it is now well understood that hedge fund return data may suffer from various biases, which, when not corrected for, will produce artificially high Sharpe ratios and alphas. In addition, hedge fund returns are typically not normally distributed (see also section 6) and may derive from exposure to sometimes very unusual risk factors. This makes traditional performance measures unsuitable for hedge funds, as deviations from normality as well as every risk factor that is incorrectly specified or left out altogether, will show up as alpha, thereby suggesting superior performance where there actually may be none.

In theory, once the relevant risk factors have been identified, factor model based performance evaluation of hedge fund returns should work well. In practice, however, we don't know enough about hedge fund return generation to be sufficiently certain that all the relevant risk factors are included and correctly specified. As a result, factor

[1] A recent SSRN (www.ssrn.com) search for papers including "hedge fund" in the title and/or abstract yielded a total of 330 papers.

models typically explain only 25-30% of the variation in individual hedge fund returns, which compares very unfavourably with the 90-95% that is typical for mutual funds. Although the procedure works better for portfolios of hedge funds, funds of funds and hedge fund indices, where most of the idiosyncratic risk is diversified away, the low determination coefficients of these models make it impossible to arrive at a firm conclusion with respect to the superiority of hedge fund returns.

In a way, it is quite surprising that so many people, on the buy-side as well as in academia, are so eager to believe that the, sometimes huge, alphas reported for hedge funds are truly there. Anyone who is well calibrated to the world we live in and the global capital markets in particular, knows how difficult it is to consistently beat the market, i.e. systematically obtain a better return than what would be fair given the risks taken. Over time, hundreds, if not thousands, of studies have confirmed this. Is it therefore likely that suddenly we are facing a whole new breed of super-managers; not one or two, but literally thousands of them? Of course not! And if anything, the rise of the hedge fund industry has made the global capital markets even more efficient, not less.

Although by far the most popular, factor models are not the only way to evaluate hedge fund performance. Kat and Palaro (2005), or KP for short, recently introduced a technique that allows the derivation of dynamic futures trading strategies, trading stock index, bond, and interest rate futures, which generate returns with predefined statistical properties. The technique is not only capable of replicating (the statistical properties of) fund of funds returns, but works equally well for individual hedge fund returns. Since the KP replicating strategies are explicitly constructed to replicate the complete risk and dependence profile of a fund, the average return on these strategies can be used as a performance measure. When the average fund return is significantly higher than the average return on the replication strategy, the fund is the most efficient alternative and vice versa.

The KP replication technique is similar to that used previously in Amin and Kat (2003b) to evaluate hedge fund performance. The important difference, however, is that the latter only replicated the marginal distribution of the fund return, while KP also replicate its dependence structure with an investor's existing portfolio. This is a significant step forwards as most investors nowadays are attracted to hedge funds because of their relatively weak relationship with traditional asset classes, i.e. their diversification potential. Only replicating the marginal distribution without giving any consideration to the dependence structure between the fund and the investor's existing portfolio would therefore be insufficient.

From a more theoretical perspective, replication of a fund's dependence pattern with other asset classes is also a necessity. According to theory as well as casual empirical observation, expected return and systematic co-variance, co-skewness and co-kurtosis are directly related. In other words, it is not so much the marginal distribution, but its dependence structure with other assets that determines an asset's expected return. An asset, which is highly correlated with stocks and bonds, offers investors very little in terms of diversification potential. As a consequence, there will be little demand for this asset. Its price will be low and its expected return therefore relatively high. On the other hand, an asset that offers substantial diversification potential will be in high demand. Its price will be high and its expected return relatively low. Although hedge funds are not priced by market forces in the same way as primitive securities are, they do operate in the latter markets. It therefore seems plausible that a similar phenomenon is present in hedge fund returns as well[2].

## 2. The KP Efficiency Measure

Applying the KP replication technique to hedge funds, the goal is to create a dynamic trading strategy, which generates returns with the same statistical properties as a given hedge fund, i.e. returns that are drawings from the same distribution as the distribution from which the actual fund returns are drawn. The basic idea behind the replication procedure is straightforward. From the theory of dynamic trading it is well known that in the standard theoretical model with complete markets any payoff function can be hedged perfectly. This observation forms the foundation of arbitrage-based option pricing theory. If it is possible to find a payoff function which, given the distribution

[2] This is confirmed by the results in Kat and Miffre (2005).

of the underlying assets, implies the same distribution as the one from which the fund returns are drawn, then the accompanying dynamic trading strategy will generate (returns that are drawings from) that distribution.

Given the KP replication technique and following the same reasoning as in Amin and Kat (2003b), we derived the following evaluation procedure, consisting of five distinct steps.

1. Monthly return data are collected on the fund to be evaluated, the representative investor's portfolio, and a so-called reserve asset. The latter is the main source of uncertainty in the replication strategy. As we want to know whether the returns that investors obtain from hedge funds are superior, fund returns should be net of all fees.

2. From the available return data, the bivariate distribution of the fund return and the representative investor's portfolio return is inferred (KP refer to this as the 'desired distribution'). The same is done for the bivariate distribution of the investor's portfolio return and the return on the reserve asset (the 'building block distribution'). In line with KP, we allow for 54 different joint distributions, choosing between them using the Akaike Information Criterion (AIC)[3].

3. Assuming an initial investment in the fund of 100, we determine the cheapest payoff function, which is able to turn the building block distribution into the desired distribution. This payoff function is known as the 'desired payoff function' and lies at the basis of the KP replication strategies.

4. The desired payoff function is priced using the multivariate option pricing model of Boyle and Lin (1997), which explicitly allows for transaction costs. For the pricing of the payoff function, we estimate the required volatility and correlation inputs over the period covered by the track record of the fund being evaluated. We use the average 1-month interest rate over the same period for

[3] See Akaike (1973) for details.

the interest rate input. We will interchangeably refer to the price thus obtained as ‘the KP efficiency measure’, ‘the efficiency measure’ or ‘the KP measure’.

5. Finally, we compare the KP efficiency measure with the 100 initially invested in the fund. If the efficiency measure is 100 as well, then the replication strategy and the fund are equivalent. If the efficiency measure is less (more) than 100, the strategy is cheaper (more expensive) than the fund and the fund therefore inefficient (efficient).

