# The Art of Investing in Hedge Funds: Fund Selection and Optimal Allocations

**Carol Alexander^ and Anca Dimitriu***

ISMA Centre, University of Reading, UK

*Preliminary version: January 2004*

*Abstract*

With institutional investors increasingly involved in alternative investments, portfolio optimisation within a large universe of hedge funds has become a key area for research. This paper develops a portfolio construction model that is specifically designed for funds of hedge funds, incorporating specific controls for operational limitations, data biases and incompleteness. Absolute performance is targeted by selecting funds' according to their alpha estimated with factor models. Whilst different factor models provide quite different estimates of a hedge fund's alpha, we can still use the ranking produced by them in the fund selection process. In an extensive out-of-sample historical analysis, funds of funds that are selected in this way and then allocated using constrained minimum variance optimisation are shown to perform much better than the equally weighted portfolio of all funds, or minimum variance portfolios of randomly selected funds. This is true even when hedge funds are selected according to their alphas produced by simple factor models. The best out-of-sample performance is obtained, out of the four factor models considered in this analysis, with the statistical factor model.

*EFM classification codes: 370, 380*

^ Carol Alexander, ISMA Centre, The University of Reading, Whiteknights Park, Reading RG6 6BA, UK, Tel +44 (0)118 9316494 Fax +44 (0)118 9314741, Email: c.alexander@ismacentre.rdg.ac.uk

* Anca Dimitriu, ISMA Centre, The University of Reading, Whiteknights Park, Reading RG6 6BA United Kingdom, Tel +44 (0)118 9316494 Fax +44 (0)118 9314741, Email: a.dimitriu@ismacentre.rdg.ac.uk

## Introduction

During the bear stock market of the last few years institutional investors as well as high net worth individual investors have found a new source of returns in alternative investment strategies. A rough comparison illustrates the difference in performance between traditional and alternative investments: over the period Jan-00 to May-03, the Wilshire 5000 equity index lost on average 10.5% per year, while the HFR fund weighted composite index gained on average 4.7% per year, with much lower volatility. Despite some high-profile losses, investors are committing more assets to the alternative investment industry. According to TASS Research, in the second quarter of 2003, the net flow into the hedge fund industry was estimated at 13.83 billion USD.

However, there is no 'free lunch'. Due to the myriad of strategies employed by hedge funds, their highly dynamic nature and the extensive use of derivatives and leverage inducing non-linear relationships with the traditional asset classes, models for hedge fund returns are inherently complex. This complexity is compounded by data reliability issues and operational limitations. In terms of data, the fact that hedge funds report to commercial database providers on a voluntary basis creates a number of sampling issues. First, even if one combines the largest commercial databases available,[1] there is no indication of the size of the population of hedge funds and the degree to which the reporting funds are representative of this population. Beyond the small sample bias, the databases comprise performance estimates rather than liquid market prices and this generates both autocorrelation and a significant amount of noise in the data. Regarding the operational limitations, short sales are not possible, there are minimum investment limits, long lock-up periods and advance notice, regular subscriptions/redemption as rare as once per year, sales and early redemption fees.

Despite these difficulties there is no doubt that alternative investments present attractive opportunities and their popularity is increasing. Thus currently, one of the most challenging problems in portfolio management is to refine the traditional portfolio models to optimise investments in a large universe of hedge funds. Given the special features of hedge funds, the portfolio management tools developed for traditional investments, being heavily reliant on the existence of liquid markets and efficient prices, need refining. In these circumstances, hedge fund portfolio management is as close to an art as any field in finance can be. High quality data is an essential ingredient for any portfolio optimisation model but this is not likely to be achieved for

[1] In 2003 Fauchier Partners counted 4,589 funds reporting to at least one of the three most important databases. providers, TASS, HFR and CISDM.

hedge funds in the near future. We are not advocating informal or ad-hoc solutions. The most one can hope for is that a skilful application of refinements to traditional portfolio management tools with limited data can produce a better solution than a naive diversification.

The aim of this paper is to investigate the features of the investment process in hedge funds, considered to be a source of alpha that can be further transported to other asset classes.[2] To this end, we analyse the out-of-sample performance of a portfolio construction model designed to target alpha through fund selection based on factor models and asset allocation based on traditional optimisation. We follow through all the steps of portfolio management: the set-up of the database and handling the biases in the data; the estimation of 'alpha', the risk adjusted performance; the detection of the 'true' correlation structure of fund returns; optimal rebalancing and the out-of-sample portfolio performance assessment. Following standard practice we use a database of both dead and alive funds to diminish the impact of survivorship bias. We account for the instant history bias through the use of dummy variables in factor models and report performance on a relative basis, benchmarked against portfolios that are affected by the same biases. Operational limitations are addressed by imposing constraints on the optimisation and by including an estimate of annual turnover as a key diagnostic of the portfolio's performance.

Traditional portfolio optimisation models require forecasts of the portfolio expected returns and/or an estimate of their covariance matrix. Often expected returns are estimated using a factor model, so difficulties arise when there is no consensus on the most appropriate factor model. Given the highly dynamic and heterogeneous styles used in alternative investments, the non-uniqueness of the factor model representation is one of the most important problems to address when optimising portfolios of hedge funds. Amenc and Martellini (2003a) show that different factor models can generate very different estimates of a hedge fund's 'alpha' and consequently argue that the hedge fund industry should promote its diversification potential rather than its debatable alpha benefits.

We use four different factor models to estimate the alpha of a hedge fund: our 'base case' model is the simplest representation of the fund returns as a function of the two most important underlying asset classes, equities and bonds; the 'broad fundamental' factor model employs indices to capture the performance of the main asset classes, and other factors representing

[2] In a forthcoming paper we investigate the diversification, or beta benefits of the hedge funds universe, considered also a potential destination of alpha.

specific types of non-linear strategies such as market timing, volatility trading and equilibrium trading; the 'multi-factor' model is based on hedge fund indices; and finally the 'statistical' factor model is based on factors extracted from fund returns through principal component analysis. The alpha estimates from the four factor models are used in the fund selection process and subsequently we show that the best allocation optimiser is based on a minimum variance setting.

Our hedge fund selection process is determined by the fund's rank and not by its absolute alpha estimate from a factor model. Different factor models are shown to generate very different alphas, yet we find significant agreement on the ranking of funds based on their alphas from different factor models. Interestingly, Amenc and Martellini (2003a) report similar results using a different database. The implication for the construction of portfolios of hedge funds is that, whilst models that depend on the accuracy of alpha estimates have significant model risk, there is considerable scope for models that are based only on the ranking of funds.

The classic minimum variance portfolio optimisation is based solely on the funds' covariance matrix. We find that this produces better results than a mean-variance maximum information ratio optimisation. As expected, given the reporting requirements of hedge fund databases, there is a high degree of randomness in the sample covariance matrix. Some authors advocate 'cleaning' this by imposing some factor structure before using it in the portfolio optimisation (Plerou et al., 2002; Amenc and Martellini, 2002). However portfolio weight constraints are essential for hedge funds investing and we show that with a constrained portfolio optimisation there appears to be no further benefit from imposing a factor structure on the correlation matrix, as the sampling errors appear to already be significantly reduced through the weights' constraining.

This approach results in a greatly improved performance compared to both randomly selected minimum variance portfolios and to the equally weighted portfolio of all hedge funds. The out-of-sample performance of this class of models highlights considerable alpha opportunities in the alternative investment world that can be exploited using simple constrained optimisation models.

The reminder of the paper is organised as follows: section one describes the hedge fund data and the treatment of biases, section two introduces the factor models used for measuring relative abnormal return and reports the estimation results, section three presents several portfolio optimisation models and the out-of-sample performance analysis, and section four summarises and concludes.

## I Hedge Fund Data and Biases

Hedge fund data are subject to several measurement biases caused by the data collection process and by the nature of the industry: *survivorship bias*, when a database does not include the performance of funds that ceased operating during the sample period;[3] *selection or self-reporting bias*, when the hedge funds in the database are not representative of the population of hedge funds;[4] *instant history bias*, when the funds entering the database are allowed to back-fill their results;[5] and *multi-period sampling bias*, when the analysis is restricted to funds having a minimum amount of history available. These biases are commonly estimated as the difference between the average returns in the database of funds controlled for the feature causing the bias (e.g. including only funds which have survived at the end of the period, or only funds which have a minimum amount of history available) and the average returns of funds in the complete database.

Fung and Hsieh (2000) provide an extensive analysis of biases in the TASS hedge fund database. They estimate a survivorship bias of approximately 3% per annum.[6] Regarding the instant history bias they found an average incubation (back-filled) period of one year with an associated bias of 1.4% p.a., while the multi-period sampling bias was negligible. In order to reduce selection bias, Fung and Hsieh (2002) recommend the use of indices of funds of funds to represent the hedge fund investment experience because it is less susceptible to selection bias than the returns on individual hedge funds.

Our fund data comes from the Hedge Fund Research (HFR) dead and alive funds databases, from which we select the period January 1990 to May 2003. We restrict our analysis to US domiciled funds reporting net of all fees in USD, having funds under management above 10 million USD and not using leverage. Additionally, since we shall be estimating multi-factor models of each

[3] Most database providers only started collecting data in the early or mid-90's so there is little information on funds that ceased operating before then. The early years of any hedge fund database will implicitly have a high survivorship bias.
[4] Selection bias is the most difficult to estimate because the population of hedge funds is not observable. The main reason for funds to report their performance is to attract investors so funds will report as long as they have not reached full investment capacity and continue to achieve an attractive performance. Funds cease reporting for a variety of reasons, including poor performance or liquidation, reorganisation, or because they have reached capacity and are no longer interested in new investors. Thus the selection bias can work both ways – some funds report superior performance in order to attract new investors, which creates an upward bias, while other funds, having reached full capacity as a result of extraordinary performance, choose to stop reporting, thus creating a downward bias. Also, different data providers have different criteria for including a fund in their database and this can induce additional selection biases.
[5] Most funds enter the database with an existing track record from their incubation period or previous organisational form and their results are usually back-filled over a year or several months. Since funds can choose, post factum, the moment when they enter the database, the back-filled history is likely to be affected by an upward bias in results.
[6] Percentage of funds that have ceased to report during a given period, relative to the number of funds existing at the beginning of the data period.

fund's excess returns, we must require that each fund has at least 60 months of reporting history. After imposing these selection criteria our database comprises 282 funds of which 55 had ceased reporting before May 2003.

The attrition rate for the HFR database is known to be smaller than that of other databases so the HFR survivorship bias could be higher. We include in our analysis the funds that ceased reporting during the sample period but this does not ensure that the portfolio performance is identical to the experience of an investor in these funds because there is no information on the performance of funds after having ceased reporting. If some funds were liquidated, their investors probably recovered only part of the net asset value last reported. Therefore by including 'dead' funds we only minimise the survivorship bias and do not eliminate it.[7]

In order to determine the impact of the instant history bias in our data base, we have computed for each fund the difference between the monthly average of the excess return (over SP500) in the first year and the monthly average of the excess return in the first five years. The mean of the difference is 0.33%, equivalent to an annual difference of 3.97%. The standard deviation of the difference is 1.01%. The distribution of differences is positively skewed (with an estimated coefficient of skewness of 0.48), suggesting the existence of a small number of funds having much higher returns in the first year than in the rest of the reporting period. Separately, we have estimated the average monthly excess returns for the first 5 to 60 months from the moment when the fund started reporting. Figure 2 plots the average excess return across all funds. The decreasing pattern reveals that there is a clear 'first year' bias in the reported fund performance. In order to isolate the instant history bias we use dummy variables for the first year of reporting in all factor models.

To identify the most likely causes of funds ceasing to report, we have analysed the performance of the 'dead' funds prior to the moment of exiting the database. For each 'dead' fund we computed the difference between the monthly average excess return over SP500 in the last year of reporting and the monthly average excess return over the five years prior to ceasing reports. The

[7] To identify the impact of survivorship bias on our selected database, Figure 1 shows the number of funds entering the database (left hand scale) and the number of funds that ceased reporting over the previous 12 months (right hand scale) for the period Jan-90 to Dec-02. Since we require all funds to have at least 5 years of reporting our database does not include any fund entering after Jun-99. Until the middle of the 1990s, during the set-up of the database, the percentage of funds entering the database is huge, starting from as high as 70% in 1990. By 1996, the percentage of funds added to the database settles at around 20%. The evolution of the number of 'dead' funds is the opposite: the first fund ceases reporting in Dec-95 (there is no information of funds that ceased operating before the moment of the database set-up), and the percentage increases steadily to 5% in the last part of our sample.

mean of the difference is -0.50%, which is equivalent to an annual difference in the excess returns of -6.10%. The standard deviation of the difference is 1.65%. Figure 3 shows the distribution of these differences: it is positively skewed with a heavy upper tail, indicating that some funds stopped reporting because of negative performance but some also because of extraordinary good performance. A similar explanation is given by the analysis of the relationship between the overall performance as measured by the information ratio and the number of dead funds in each category.

As already mentioned, we require a minimum number of reporting months: the longer the minimum reporting period, the more accurate our factor model estimates for each fund, but also fewer funds in the sample. Longer minimum reporting periods may increase some of the biases. With no minimum number of reporting months, the database has more than 600 funds of which 25% had ceased reporting during the sample period. For a minimum of 60 months of reporting, the number of funds goes down to 282 of which 19% had ceased reporting: this suggests that the survivorship bias could increase when we impose a minimum number of reporting months. However, the estimated multi-period bias is negative (but negligible), at -.33% p.a.