The difference between the KP measure and 100 reflects the difference in mean return between the fund and the replication strategy. Suppose the fund had a mean return of 2%. In that case investing 100 could be expected to turn into 102 at the end of the month. Now suppose the investor’s portfolio and the reserve asset both had a mean return of 1%. In that case, to generate 102, we would have to invest 101 in the replication strategy. Likewise, when the fund had a mean of –2% it would turn 100 into 98. If the investor’s portfolio and the reserve asset both had a mean return of +1%, generating a payoff of 98 would require an investment in the replication strategy of 97. In both cases the difference between 100 and the required replicating investment reflects the difference in mean return between the fund and its replication strategy.

All performance evaluation studies in finance follow the same general procedure. First, using a fund’s track record and possibly some additional data over the same period as well, the fund return is characterized in some way. With the Sharpe ratio this is done by calculating the volatility of the fund return. With alphas this is done by estimating a fund’s exposure to the relevant risk factors. Second, based on this characterization, a benchmark return is determined and compared with the actual average fund return over its track record. With the Sharpe ratio the benchmark return is derived from the average index return and the volatility of the index, while with alphas it derives from the average returns of the risk factors.

Our procedure is not different. We just use a different characterization. Where others use volatility or factor loadings, we use the desired payoff function. Where others use the average return on the index or the chosen risk factors, we use the average interest rate, building block volatilities and correlations over a fund's track record to set a benchmark. What is different, however, is that we do not need to make strong assumptions concerning the exact nature of a fund's risk exposure or the behaviour of markets in general. As shown by KP, a fairly limited set of returns will often be enough to obtain a sufficiently good estimate of the desired distribution and the efficiency measure. As such, our procedure is quite robust.

Another point worth noting about the above evaluation procedure is the fact that it explicitly takes transaction costs into account by, instead of a Black-Scholes type option pricing model, using the Boyle and Lin (1997) model. In factor model based evaluations, transaction costs are typically ignored, despite the fact that maintaining the replicating portfolio's factor loadings at their desired levels is likely to require periodic rebalancing. In addition, when dealing with hedge funds the risk factors used may be quite unusual and may therefore be accompanied by very significant levels of transaction costs.

In the evaluations, we do not use hedge funds' raw returns. The reason is that, as shown in Brooks and Kat (2002) and Lo et al. (2004) for example, monthly hedge fund returns may exhibit high levels of autocorrelation. This primarily results from the fact that many hedge funds invest in illiquid securities, which are hard to mark to market. When confronted with this problem, hedge fund administrators will either use the last reported transaction price or a conservative estimate of the current market price. This creates artificial lags in the evolution of hedge funds' net asset values, i.e. artificial smoothing of the reported returns. As a result, volatility estimates from monthly data will be biased downwards for example.

One possible method to correct for this bias is found in the real estate finance literature. Due to smoothing in appraisals and infrequent valuations of properties, the returns of direct property investment indices suffer from similar problems as hedge fund returns. The approach employed in this literature has been to "unsmooth" the observed returns to create a new set of returns which are more volatile and whose characteristics are believed to more accurately capture the characteristics of the underlying property values. Nowadays, there are several unsmoothing methodologies available. In this study we use the method originally proposed by Geltner (1991).

Finally, the track record of individual hedge funds sometimes contains one or more, typically negative, outliers. The latter can present a problem to the procedure as it may lead one to overestimate the skewness of the desired distribution[4]. For that reason outliers are removed from the sample before running the evaluation procedure[5]. Note that since the majority of outliers in hedge fund returns are negative, this biases our test results in favour of funds that had one or more outliers removed.

## 3. An Example

To clarify the above, let's look at a worked-out example. ABC is a well-known hedge fund, which started in 1987. Given ABC's monthly, net-of-fee returns since 1987, the first step is to model the joint distribution of ABC and the investor's portfolio, as well as the joint distribution of the investor's portfolio and the reserve asset. Before we can do so we need to decide what exactly the investor's portfolio and the reserve asset are, as well as remove possible outliers and unsmooth the raw fund return data.

Let's assume that the representative investor's portfolio consists of 50% S&P 500 and 50% long-dated US Treasury bonds. Let's also assume that all exposure management is done in the futures markets. So instead of investing in the cash market, we will hold fully collateralised (nearby) futures contracts. We use nearby Eurodollar futures as the reserve asset. Futures have several advantages over cash, in particular high liquidity and low transaction costs, which is extremely important given the dynamic nature of the KP replication strategies.

<< Insert Table 1 Here >>

[4] In addition, in an evaluation of the accuracy of the replication, the outlier(s) will pull the conventional skewness measure strongly away from zero, incorrectly suggesting that the replication of the fund return distribution's overall skewness was unsuccessful.

[5] To detect outliers in the left and right tails we use the Dixon (1950, 1951) $r_{10}$ statistic at a 1% significance level.

Taking a closer look at the monthly ABC returns, we notice one clear outlier: -46.2% in October 1987. Since this one extreme observation will have a disproportionate impact on the replication, October 1987 is eliminated from the sample, after which the remaining returns are unsmoothed. Table 1 shows the marginal risk characteristics of the raw and unsmoothed ABC returns, with and without the outlier. It shows that with the outlier the conventional skewness measure will conclude that the fund return distribution exhibits a high degree of negative skewness. After removing the outlier, however, skewness goes up very substantially and even turns slightly positive. From the table, we see that, apart from some positive skewness and excess kurtosis, ABC's raw returns (ex. outlier) exhibit positive autocorrelation. Application of the unsmoothing procedure eliminates the autocorrelation and produces returns with the same degree of skewness and kurtosis, but with a substantially higher volatility (annualised 14.21% vs. 12.47% for the raw returns).

We are now ready to infer the desired and the building block distribution. Using the same methodology as KP, we find that the best fit (according to the AIC) is provided by the following set of marginals and copulas[6]:

**ABC**: Student-t ($\mu = 0.0115$, $\sigma = 0.0402$, df = 5.59)
**Portfolio**: Normal ($\mu = 0.0090$, $\sigma = 0.0261$)
**Reserve**: Johnson ($\xi = 0.0034$, $\lambda = 0.0041$, $\gamma = -0.46$, $\delta = 1.60$)
**Copula (ABC, portfolio)**: Cook-Johnson ($\alpha = 0.62$)
**Copula (portfolio, reserve)**: Gumbel (a = 1.32)

Not unexpectedly, the investor's portfolio is best modelled by a normal distribution. ABC's returns, however, seem best described by a Student-t distribution, while the reserve asset is best modelled by a Johnson distribution. We also see that the relationship of the investor's portfolio with the fund is quite different from its relationship with the reserve asset, with the former modelled as a Cook-Johnson copula and the latter as a Gumbel copula.

[6] Distributions and copulas as defined in Kat and Palaro (2005).

<< Insert Figure 1 Here >>

Given the above distributions, we can derive the desired payoff function following the methodology developed in KP. The result is depicted in Figure 1 and shows that the desired payoff is an increasing function of the reserve asset. The relationship with the investor's portfolio is somewhat more complex though; positive for low values of the investor's portfolio, but turning negative for higher values of the investor's portfolio. This means that the replication strategy will always take a long position in the reserve asset, but may take a short position in the investor's portfolio when the latter and the reserve asset both end up at relatively high levels.

Subsequently, we price this payoff function using the Boyle and Lin (1997) model, assuming transaction costs in the futures markets are 1bp one-way. This produces a value for the KP efficiency measure of 99.89, meaning that, seen over the whole life of the fund, ABC's returns were not as miraculous as many investors may have thought[7]. Trading S&P 500, T-bond and Eurodollar futures, investors could have generated the same risk profile as ABC and obtained a higher average return at the same time.

<< Insert Figure 2 Here >>

To see how well the derived payoff function succeeds in replicating the desired distribution, Figure 2 shows a scatter plot of the investor's portfolio return versus the ABC return (left) as well as a plot of the portfolio return versus the replicated return (right). The two plots are quite similar, suggesting that the replication has indeed been successful. A further indication of the accuracy of the replication strategy comes from comparing the mean, standard deviation, skewness and kurtosis of ABC's returns with those of the replicated returns. The latter statistics can be found in Table 2, together with the correlation and Kendall's Tau with the investor's portfolio. To test whether the marginal distribution of the replicated returns and the joint distribution of the replicated returns and the investor's portfolio are significantly different from the

[7] Note that this is especially true since we deleted the large negative outlier.

original distributions, we use the univariate and bivariate Kolmogorov-Smirnov (K-S) tests[8].

<< Insert Table 2 Here >>

Comparing the entries in Table 2, it is clear that the statistical properties of ABC's returns have been quite successfully replicated. The replication strategy has not only replicated the marginal distribution of ABC's returns but also its relationship with the investor's portfolio. The same conclusion follows from both the K-S tests.

# 4. Data Description

Having introduced the evaluation procedure in detail, we are now ready to turn to the evaluation results. We do so for all funds taken together, as well as for the various strategy groups separately, so we can detect possible differences between them. The strategy classification used and the number of funds within each group can be found in Table 3.

<< Insert Table 3 Here >>

Our total sample consists of 1917 individual hedge funds[9] with a minimum of 4 years of history available. All data were obtained from TASS as of November 2004. Note that this implies that the somewhat disappointing results of 2005 were not taken into account in the evaluation. This is especially relevant for convertible arbitrage funds, which performed quite badly in 2005. Funds denominated in another currency than USD were converted to USD, i.e. the perspective taken is that of a USD-based investor. Table 4 and 5 provide some information on the start and end dates of the track records of the funds in our sample.

<< Insert Table 4 and 5 Here >>

---

[8] See Fasano and Franceschini (1987) for details.

[9] The performance of funds of funds was evaluated separately earlier. See Kat and Palaro (2006) for details and results.

Table 4 shows that, reflecting the increasing popularity of hedge funds in the second half of the 1990s, the majority of funds started after 1994. Only 19 funds started before January 1985. Most hedge fund databases, including TASS, first started collecting data around 1994. As a result, our sample contains no funds that stopped reporting before that date. This is shown in Table 5, from which we also see that out of 1917 funds, no less than 847 funds stopped reporting before October 2004. This confirms that the attrition rate in hedge funds is quite substantial[10].

<< Insert Table 6 Here >>

Table 6 provides details on the length of the available hedge fund track records. Out of the 1917 funds in the sample, only 359 have more than 10 years of history. This again reflects the fact that most funds are still relatively young and attrition levels can be very significant.

As in the example in section 3, in the evaluations we assume that the representative investor's portfolio consists of 50% S&P 500 and 50% long-dated US Treasury bonds, with all exposure management done through fully collateralised (nearby) futures contracts[11]. Since it is one of the most traded futures contracts in the world, we use nearby Eurodollar futures (trading on the CME) as the reserve asset. Transaction costs on all futures contracts are assumed to be 1bp one-way. For the pricing of the payoff functions, we use 1-month USD Libor as the relevant interest rate, while estimating the required volatilities and correlations over the period covered by the track record of the fund that is being evaluated. The interest rate data was obtained from Datastream. The futures data was obtained from Commodity Systems Inc. (CSI).

---

[10] Not all funds that stop reporting into a database do so because they close down. The majority does so, however. For more details on hedge fund and fund of funds attrition see Kat and Amin (2003a).

[11] More in particular, we trade S&P 500 futures on the CME and T-bond futures on the CBOT. Both contracts are in the top 10 of most traded futures contracts in the US.

## 5. Evaluation Results

Before we present the evaluation results, it is interesting to take a closer look at the accuracy of the replication procedure. We therefore plotted the fund standard deviation, conventional skewness and correlation with the investor's portfolio versus the replicated values for all 1917 funds. The results are shown in Figure 3-5. As is clear from these graphs, on average the replication of these parameters is unbiased and quite accurate. Not surprisingly, the replication of skewness can be somewhat difficult at times as the conventional skewness measure is extremely sensitive. Taking that into account, however, the results in Figure 4 are quite satisfactory.

<< Insert Figure 3 - 5 Here >>

Figures 3-5 also provide additional information on the risk-return profile of the funds in our sample. From Figure 4 for example, we see that for most hedge funds estimated skewness lies somewhere between –1 and +1. Likewise, from Figure 5 we see that the majority of hedge funds are positively correlated with a portfolio of 50% stocks and 50% bonds. Most correlation coefficients lie between 0 and 0.6, indicating that many hedge funds' returns are a lot less 'market neutral' than the popular term 'absolute returns' suggests[12].

<< Insert Table 7 Here >>

We tested the statistical significance of the KP efficiency measure results by calculating bootstrapped confidence intervals, distinguishing between three cases: (1) Inefficient, i.e. confidence interval entirely lower than 100, (2) Efficient, i.e. confidence interval entirely higher than 100, and (3) Equivalent, i.e. confidence interval contains 100. Table 7 summarizes the evaluation outcomes. From the table we see that the majority of funds produce a value for the KP measure that is below 100. In other words, over the period under consideration, the majority of hedge funds have not provided their investors with returns, which they could not have generated themselves in the futures market.