In summary, the instant history bias is significant but this will be eliminated through the use of dummy variables. The multi-period sampling bias is negligible. However, there may be an unavoidable selection bias in our database and a survivorship bias that, though minimised by including in our analysis the funds which have ceased operating during our data sample, may be increased by the necessary imposition of a minimum reporting period. We therefore present all performance results on a relative basis, benchmarked against the equally weighted index of all funds. The relative performance can be interpreted as bias-free since both the portfolios and their benchmark are affected by the same biases.

## II Factor Models for Hedge Funds

The development of a parsimonious factor model that adequately explains hedge fund returns is a great challenge for alternative investment research. Such a model would allow one to measure risk adjusted performance and manager's skill, identify the style mix employed by funds and eventually devise optimal portfolios. Assuming that investors are only willing to reward managers for superior performance that cannot be easily replicated, the fund returns may be decomposed into the part explained by the factor model, which can be replicated by standard asset baskets and common trading strategies, and the factor model residuals being attributed to the fund manager's skill.

A number of factor models have been proposed to capture risk adjusted hedge fund performance: Ackermann, McEnally and Ravenscraft (1999), Brown, Goetzmann and Ibbotson (1999), Agrawal and Naik (2000b), and Liang (2000) among others have used one factor models, while Fung and Hsieh (1997), Schneeweis and Spurgin (1998), Liang (2001), Agarwal and Naik (2000), and Edwards and Caglayan (2001) have used fundamental or statistical multi-factor models. The wide range of models developed and the fact that no single model dominates the others is usually explained by the large diversity of the strategies employed by hedge funds and their highly dynamic nature. Whilst the traditional asset pricing literature is based on linear factor models, the highly dynamic derivatives strategies that are often levered in hedge fund management have non-linear relationships with traditional asset class returns (Fung and Hsieh, 1997; Agarwal and Naik, 2003; Amin and Kat, 2001).[8] Consequently, additional factors have been suggested for explaining hedge fund returns, such as dynamic derivative strategies payoff (Schneeweis and Spurgin, 2000; Agarwal and Naik, 2003; Mitchell and Pulvino, 2001; Fung and Hsieh, 2001). Hedge fund indices have also been used as alternative benchmarks for measuring performance (Lhabitant, 2001).

Since there is no consensus on the best model[9] we estimate four factor models on our database of 282 US hedge funds:

(1) a two-index model that considers the two main asset classes, US equities and bonds: this is the simplest possible representation of risk factors and is the base case model for our analysis;
(2) a broad fundamental model, including as factors: international equity and bond indices representing US and worldwide markets; investment style factors; commodities and foreign exchange risk factors, and other factors representing specific types of non-linear strategies such as market timing, volatility trading and equilibrium trading;
(3) a multi-factor model using the HFR hedge funds indices as factors;
(4) a statistical factor model using as factors portfolios replicating the first four principal components of the system of all funds' returns.

[8] For pricing securities whose payoffs are non-linear functions of the risk factors two strands of literature have emerged: non-linear factor models (Bansal and Viswanathan, 1993; Bansal, Hsieh and Viswanathan, 1993) and the use of derivatives strategies and other non-linear factors to capture the non-linearities in securities payoff (Breeden and Litzberger, 1978; Glosten and Jagannathan, 1994; Harvey and Siddique, 2000).

[9] As is the case for mutual funds, with Sharpe's (1992) model.

The general factor model representation is:

$$r_{it} = \alpha_i + \sum_{j=1}^{k} \beta_{ik} F_{kt} + \varepsilon_{it}$$

where $r_{it}$ is the net of fees excess return on fund *i* during month *t*; $\alpha_i$ is the risk adjusted performance, 'alpha', of fund *i* over the estimation sample; $F_{it}$ is the excess return on the $k^{th}$ risk factor over the month *t*;[10] $\beta_{ik}$ is the loading of the fund *i* on $k^{th}$ factor, i.e. the sensitivity of the fund *i* to the factor *k* over the estimation sample; and $\varepsilon_{it}$ is the error term.

Each of the four models were estimated by least square regressions over the period Jan-90 to May-03 on each of the 282 fund returns in excess of the risk free rate.[11] For each fund we used the entire data sample available: some funds entered the database after Jan-90 or ceased reporting before May-03. To account for the back-filling bias previously identified, all models included a dummy variable taking a value of one for the first twelve months of reporting. Potential multicollinearity problems were addressed by identifying any pairs of factors that were highly correlated over the sample and dropping the factor having lower correlation with the returns of the fund.[12] To select the significant factors for each fund a backward step-wise regression method was applied: starting with the most complete model which passed the multicollinearity filter, one-by-one the non-significant factors were removed until a parsimonious model was obtained.

The results of estimating the four factor models in this way are grouped for presentation, combining results for all funds within the same HFR strategy (strategy definitions are given in Appendix 1).[13] Each cell in Tables 1 to 4 reports two figures: the average coefficient estimate

[10] With a few exceptions when the risk factors are not investable indices or portfolios (e.g. for volatility and price dispersion risk factors)

[11] The 3 months US T-bill rate was used as a proxi for the risk free rate. We also attempted to use unsmoothed fund returns (Geltner, 1991 and 1993), given some evidence of autocorrelation in returns for approximately 30% of the funds in our database. However, such unsmoothed returns cannot be used in the second stage of our analysis, portfolio construction, since in back-tests we require returns as close to real market circumstances as possible. For coherency purposes, the factor models need to be estimated on the same type of returns which will be further used to construct portfolios in back-tests.

[12] The 'rule-of-thumb' used was if the pair-wise factor correlation coefficient was above the square root of the coefficient of determination of the model including all factors, then one of the factors was excluded. However, for some funds the model's coefficient of determination was low so we applied this rule only for correlation coefficients exceeding 0.5. Despite its wide-spread use this procedure is rather ad-hoc. A more sound approach would be to orthogonalise the factors and preserve the explanatory power of the model. However, the orthogonalisation of highly correlated explanatory variables has no impact on measuring alpha and its standard error, and we prefer the first approach for its intuitive interpretation and straightforward hypothesis testing.

[13] For reporting purposes, we have regrouped the HFR strategies that were weakly represented in our selected database (i.e. with less than 4 funds from that strategy type). The three funds from 'regulation D' and the only fund from 'relative value' were put together with the convertible arbitrage funds. Also, the four 'macro' funds and the two 'short selling' funds were grouped with 'market timing' funds.

over all funds in that strategy type (above) and the percentage of these funds for which the coefficient was statistically significant at 10% (below).[14] The coefficient standard errors were computed using the Newey-West (1987) heteroscedasticity and autocorrelation consistent covariance matrix.

*Two-index model*

This is the simplest representation of the fund excess returns as a function of the excess returns on the two most important underlying asset classes, equities and bonds. We are fully aware of its inappropriateness and use it merely as a base case model. The indices used to proxy the equity and bond markets were the Wilshire 5000 and the Lehman Government/Credit Intermediate indices. Given evidence of autocorrelation for approximately 30% of the funds in our database, which could be caused by their pricing practices for illiquid securities, there is a non-synchronous trading measurement risk for beta. Asness, Krail and Liew (2001) use contemporaneous and lagged market betas to show that hedge funds may have more market exposure than one expects due to stale prices or illiquidity of the securities they trade. Following Dimson's (1979) arguments, we included the lagged equity index excess returns in the factor model to account for potentially stale prices.

The results are summarised in Table 1. On average, the two-index model explains only 27% of the total variance of fund excess returns. This is in line with previous results (e.g. Fung and Hsieh, 1997) and can be attributed to the diverse dynamic strategies employed in the alternative investment industry which induce non-linear exposures to traditional asset classes. Still, 80% of funds have a significant relationship with the Wilshire 5000 excess returns (average beta = 0.3) and for 38% of funds the lagged Wilshire excess returns are also significant determinants. However, the bond index returns are significant for only 20% of funds. The average alpha is positive and significant for 48% of funds, and negative and significant for only three funds. The distribution of alphas has a mean of 7% and is both positively skewed and leptokurtic.

*Broad fundamental factor model*

The broad fundamental factor model uses excess returns on indices to capture the performance of the main traditional asset classes, and other factors to model specific types of strategies, such as

[14] The alphas on individual funds are not independent so to assess the significance of an average alpha over all the funds in a particular strategy type single-sample mean t-tests are not appropriate. Following Amenc and Martellini (2003a) we estimated the factor models on an 'average fund' returns series for each strategy type. We then used the least squares standard error of the intercept to determine the statistical significance of the average alpha.

market timing, volatility trading and equilibrium based trading models. Following Sharpe (1992), Schneeweis and Spurgin (1998), Agarwal and Naik (2003) and others, we include in the broad fundamental factor model *equity* indices (Wilshire 5000, SP500 growth and value, SP mid-cap and small-cap to capture differences in equity investment styles, MSCI world index excluding US to account for the investment opportunities outside US and MSCI emerging markets index to capture the emerging markets investment opportunities as a separate asset class); *bond* indices (Lehman Government, Lehman Credit Bond, Lehman High Yield and Lehman Mortgage Backed Securities); the FED trade weighted *foreign exchange* rate index as a proxy for foreign exchange risk; the GS Commodity index to capture *commodity* related investment risk factors. In the spirit of Sharpe's (1992) model, the slope coefficients of the regression will indicate the replicating static mix of asset classes that would capture the fund's performance. It is common practice to go beyond static asset class mixes and analyse the performance of funds using simple trading strategies. As suggested by Treynor and Mazuy (1966), squared market returns can proxy for market timing abilities. Therefore, in order to account for potential asymmetries in the relationship of the fund returns with the main asset classes, which are expected to occur when derivatives or dynamic timing strategies are used, we include in the regressions the squared excess returns of the main indices. Additionally, we include two factors capturing specific trading strategies: the change in the equity implied *volatility* index (VIX) to account for volatility trades (Schneeweis and Spurgin, 1998),[15] and the prices' *dispersion* as a leading indicator of price equilibrium trading strategies (Alexander and Dimitriu, 2003).[16]

The estimation results are presented in Table 2. Although 17 factors were considered in total, the average number of significant factors for an individual fund was only 2.5. Nevertheless the average $R^2$ across all funds was 36%, a considerable increase from the base case model. The broad fundamental model better explained the returns of funds in the following classes: emerging markets, equity hedge and non-hedge, event driven, convertible bonds, financial and technology sectors. However, the returns for some fixed income, macro and relative value funds were not well modelled by this approach. The most significant factor, determining the excess returns of

[15] A quick note on the importance of using the right measure of volatility. If funds are indeed using derivatives and timing strategies, a positive relationship with volatility should be revealed. However, by looking at several other volatility measures, e.g. intra-month SP volatility, or maximum draw-down, one would get the opposite intuition – that is, there is a negative relationship volatility – fund returns, which would cast serious doubt on the effectiveness of the sophisticated strategies used by funds.

[16] Alexander and Dimitriu (2003) argue that prices' cross sectional dispersion represents a measure of prices equilibrium. Strategies based on historical equilibrium will realise relative losses when the dispersion of prices is increasing and relative gains when the dispersion of prices decreases. However, in special market circumstances, e.g. following a crash period, the relationship changes sign – relative gains are realised when prices dispersion increases.

38% of the funds is the *small cap SP index*. As expected, this influences funds trading on distressed securities, equity hedge and non-hedge, event driven, funds of funds and technology. Additionally, the squared returns of the small cap SP index are significant in 40% of models, indicating use of leverage and market timing abilities.[17] The other equity style indices are only significant in 6-10% of the funds. Another important factor is the Lehman High Yield index (for convertible arbitrage, distressed securities, event driven, managed futures, merger arbitrage and funds of funds, as well as equity hedge and non-hedge funds, the latter two having a negative average beta) and its squared excess returns (highly significant and positive for funds in distressed securities, equity non-hedge and managed futures, and significant but negative for convertible arbitrage, equity hedge, funds of funds, market timing and technology funds). The MSCI emerging markets index was significant for 20% of the funds: in addition to the funds primarily trading in emerging markets and funds of funds investing in these, distressed securities, equity hedge and non-hedge, event driven, and technology funds all have significant exposure to emerging markets. The GS commodity index is a significant factor for emerging markets, event driven, funds of funds, equity hedge and non-hedge, and, as expected, managed futures.

Whilst the broad equity market indices and the FED trade weighted forex index are generally less significant than other factors, the change in SP500 implied volatility and the Dow Jones price dispersion index are among the most significant factors. Each of these factors has positive coefficients and is significant for almost 30% of funds. The first year reporting dummy confirmed a significant positive bias for most strategies, especially convertible arbitrage, equity market neutral, event driven, funds of funds, market timing and managed futures. However, there is a negative first year of reporting dummy coefficient for funds in the equity non-hedge, emerging markets and technology categories. Alpha was most significant for emerging markets and financial sector funds (significant negative alpha only occurred for funds investing in emerging Asian markets) and also for convertible arbitrage, relative value and short-selling (but for these a high alpha could just result from poor explanatory power of the model).

*Hedge fund indexes model*

The hedge fund indices provided by HFR represent investable portfolios having non-linear exposures to traditional asset classes. Hence a factor model based on them should explain the

---

Considering the wide spread use of historical equilibrium models we use dispersion as an indicator of their performance.