---

12 In this context it is important to note that at least for some of the more complex distributions encountered, the correlation coefficient will not be a particularly good measure of dependence and may underestimate the true level of dependence.

The percentage of efficient funds varies considerably between the different strategy groups, with dedicated shorts producing the least (3.8%) and the category 'other' the most (47.8%) efficient funds. Convertible arbitrage (34.7% efficient) and fixed income (43.5% efficient) stand out as well. When interpreting these results, one has to keep in mind that the available dataset on hedge funds is limited and that, as shown in Table 6, most funds have relatively short track records. The idea behind the KP measure is that in the longer run investors receive a return that is fair compensation for the bottom-line risk that they have taken, irrespective of how that risk profile is obtained. For many hedge funds, however, we may not have enough data to be able to properly observe 'the longer run'. The shorter the track record, the more the efficiency measure may be influenced by sampling error[13], in both the fund and the assets traded in the replication strategy. The relatively high proportion of efficient funds in convertible and fixed income arbitrage for example may have been partly due to the combination of falling interest rates and shrinking credit spreads observed over recent years.

The last three columns of Table 7 show the mean, standard deviation and skewness of the frequency distribution of the efficiency measure values observed within each strategy group. For all strategies the distribution is negatively skewed, implying that within each group some funds have shown extremely bad performance relative to what could have been achieved trading S&P 500, T-bond and Eurodollar futures. This should be taken into account when interpreting the mean values. From the table we also see that especially the efficiency measures of strategies whose returns are known to be relatively volatile exhibit a relatively high standard deviation. Managed futures, global macro and emerging markets exhibit relatively high standard deviations, while the opposite is true for convertible arbitrage and equity market neutral. Likewise, highly negative skewness is observed in exactly those strategies that are known to be most susceptible to shocks, i.e. convertible arbitrage, event driven, fixed income and global macro.

[13] Note that this applies to all performance evaluation procedures, not just the KP measure.

<< Insert Figure 6 Here >>

Figure 6 shows the frequency distribution of the efficiency measure values found in all 1917 hedge funds[14]. Since lack of performance is one of the main reasons for hedge funds to close down, Figure 6 also separates out funds that stopped reporting to the database before October 2004. From the graph we see that there is a strong relationship. Out of the 567 funds with a KP measure below 99, no less than 337 (60%) stopped reporting. Out of the 427 funds with a KP measure higher than 100, only 173 (40%) did so, with probably a significant number of these not really closing down but simply stopping reporting. A similar relationship is observed in the average efficiency measure values of live and dead funds. The average KP measure over the 848 dead funds is 98.94. If we assume that all funds with a KP measure above 100 did not really die, but simply stopped reporting, the average KP measure for dead funds drops to 98.54. Over the 1069 funds still alive on the other hand, the average is 99.44.

<< Insert Table 8 Here >>

To investigate whether there is any indication of older funds doing better than younger funds or vice versa, we sorted the funds in our sample on the length of their track record. Table 8 shows the statistics of the resulting frequency distributions of the KP measure. From the means we see that on average age has no impact on performance. The standard deviation, skewness, and kurtosis measures all drop when we move to funds with a longer track record. This is not surprising as with more data available, sampling error will tend to be less of an issue, which is reflected in the declining dispersion of observed KP measure values.

<< Insert Figure 7 Here >>

Another question concerns the performance of hedge funds through time. Especially after two years of somewhat disappointing results, it is often claimed that overall hedge fund performance is deteriorating, with the massive inflow of capital over recent years being the most obvious cause. We therefore split the track record of all

[14] Histograms for the various strategies show a similar picture and are therefore not reported.

funds with 8 or more years of history in two equal parts and calculated the KP measure over each part. Over all funds, the average over the first period was 100.56, while over the second, more recent, period the average was only 98.28. This indicates a substantial deterioration in hedge fund performance over time. In addition, Figure 7, which shows a plot of the results for the two sub-periods, reveals a tendency for funds with a relatively high (low) KP measure in the first period to produce a relatively low (high) KP measure in the second. Subdividing between strategy groups (not reported) this phenomenon appears strongest in emerging markets, equity market neutral and global macro, and weakest in convertible arbitrage and managed futures[15]. As the assets under management of funds that do well can be expected to grow substantially (organically as well as through additional inflows) and vice versa, this finding supports the idea that increased fund size has a negative impact on future performance.

## 6. Distributional Analysis

A crucial stage in the evaluation procedure is the proper modelling of the distributional characteristics of the fund, the investor's portfolio and the reserve asset. This means that, although not explicitly designed to do so, the evaluations provide a wealth of information on the distributional properties of fund of funds returns. Table 9 summarizes how often a given marginal distribution or copula was used in the evaluations for modelling the fund return marginal and the joint distribution of the fund and the investor's portfolio return.

<< Insert Table 9 Here >>

Table 9 confirms that the majority of individual hedge fund returns are far from normally distributed. Out of 1917 funds, 1374 funds' marginal return is better modelled by a Student-t or Johnson distribution than a normal distribution. In addition, for only 334 of the 1917 funds is the relationship with the investor's

[15] Note, however, that the 2005 performance of convertible arbitrage and managed futures, which was not included in the evaluations, was quite bad, with the Barclay Convertible Arbitrage Index for example dropping 3.28% and the Barclay CTA Index up by only 1.71% .

portfolio of 50% S&P 500 and 50% T-bonds best modelled by the normal copula. This emphasizes once more how important it is to evaluate hedge fund performance using a methodology, which does not rely on the assumption of normally distributed returns.

## 7. Conclusion

In this paper we have used the hedge fund return replication technique recently introduced in Kat and Palaro (2005) to evaluate the net-of-fee performance of 1917 individual hedge funds. The results indicate that the majority of hedge funds have not provided their investors with returns, which they could not have generated themselves by mechanically trading S&P 500, T-bond and Eurodollar futures. Overall, only 17.7% of the funds studied beat the benchmark. Over time, we observe a substantial deterioration in overall hedge fund performance. In addition, we find a tendency for the performance of successful funds to deteriorate over time. This supports the hypothesis that increased fund size may hurt future performance.

Overall, only 17.7% of the 1917 funds in our sample were able to beat the benchmark. Compared to the fund of funds results reported in Kat and Palaro (2006), where it was found that only 11.3% of funds of funds were efficient, this means that in terms of the KP measure individual hedge funds and funds of hedge funds are not too different. At first sight this may seem odd. With funds of funds putting on an additional layer of fees, one would expect the results for funds of funds to be substantially worse than for individual hedge funds. However, funds of funds diversify and given the low correlation between hedge funds, this means that the risk characteristics of fund of funds returns are typically a lot more conservative than those of individual hedge funds, which is reflected in the efficiency measure outcomes.

Compared with the various hedge fund performance evaluation studies that have been carried out over the last couple of years, our results are quite unusual. Often, the conclusion from hedge fund performance studies is that hedge funds generate superior returns, not inferior. This once again indicates how tricky factor model based performance evaluation can be. As long as one can't be sure that all relevant risk factors are accounted for, it is impossible to know whether unexplained returns are indeed true alpha or just unexplained because one or more risk factors were left out or specified incorrectly. Our methodology is more robust, as it relies on a simple principle: "if it can be replicated, it can't be superior". Of course, we need to make assumptions as well, but these are less crucial for the final outcome of the evaluation than the kind of assumptions required to make factor model based alphas work.

Should investors rush out to buy into those funds with the highest KP measures? Although tempting, the answer is no. The core problem of performance evaluation is separating luck and skill. With a limited set of data, however, it is impossible to make a clean cut, whatever the method used. The KP measure is founded on the idea that in the longer run, risk and return are related, irrespective of how a given risk profile is obtained. When there are not enough data available to properly observe 'the longer run', however, the efficiency measure becomes prone to sampling error. If the available dataset is limited, it is very hard to identify the presence of any extreme (but compensated) risks for example, since by definition extreme events only occur infrequently. A fund manager may have been taking the most horrific risks, but if so far he has been lucky, the premium collected for taking on those risks will show from his track record, but the risk won't. Likewise, one or more risk factors may have done extremely well over a prolonged period of time. This will bias the available sample, which in turn may have a significant impact on the outcome of the evaluation.

Since performance evaluations over relatively short time periods will always leave us with a considerable degree of uncertainty, a high KP measure should first and foremost be interpreted as a signal that further due diligence is warranted. One can only speak of truly superior performance if such follow-up research shows that the good evaluation outcome was not simply due to luck. In other words, that the manager in question has generated the observed excess return without taking any extreme risks and that all the relevant risk factors behaved in a more or less representative manner during the period under consideration. Questions like these can typically not be answered satisfactorily within a purely quantitative framework and require a thorough understanding of hedge fund strategies. No matter how sophisticated the econometrics, proper performance evaluation will therefore always remain a combination of science and art.

# References


**Akaike, H. (1973).** Information Theory and an Extension of the Maximum Likelihood Principle, in: B. Petrov and F. Csaki (eds.), *Second International Symposium on Information Theory*, Academiae Kiado, Budapest, pp. 267-281.

**Amin, G. and H. Kat (2003a).** Welcome to the Dark Side: Hedge Fund Attrition and Survivorship Bias 1994-2001, *Journal of Alternative Investments*, Summer, pp. 57-73.

**Amin, G. and H. Kat (2003b).** Hedge Fund Performance 1990-2000: Do the Money Machines Really Add Value? *Journal of Financial and Quantitative Analysis*, Vol. 38, pp. 1-24.

**Boyle, P. and X. Lin (1997).** Valuation of Options on Several Risky Assets When There are Transaction Costs, in: P. Boyle, G. Pennacchi and P. Ritchken (eds.), *Advances in Futures and Options Research,* Vol. 9, Jai Press, pp. 111-127.

**Brooks, C. and H. Kat (2002).** The Statistical Properties of Hedge Fund Index Returns and Their Implications for Investors. *Journal of Alternative Investments,* Fall, pp. 26-44.

**Crow, E. and M. Siddiqui (1967).** Robust Estimation of Location. *Journal of the American Statistical Association*, Vol. 62, pp. 353-389.

**Dixon, W. (1950).** Analysis of extreme values. *Annals of Mathematical Statistics,* Vol. 21, pp. 488-506.

**Dixon, W. (1951).** Ratios involving extreme values. *Annals of Mathematical Statistics,* Vol. 22, pp. 68-78.

**Fasano, G. and A. Franceschini (1987).** *Monthly Notices of the Royal Astronomical Society*, Vol. 225, pp. 155-170.

**Geltner, D. (1991).** Smoothing in Appraisal-Based Returns. *Journal of Real Estate Finance and Economics*, Vol. 4, pp. 327-345.

**Hinkley, D. (1975).** On Power Transformations to Symmetry. *Biometrika*, Vol. 62, pp. 101-111.

**Kat, H. and J. Miffre (2005).** Hedge Fund Performance: The Role of Non-Normality Risks and Conditional Asset Allocation, Alternative Investment Research Centre Working Paper No. 26, Cass Business School, City University, London.

**Kat, H. and H. Palaro (2005).** Who Needs Hedge Funds? A Copula-Based Approach to Hedge Fund Return Replication, Alternative Investment Research Centre Working paper No. 27, Cass Business School, City University, London.

**Kat, H. and H. Palaro (2006).** Replication and Evaluation of Fund of Hedge Funds Returns, Alternative Investment Research Centre Working paper No. 28, Cass Business School, City University, London.

**Lo, A., M. Getmansky and I. Makarov (2004).** An Econometric Analysis of Serial Correlation and Illiquidity in Hedge Fund Returns. *Journal of Financial Economics*, Vol. 74, pp. 529-609.

| | Standard Deviation | Skewness | Excess Kurtosis | 1M Auto Correlation |
|---|---|---|---|---|
| **ABC smooth, incl. outlier** | **0.0487** | **-4.154** | **42.583** | **0.156** |
| **ABC smooth, outlier removed** | **0.0360** | **0.461** | **4.598** | **0.143** |
| **ABC unsmooth, outlier removed** | **0.0410** | **0.507** | **4.475** | **0.011** |

**Table 1: Risk statistics ABC.**

| | Mean | St. Dev | Skewness | Excess Kurtosis | Corr. with Portfolio | Kendall's Tau |
|---|---|---|---|---|---|---|
| **ABC** | **0.0115** | **0.0410** | **0.5074** | **4.4749** | **0.359** | **0.254** |
| **Replica** | **0.0127** | **0.0412** | **0.5453** | **1.4493** | **0.398** | **0.284** |
| **Univariate K-S Statistic = 0.053, (approximated) p-value = 0.925**<br>**Bivariate K-S Statistic =0.065, (approximated) p-value = 0.901** | | | | | | |