[17] The squared returns of small caps have a positive relationship with funds from the equity hedge, market neutral and non-hedge, funds of funds, managed futures, market timing and technology and a negative one with the funds trading on distressed securities and event driven.

returns on individual funds better than the two factor model considered previously (Lhabitant, 2001). The HFR indices are based on clustering analysis and possible re-classification of funds if their self-stated strategy is not consistent with HFR's statistical analysis. One approach is to classify funds into groups and use only the relevant group index to explain their returns. Instead we allow, through step-wise regression, the selection of the most significant hedge fund indices for each fund. Also, we do not reconstruct the indices from the funds in our sub-sample: the HFR indices are a better choice as factors since they are more representative for the entire population of funds and are also investable.

The results of estimating this factor model are presented in Table 3. Clearly there are systematic risk factors, beyond the ones included in the fundamental factor model that are captured by these HFR style indices.[18] The model explains an overall average of 46% of the variance in fund excess returns (ranging from 37% for equity market neutral funds, to over 60% for equity non-hedge funds, event driven, funds of funds and technology funds). Each fund's excess returns tend to be determined by the relevant index for their self-stated strategy, indicating no large errors in the classification. At an individual fund level, 17% of funds have negative and significant alpha, while only 11% of funds have positive and significant alpha. The funds with positive alpha come from the emerging markets group, some fixed income strategies, market timing and managed futures. The funds with the largest negative alphas are equity non-hedge, event driven and some sectors.

*Statistical model*

Statistical factors extracted from the funds' returns data through principal component analysis were first used to model hedge fund returns by Fung and Hsieh (1997). As opposed to mutual funds, for which the asset class they trade in is the determinant factor, the more important factor for hedge funds is 'how' they trade, i.e. which type of dynamic strategy they employ for a particular asset class. The intuition behind statistical factor analysis is that if a group of funds use similar strategies in the same markets, their returns should be correlated. Through factor analysis, the major common styles can be extracted from fund returns.

Appendix 2 describes the methodology used to identify the optimal number of principal component factors and to construct investable portfolios replicating these factors. Our statistical factor model uses four principal component factors, and the portfolios replicating these are denoted by PC1 to PC4. These portfolios preserve well the theoretical features of the principal components: PC1 captures the common trend in fund returns and has a very similar behaviour to an equally-weighted index of all funds, to which it is also highly correlated.[19] It has negative but not very significant skewness, and an excess kurtosis of 1.4. The other three portfolios have returns distributions that are nearer to normality.[20]

From their portfolio structure, we can infer that the first two principal components have an intuitive interpretation: the 'common trend' PC1 portfolio is well diversified across all strategies (with particular emphasis on funds of funds and equity funds because most funds are of these types) and the PC2 portfolio is clearly dominated by managed futures, which stand out as an investment style with returns uncorrelated to the common trend but representative of a significant part of the hedge fund population. PC3 comprises mainly equity market neutral and funds of funds, while in PC4 there are technology funds and again equity market neutral funds.[21] In summary, the structure of the portfolios replicating the first two principal components is well defined, whilst the third and fourth principal components portfolios have a mixed composition providing little insight on their characteristics.

The relationship of the first principal components with the traditional asset classes has been investigated previously, both from a linear and non-linear perspective (Fung and Hsieh, 1997). Their results indicate that, if the first principal component is explained to a fair degree by a linear representation of the traditional asset classes, this is not the case for the other principal components, which have non-linear relationships with the traditional asset classes. Indeed, we find that the first principal component has strong linear relationships with our broad fundamental factors ($R^2$ = 0.79) and hedge fund indices ($R^2$ = 0.89), while the other three principal components have no obvious significant linear relationship with fundamental factors (maximum $R^2$ = 0.20), and a rather weak one with the hedge fund indices (maximum $R^2$ = 0.35).[22] Appendix 3 provides

[18] Following Fung and Hsieh's (1997) argument, the hedge fund indices can be interpreted as style factors, as opposed to the location factors which are captured by the fundamental factors.
[19] The equally-weighted index is less volatile, but this is to be expected, considering that the number of funds entering the portfolio replicating the first principal component is smaller, and the first principal component is constructed by maximising the variance of the underlying linear combination of funds.
[20] The portfolios replicating the other three principal components also have progressively lower volatility, by construction, and generally higher information ratios than PC1.
[21] The presence of equity market neutral funds in three of the four orthogonal factors indicates that the funds in this category are rather heterogeneous, having low correlation with each other.
[22] This low dependency of higher order principal components to traditional asset classes is thought to be caused by the use of dynamic trading strategies and derivatives, as well as style switching. A well-known example is the strategy of a market timer with perfect forecasting abilities, which will display non-significant correlation with the market, but this is just an artefact of the correlation computation method. In fact, the strategy is perfectly correlated with up-markets and

clear evidence that the portfolios replicating the higher principal components are in fact capturing dynamic hedge fund strategies having non-linear relationships with the traditional asset classes.

The estimation results for the statistical factor model are presented in Table 4. The only strategies with positive and significant alphas are convertible arbitrage and merger arbitrage, but they also have a low average $R^2$, so the abnormal return could be due to omitted risk factors. The average $R^2$ across all strategies is 0.39, greater than for the fundamental factor model, suggesting that there is indeed value in using statistical factors capturing dynamic hedge fund strategies. The strategies with the highest $R^2$ are the equity hedge and non-hedge, funds of funds and technology funds, but this is clearly an artefact of the structure of our database, which is dominated by these types of funds. The PC1 portfolio is a significant factor for 79% of funds, PC2 portfolio for 39% funds, PC3 for 44% funds and PC4 is significant for 29% of funds. All strategies, except for managed futures have positive average betas on the PC1 portfolio. Also, most strategies have positive betas on the PC2 portfolio and negative betas on PC3 portfolio. The first year reporting dummy remains positive and significant for most strategies, except for event driven, merger arbitrage and technology funds.

*A Comparison of Alphas from Different Factor Models*

At the strategy level, substantial differences in alphas estimated from the four factors models have been identified in Tables 1 – 4. Table 5 summarises the alpha estimates of the 'average fund'. For the two index model an estimated alpha of 0.69% per month is highly significant and is equivalent to 8.5% per annum. The risk adjusted performance for the 'average fund' is again highly significant in the broad fundamental model: at 5.1% in annual terms this model still implies a positive average alpha, even if significantly reduced from the base-case two index model. However, for the multi-factor model, the 'average fund' risk adjusted performance is negative, even if not statistically significant. More importantly, the 'average fund' has a negative and significant alpha when benchmarked against the four PC portfolios.