**Table 2: Statistics ABC and replicated returns.**

| Category | No. funds |
|---|---|
| **Convertible Arbitrage** | **98** |
| **Dedicated Short Bias** | **26** |
| **Emerging Markets** | **176** |
| **Equity Market Neutral** | **94** |
| **Event Driven** | **209** |
| **Fixed Income Arbitrage** | **85** |
| **Global Macro** | **111** |
| **Long/Short Equity** | **758** |
| **Managed Futures** | **293** |
| **Other** | **67** |
| **Total** | **1917** |

**Table 3: Strategy classification and number of hedge funds.**

| Start after: | Jan 1985 | Jan 1988 | Jan 1991 | Jan 1994 | Jan 1997 | Jan 2000 | Jan 2003 |
|---|---|---|---|---|---|---|---|
| Convertible Arb. | 98 | 98 | 89 | 72 | 44 | 12 | 0 |
| Dedicated Short | 26 | 24 | 19 | 14 | 7 | 5 | 0 |
| Emerging Markets | 176 | 175 | 166 | 132 | 72 | 13 | 0 |
| Equity Neutral | 94 | 94 | 90 | 83 | 58 | 9 | 0 |
| Event Driven | 207 | 202 | 185 | 149 | 88 | 23 | 0 |
| Fixed Income Arb. | 84 | 83 | 82 | 70 | 38 | 8 | 0 |
| Global Macro | 110 | 105 | 93 | 66 | 35 | 12 | 0 |
| Long/Short Equity | 748 | 733 | 697 | 605 | 407 | 113 | 0 |
| Managed Futures | 288 | 271 | 232 | 149 | 72 | 10 | 0 |
| Other | 67 | 66 | 61 | 51 | 33 | 7 | 0 |
| Total | 1898 | 1851 | 1714 | 1391 | 854 | 212 | 0 |

**Table 4: Hedge funds start date details.**

| End before: | Jan 1994 | Jan 1997 | Jan 2000 | Jan 2003 | Oct 2004 |
|---|---|---|---|---|---|
| Convertible Arb. | 0 | 6 | 10 | 19 | 33 |
| Dedicated Short | 0 | 2 | 3 | 8 | 12 |
| Emerging Markets | 0 | 2 | 20 | 61 | 85 |
| Equity Neutral | 0 | 0 | 5 | 15 | 36 |
| Event Driven | 0 | 1 | 10 | 49 | 85 |
| Fixed Income Arb. | 0 | 1 | 12 | 29 | 42 |
| Global Macro | 0 | 6 | 20 | 46 | 56 |
| Long/Short Equity | 0 | 8 | 39 | 156 | 297 |
| Managed Futures | 0 | 22 | 90 | 150 | 182 |
| Other | 0 | 5 | 6 | 8 | 19 |
| Total | 0 | 53 | 215 | 541 | 847 |

**Table 5: Hedge funds end date details.**

| Length Track record: | 4-5Y | 5-6Y | 6-7Y | 7-8Y | 8-9Y | 9-10Y | 10-12Y | 12-14Y | 14+ |
|---|---|---|---|---|---|---|---|---|---|
| Convertible Arb. | 19 | 9 | 17 | 13 | 11 | 8 | 13 | 6 | 2 |
| Dedicated Short | 5 | 5 | 0 | 6 | 0 | 1 | 4 | 1 | 4 |
| Emerging Markets | 36 | 32 | 23 | 27 | 18 | 16 | 17 | 5 | 2 |
| Equity Neutral | 24 | 19 | 16 | 9 | 6 | 10 | 5 | 4 | 1 |
| Event Driven | 40 | 41 | 18 | 18 | 16 | 19 | 26 | 11 | 20 |
| Fixed Income Arb. | 26 | 14 | 11 | 10 | 9 | 6 | 6 | 1 | 2 |
| Global Macro | 18 | 16 | 17 | 16 | 13 | 9 | 9 | 6 | 7 |
| Long/Short Equity | 206 | 120 | 100 | 86 | 71 | 49 | 55 | 28 | 43 |
| Managed Futures | 64 | 42 | 34 | 37 | 26 | 23 | 28 | 14 | 25 |
| Other | 16 | 10 | 9 | 6 | 9 | 3 | 11 | 2 | 1 |
| Total | 454 | 308 | 245 | 228 | 179 | 144 | 174 | 78 | 107 |

**Table 6: Length hedge funds track records.**

| | No. | Inefficient | Equivalent | Efficient | Mean | StDev | Skew |
|---|---|---|---|---|---|---|---|
| Convertible Arbitrage | 98 | 45 (45.9%) | 19 (19.4%) | 34 (34.7%) | 99.830 | 0.597 | -4.212 |
| Dedicated Short Bias | 26 | 19 (73.1%) | 6 (23.1%) | 1 (3.8%) | 99.245 | 0.877 | -2.835 |
| Emerging Markets | 176 | 123 (69.9%) | 18 (10.2%) | 35 (19.9%) | 98.821 | 1.663 | -1.595 |
| Equity Market Neutral | 94 | 72 (76.6%) | 7 (7.4%) | 15 (16.0%) | 99.500 | 0.709 | -1.606 |
| Event Driven | 209 | 123 (58.9%) | 34 (16.3%) | 52 (24.9%) | 99.636 | 1.096 | -6.301 |
| Fixed Income Arbitrage | 85 | 34 (40.0%) | 14 (16.5%) | 37 (43.5%) | 99.573 | 1.274 | -4.326 |
| Global Macro | 111 | 87 (78.4%) | 12 (10.8%) | 12 (10.8%) | 99.062 | 1.710 | -6.465 |
| Long/Short Equity | 758 | 600 (79.2%) | 70 (9.2%) | 88 (11.6%) | 99.077 | 1.148 | -1.597 |
| Managed Futures | 293 | 240 (81.9%) | 20 (6.8%) | 33 (11.3%) | 98.842 | 1.289 | -1.899 |
| Other | 67 | 30 (44.8%) | 5 (7.5%) | 32 (47.8%) | 99.417 | 1.358 | -2.753 |
| Total | 1917 | 1373 (71.62%) | 205 (10.69%) | 339 (17.68%) | 99.218 | 1.239 | -2.884 |

**Table 7: Overall hedge fund evaluation results.**

| Track Record (months) | Mean | StDev | Skew | Kurt |
|---|---|---|---|---|
| 48 - 71 | 99.12 | 1.42 | -3.03 | 19.66 |
| 72 - 95 | 99.19 | 1.27 | -2.71 | 13.55 |
| 96 - 119 | 99.33 | 1.06 | -1.96 | 8.22 |
| 120 - 143 | 99.39 | 0.89 | -1.30 | 1.50 |
| 144+ | 99.32 | 0.86 | -1.33 | 3.58 |

**Table 8: Distribution KP measure as function length of track record.**

| | Marginal Distribution | | | Copula | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Norm | Stud | John | Norm | Stud | Gumb | SJC | Cook-J | Frank |
| Convertible Arbitrage | 20.41% | 39.80% | 39.80% | 20.41% | 3.06% | 6.12% | 33.67% | 25.51% | 11.22% |
| Dedicated Short Bias | 30.77% | 61.54% | 7.69% | 53.85% | 7.69% | 0.00% | 3.85% | 0.00% | 34.62% |
| Emerging Markets | 23.86% | 47.16% | 28.98% | 14.20% | 1.70% | 10.23% | 10.80% | 31.25% | 31.82% |
| Equity Market Neutral | 36.17% | 29.79% | 34.04% | 21.28% | 7.45% | 12.77% | 22.34% | 9.57% | 26.60% |
| Event Driven | 12.44% | 47.37% | 40.19% | 10.53% | 5.26% | 4.31% | 28.23% | 38.28% | 13.40% |
| Fixed Income Arbitrage | 10.59% | 42.35% | 47.06% | 11.76% | 5.88% | 10.59% | 34.12% | 17.65% | 20.00% |
| Global Macro | 29.73% | 31.53% | 38.74% | 16.22% | 9.01% | 14.41% | 18.92% | 12.61% | 28.83% |
| Long/Short Equity | 30.60% | 47.23% | 22.16% | 20.05% | 7.12% | 9.63% | 9.10% | 19.79% | 34.30% |
| Managed Futures | 42.66% | 36.17% | 21.16% | 14.68% | 27.65% | 20.48% | 7.85% | 9.22% | 20.14% |
| Other | 20.90% | 40.30% | 38.81% | 14.93% | 7.46% | 7.46% | 28.36% | 5.97% | 35.82% |
| Total | 28.33% | 43.14% | 28.54% | 17.42% | 9.44% | 10.85% | 15.34% | 19.77% | 27.18% |