In summary, there is a significant disagreement between the factor models on the 'average fund' alpha, ranging from –2.5% to 7% per annum, and the dispersion of alpha estimates is even higher at the level of individual funds. This leads us to conclude that an accurate estimate of the abnormal return is very difficult to obtain from any single model with any degree of certitude. However, some agreement can be achieved on the sign of alpha for individual funds. Of the funds

---

perfectly negatively correlated with down-markets. In the real world, there are of course no perfect market timers, having at least one positive alpha estimate, in 30% of cases there is perfect agreement between all models in terms of alpha's sign.

After investigating an exhaustive range of factor models, Amenc and Martellini (2003a) conclude that the models tend to agree on the relative ranking of funds despite the fact that the range of alpha estimates produced by different models is wide. We also find significant agreement between different models in terms of funds' ranking based on alpha. Table 6 shows the correlation, the rank correlation and probability of agreement between different models in terms of alpha estimates. All probabilities of agreement are above 0.58, even though the correlation in individual alpha's estimates can be as low as 0.27 (for the statistical model and the base case model).

The implication of these results for funds of funds portfolio construction is that methods relying heavily on the accuracy of alpha estimates cannot be implemented without significant model risk. However, there is considerable scope for more flexible models, such as implied alpha models, based precisely on funds ranking. In the following, we will examine the power of such portfolio construction models for hedge funds.

## III Constructing Optimal Portfolios of Hedge Funds

Optimisers are well known to be error enhancers (Michaud, 1989), so the quality of the data is essential. Since this is not likely to be achieved with hedge funds in the near future one can only aim for a better solution than naïve diversification. Before proceeding to optimisation, we consider the optimal size of diversified hedge funds portfolios by testing the benefits of naïve diversification through simulation. We randomly draw without replacement 5, 10, and up to 80 funds (in multiples of 5) from our database and independently form equally weighted portfolios having no style consideration. For each size of portfolio we repeat the experiment 1,000 times, and estimate the first four moments of the fund of funds portfolio returns distributions for each portfolio size. Our results, summarised in Figures 4 and 5 indicate that diversification across all strategies works well, with most of the diversification benefits obtained at around 30 funds in the portfolio.[23] Practitioner standards also appear to favour portfolios of at least 20 to 30 funds even though other academic results on funds diversification (Henker, 1998;

which makes things even more complicated, because one has to consider also the effect of forecasting/trading errors.

[23] Our results are based on a smaller sample (282 funds) than other studies, which does not allow the investigation of the diversification benefits at the level of each style. Lhabitant and Learned (2002), on a database of 6,985 funds, find a significant trade off between the decrease in volatility and the evolution of the higher moments: as the number of

Amin and Kat, 2002) recommend a number of funds in the range of 15 to 20. Therefore, provided that holding 20 to 30 funds does allow a fair diversification across all investment styles in our database, our portfolios will include a number of funds in this range.

The classical portfolio theory of Markowitz (1952) underlies the foundations of modern finance and many of today's practitioner models. Assuming investors have quadratic preferences, its application requires knowledge of the first two moments of the returns distribution. Estimating expected returns has been shown to be a difficult task even for traditional asset classes (Merton, 1980; Jorion, 1985), and this is the main reason why mean-variance efficient portfolios perform poorly out-of-sample (Frost and Savarino 1986, 1988; Jorion, 1986; Michaud, 1989; Best and Grauer, 1991). Accurate estimates of expected returns are even more difficult to obtain for alternative investments, as shown in the previous section. On the mean-variance efficient frontier the only portfolio that does not require estimates of expected returns is the minimum variance portfolio. In the hedge funds world, this is a natural choice, given that it eliminates the estimation risk associated with expected returns. Moreover, there is empirical evidence (Jorion, 1985; Jagannathan and Ma, 2003) that out-of-sample, the minimum variance portfolio overperforms classical tangent portfolios in terms of information ratio, especially when the estimation sample is not large.[24] For these reasons, we focus our analysis on minimum variance portfolios of hedge funds.

When no restrictions are imposed, the minimum variance portfolio weights are given by:

$$W_{MV} = \frac{\Sigma^{-1}\underline{1}}{\underline{1}'\Sigma^{-1}\underline{1}}$$

where $\Sigma$ is the covariance matrix of the fund returns and $\underline{1}$ is a vector of ones. If restrictions are imposed on weights, the solution can be obtained numerically. Since the covariance matrix of fund returns is the only input required in the model, the results will strongly depend on its accuracy. Given that the number of funds in our sample is much larger than the number of data points, we are concerned with the estimation risk in the sample covariance matrix. Appendix 2

funds is increased, for strategies like fixed income arbitrage, convertible arbitrage and event driven, the skewness decreases and the excess kurtosis increases.

[24] For comparison, we also implemented a mean-variance maximum information ratio optimisation in place of a minimum variance optimisation but the lack of accuracy in the individual alpha estimates resulted in lower out-of-sample information ratios for the mean-variance portfolios than for the minimum variance portfolios, and the portfolios were less stable than the ones constructed based on minimum variance. Therefore, rather than optimising on less accurate estimates, one is better off selecting funds based on alphas and then running the optimisation on the covariance matrix solely.

compared the properties of the empirical correlation matrix[25] with those of a similar but random correlation matrix, finding significant agreement between the two. Hence part of the information content of the empirical correlation matrix is driven by randomness. In Appendix 2 we showed that the 'true' correlation structure is captured by the first four eigenvalues, while the rest can be ascribed to noise and measurement errors. Since the presence of such noise is likely to perturb the minimum variance optimisation, its reduction is essential. Several solutions have been offered, including imposing a factor structure (e.g. Sharpe, 1963; Chan, Karceski and Lakonishok, 1999), the use of an optimal shrinkage towards the mean or the single factor model (Jorion, 1985, 1986; Ledoit and Wolf, 2003; Kempf and Memmel, 2003) and the introduction of portfolio constraints (Jagannathan and Ma, 2003).

One straightforward solution that exploits the information contained in the empirical correlations whilst minimising the model risk has been recommended by Plerou *et al.* (2002) for traditional assets, and Amenc and Martellini (2002, 2003b) for alternative investments. It involves the reconstruction of the empirical correlation matrix using only the eigenvalues that deviate from those of a random correlation matrix. In our database only four eigenvalues deviated from those of a random correlation matrix, as shown in Appendix 2. We therefore construct the 'cleaned' correlation matrix as:

$$C = W\Lambda_c W'$$

where $\Lambda_c$ is the diagonal matrix of the ordered eigenvalues of the correlation matrix of fund returns with all but the first four eigenvalues replaced by zeros, and W is the matrix of eigenvectors. The diagonal elements of C were set equal to one. The 'cleaned' covariance matrix was computed as V = DCD, where D is the diagonal matrix of standard deviations of each fund returns.

In order to test the efficiency of the noise 'cleaning' process, we compared, out-of-sample, the variances of minimum variance portfolios constructed from (a) the empirical covariance matrix and (b) the 'cleaned' covariance matrix. We constructed 1,000 portfolios, each with 25 randomly selected funds, and optimised them based on both the 'cleaned' and the sample covariance matrices to achieve minimum variance. For estimating the covariance matrices we used a rolling sample of 60 months. The first portfolios were constructed in Jan-98 and rebalanced every six months until Jan-03. Between two rebalancing moments, the portfolios are left unmanaged and their out-of-sample performance monitored in order to determine the out-of-sample variance.

[25] We have performed the same analysis for the covariance matrix and obtained similar results.

The results are presented in Table 7. We found that, when no constraints are imposed on the portfolio weights, the out-of-sample variance of the portfolio constructed on the 'cleaned' covariance matrix is smaller than that based on the sample covariance matrix, indicating the effectiveness of the noise cleaning process. Our results are in line with those of Plerou *et al.* (2002), whose out-of-sample results also favoured the 'cleaned' covariance matrix. But in a hedge funds portfolio, no short sales are allowed, so the portfolio weights need to be constrained to be positive and upper bounds are also normally imposed to reduce concentration risk.[26] We therefore introduce a non-negativity constraint and an upper bound on individual fund weights at 20%, following standard practice in this respect.

When we repeated the previous analysis with the constraints in place, the out-of-sample results were in favour of the sample covariance matrix (Table 7). This feature has been previously observed by Jagannathan and Ma (2003).[27] The constraints improve significantly the performance of the portfolio based on the sample covariance matrix, which has the smallest variance overall and an average excess kurtosis that is smaller than that of the minimum variance portfolios based on the 'cleaned' covariance matrix. Hence in the following we use the ordinary sample covariance matrix instead of the 'cleaned' one.

The portfolios analysed so far were based on randomly selected funds, and give an idea of the average performance expected from minimum variance portfolios of hedge funds. Next, we enhance the minimum variance portfolios by introducing a fund selection criterion based on the alphas estimated with the four factor models presented in the previous section. If accurate estimates of alphas had been available, then a mean-variance optimisation would have been possible. But a large dispersion was associated with the alpha estimates from different factor models and the only significant agreement between the models' alphas was in terms of their

[26] The role of constraints on portfolio performance goes beyond this institutional limitation, being an alternative method for dealing with measurement errors. A recent paper by Jagannathan and Ma (2003) shows that imposing upper and lower bounds on the weights is equivalent to shrinking the covariance matrix towards zero. The sampling errors are reduced this way, improving the out-of-sample performance. Especially in large cross sections, even if the constraints do not hold in the population, they are still useful, because the sampling error is much larger than the specification error induced by imposing wrong restrictions. Additionally, since non-negativity constraints tend to concentrate the portfolio on few assets, imposing upper bounds on the portfolio weights ensures that the optimal portfolios comprise a sufficiently large number of stocks.

[27] They found that when imposing a factor structure on the covariance matrix, constraining the portfolio weights to be non-negative can result in a significant reduction of performance. This happens because the sampling error has already been reduced through imposing the factor structure, and adding constraints which do not hold in the population increases the specification error, which is penalised in terms of out-of-sample performance.

ranking. Hence we can only use alpha estimates as a selection criteria and not as a parameter in the optimisation.

Using each of the four factor models, we first select all the funds having positive alphas that are significant at 10%. Secondly, we tighten the selection criteria and require that all models indicate a positive alpha, which is also significant in at least three of them at the 10% significance level. In each case, the portfolio weights are allocated so that the fund of funds has minimum variance.

In order to test the out-of-sample performance of these two selection criteria we use the period Jan-90 to Dec-97 to calibrate the factor models and select the first set of funds. We keep the non-negativity and 20% upper bound constraints in place. The first minimum variance portfolios are set up in Jan-98 and left unmanaged for the next 6 months. The portfolios are then re-balanced every six months, based on the alphas estimated over the entire data sample available at the portfolio construction moment and the covariance matrices estimated over the previous 60 months.

Compared with the randomly selected minimum variance portfolios having a fixed number of 25 funds, the portfolios of hedge funds that employ the alpha selection criterion (each having a variable number of funds) have a significantly improved out-of-sample performance. The evolution of the portfolio values, equivalent to a dollar investment of 100 in Jan-98, is plotted in Figure 6 (for reference, we also plot the dollar value of an equally weighted portfolio of all funds in our database[28]) and the portfolio statistics are reported in Table 8. The results are very encouraging indeed: all portfolios have average annual returns in the range of 8% to 9.5%, with an annual volatility of only 1.3% to 1.7%. Their evolution is very constant, with no more than three months (out of 66) having negative returns in any of the models. The lowest volatility is displayed by the portfolio based on the alpha estimates of the statistical factor model and this portfolio also has the highest average annual information ratio (6.91). The highest returns are produced by the portfolio based on the HFR indices factor model, but given the higher volatility, the information ratio of this portfolio was the lowest (5.56). The null hypothesis of normally distributed out-of-sample returns cannot be rejected. By contrast, the equally weighted portfolio has comparable returns but much higher volatility resulting in an information ratio of 1.5, and its returns display significant excess kurtosis. We note that even the portfolio based on the alpha estimates from the base case model produces good results. The more conservative approach, based on all models alpha estimates generates average results: three of the portfolios based on single model alpha estimates produced better results in terms of both volatility and information ratio. To conclude, all factor models provided valuable information for hedge fund selection.

These results are reported before transactions costs. It is not easy to estimate the transaction costs associated with hedge funds portfolios, given the heterogeneity of sales and early redemption fees used in the hedge fund world as well as other associated costs. However, given the significant trading limitations for hedge funds portfolios it is very important to investigate the structural stability of these portfolios. Therefore we compute a measure of portfolio turnover: the absolute difference in fund weights from one rebalancing period to the next. With 10 rebalancing points in our out-of-sample test, the maximum turnover is 20 (10*200%). Table 9 reports that the estimated turnover for our portfolios ranges from 4.9 (for portfolio based on the statistical factor model, the same as for the equally weighted portfolio of all funds) to 7.6 (for the portfolio based on the HFR indices model). This is equivalent to restructuring from 12% (turnover = 4.9) to 19% (turnover = 7.6) of the portfolio at each rebalancing point.

For comparison, we also implement a maximum information ratio optimisation based on the alpha estimates from all four models and the sample covariance matrix, all the rest being the same. The results are presented in Table 9. As expected, the lack of accuracy in the individual alpha estimates results in lower out-of-sample information ratios for the mean-variance portfolios as compared to the minimum variance portfolios, even if the objective of the optimisation in the first case is exactly the information ratio. Additionally, the mean-variance portfolios are less stable than the ones constructed based on minimum variance. Therefore, rather than optimising on less accurate estimates, one is better off selecting funds based on alphas and then running the optimisation on the covariance matrix solely.