**Table 9: Overall distributional characteristics hedge fund returns.**

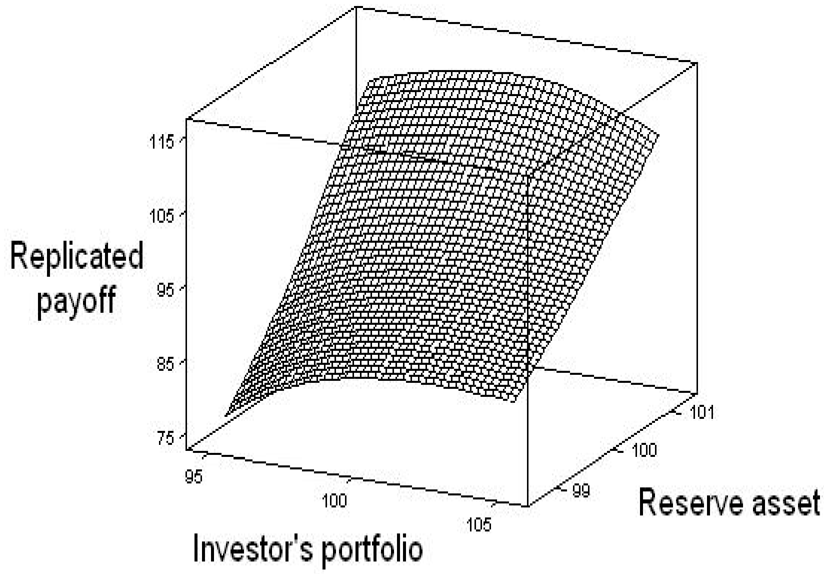


**Figure 1: Desired payoff function for replication ABC returns.**

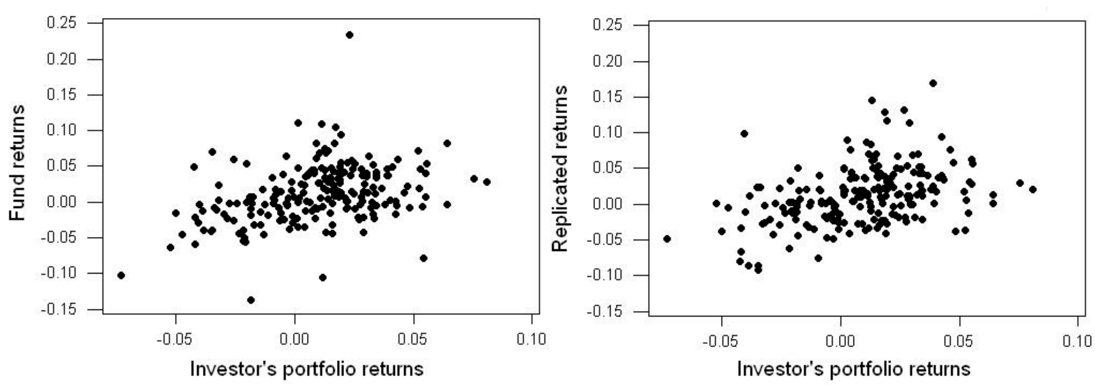


**Figure 2: Scatter plot investor's portfolio returns versus ABC returns (left) and replicated returns (right).**

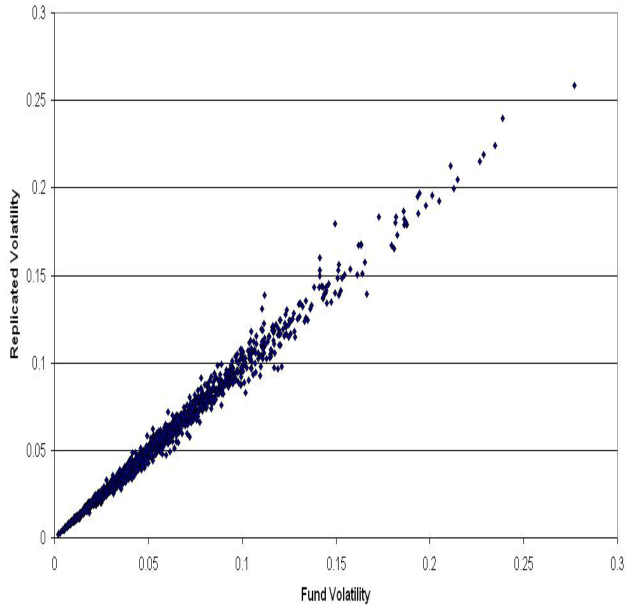


**Figure 3: Scatter plot hedge fund vs. replicated standard deviation.**

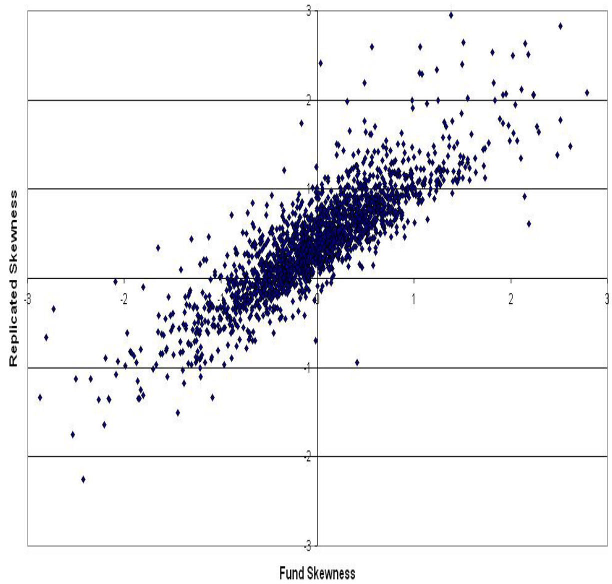


**Figure 4: Scatter plot hedge fund vs. replicated (conventional) skewness.**

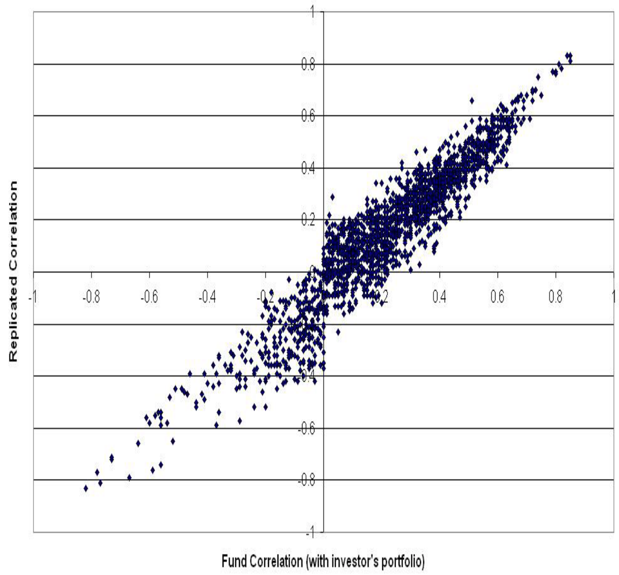


**Figure 5: Scatter plot hedge fund vs. replicated correlation with investor's portfolio.**

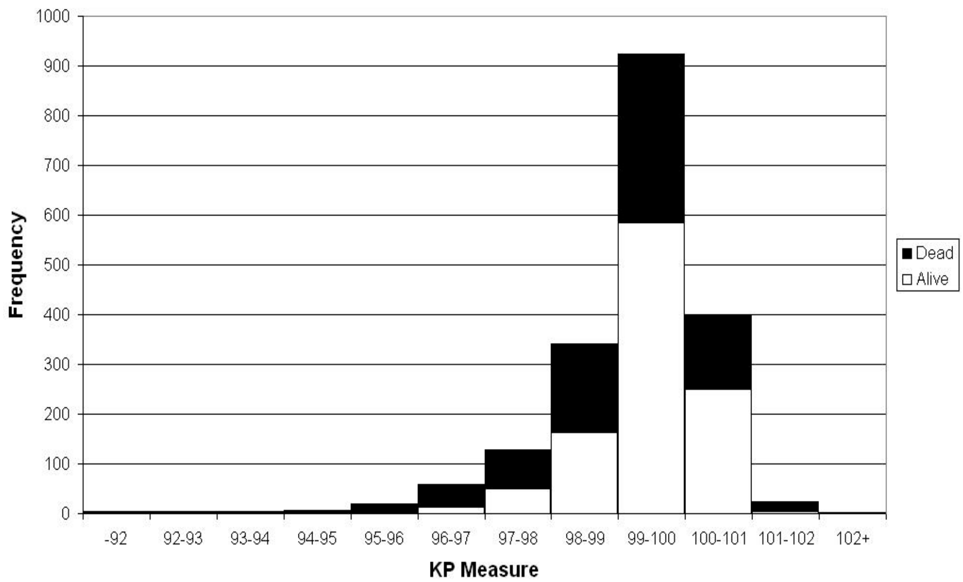


**Figure 6: Number of live and dead funds as function KP efficiency measure.**

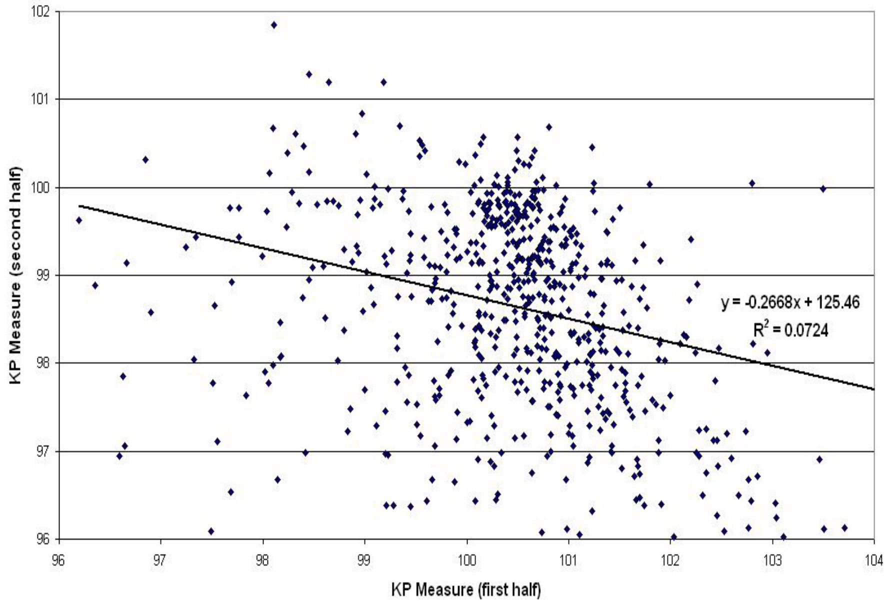


**Figure 7: KP efficiency measures over split track record.**