In summary, all performance measures favour the portfolio where hedge funds are selected using the alphas from the statistical factor model: this produces a portfolio of hedge funds with the highest average annual information ratio (6.91), the lowest turnover (12% portfolio restructured every 6 months), and its returns are very close to normality. We have shown that a hedge fund selection method based on alpha estimates from any factor model greatly improves the performance of hedge fund portfolios compared to both randomly selected portfolios and equally weighted portfolio of all funds. Also, the minimum variance optimisation produces better results

[28] A value weighted portfolio of all funds would be a better reflection of the performance of hedge funds investment,

than maximum information ratio optimisation. In the constrained portfolio optimisation, there is no benefit from imposing a factor structure on the correlation matrix, as the sampling errors appear to be significantly reduced through the weighting constraints.

## IV Summary and conclusions

Despite the modelling complexity caused by biases present in data, noisy correlation structure, alphas that are difficult to estimate and the institutional limitations to trading, there is no doubt that alternative investments present attractive opportunities, which explains their increasing popularity amongst institutional investors. Consequently, portfolio optimisation within a large universe of hedge funds has become a key area for research in portfolio management.

The purpose of this paper was to develop a fund selection and optimal allocation process for funds of hedge funds. To this end, we have analysed the out-of-sample performance of a portfolio construction model that is designed to target alpha through fund selection based on factor models. In order to deal with the data biases, we have used a database comprising both dead and alive funds and reported performance on a relative basis, benchmarked against portfolios affected by the same biases.

Since most traditional portfolio construction models require an estimate of expected returns, we have used a number of factor models to estimate the funds relative abnormal return. Despite the significant disagreement in terms of individual alpha estimates, the factor models largely agreed on the ranking of funds. Therefore, whilst methods relying heavily on the accuracy of alpha estimates cannot be implemented for hedge funds given the significant model risk, there is considerable scope for more flexible models that are based only on funds' ranking. Having obtained alphas estimates from the four factor models, we used them only to select funds.

The best portfolio optimisation, as a separate stage, was based solely on the covariance matrix. Finding that the sample covariance matrix had a high element of randomness, we considered both 'cleaning' the matrix (by imposing some factor structure before using it in the portfolio optimisation) and reducing the measurement errors by imposing weight constraints in the portfolio optimisation model. We found no benefit from imposing a factor structure on the

but assets under management series are incomplete for some of the funds in our database.

correlation matrix, as the sampling errors already appear to be significantly reduced through the weights constraining.

Our results showed that the fund selection method based on factor models alpha estimates greatly improves the performance of hedge fund portfolios optimised to have minimum variance, compared to both randomly selected portfolios and the equally weighted portfolio of all funds. Also, given the inaccuracy in alpha estimates, the minimum variance optimisation produced better results than maximum information ratio optimisation, when applied to the funds selected based on the factor model alpha estimates.

An out-of-sample performance assessment of this class of models has revealed considerable alpha opportunities in the alternative investment world, which can be exploited by simple constrained optimisation models. Thus, with some refinements, the traditional tools of portfolio management can be applied to portfolios of hedge funds to achieve excellent results.

## References:


Ackermann, C., R. McEnally and D. Ravenscraft (1999) "The performance of Hedge Funds: Risk, Return and Incentives", Journal of Finance 54, p. 833-874

Agarwal V. and N. Naik (2000) "On taking the Alternative Route: Risks, Rewards and Performance Persistence of Hedge Funds", Journal of Alternative Investments 2-4, p. 6-23

Agarwal V. and N. Naik (2003) "Risk and Portfolio Decisions Involving Hedge Funds", forthcoming Review of Financial Studies

Amenc, N., and L. Martellini (2002) "Portfolio Optimization and Hedge Fund Style Allocation Decisions", Journal of Alternative Investments 5-2, p. 7-20

Amenc, N., and L. Martellini (2003a) "The Alpha and Omega of Hedge Fund Performance Measurement", discussion paper, EDHEC Business School, www.edhec-risk.com

Amenc, N., and L. Martellini (2003b) "Optimal Mixing of Hedge Funds with Traditional Investment Vehicles", discussion paper, EDHEC Business School, www.edhec-risk.com

Amin, G. and H. Kat (2002) "Portfolios of Hedge Funds: What Investors Really Invest In" ISMA Centre Discussion Paper 07/2002

Amin, G. and H. Kat (2003) "Hedge Fund Performance 1990-2000: Do the "Money Machines" Really Add Value?" Journal of Financial and Quantitative Analysis 38-2, p. 251-274

Asness, C., R. Krail and J. Liew (2001) "Do Hedge Funds Hedge?" The Journal of Portfolio Management 28, p. 6-19

Bansal, R. and S. Viswanathan (1993) "No Arbitrage and Arbitrage Pricing: A New Approach", Journal of Finance 48, p. 1231-1262

Bansal, R., D. Hsieh and S. Viswanathan (1993) "A New Approach to International Arbitrage Pricing", Journal of Finance 48, p. 1719-1747

Best, M., and R. Grauer (1991) "On the Sensitivity of Mean-variance-efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results", Review of Financial Studies 4, p. 315-342

Breeden, D. and R. Litzenberger (1978) "Prices of State Contingent Claims Implicit in Option Prices", Journal of Business 51, p. 621-651

Brown, S., W. Goetzmann and R. Ibbotson (1999) "Offshore Hedge Funds: Survival and Performance 1989-1995", Journal of Business 72, p. 91-117

Chan, L.K.C, J. Karceski and J. Lakonishok (1999) "On Portfolio Optimization: Forecasting Covariances and Choosing the Risk Model", Review of Financial Studies 12, p. 937-974

Dimson, E. (1979) "Risk measurement when shares are subject to infrequent trading", Journal of Financial Economics 1979, 7(2), p. 197-226

Edwards, F.R. and M.O. Caglayan (2001) "Hedge Fund Performance and Manager Skill", Journal of Futures Markets 21-11, p. 1003–1028

Frost, P., and J. Savarino (1986) "An Empirical Bayes Approach to Efficient Portfolio Selection", Journal of Financial and Quantitative Analysis 31, p. 293-305

Frost, P., and J. Savarino (1988) "For Better Performance: Constrain Portfolio Weights", Journal of Portfolio Management 15, p. 29-34

Fung, W. and D. Hsieh (1997) "Empirical Characteristics of Hedge Fund Strategies", Review of Financial Studies 10, p. 425-461

Fung, W. and D. Hsieh (2000) "Performance Characteristics of Hedge Funds and Commodity Funds: Natural vs. Spurious Biases", Journal of Financial and Quantitative Analysis 35, p. 291-307

Fung, W. and D. Hsieh (2001) “The Risk in Hedge Fund Strategies: Theory and Evidence from Trend Followers”, Review of Financial Studies 14, p. 313-341

Fung, W. and D. Hsieh (2002) “Benchmarks of Hedge Fund Performance: Information Content and Measurement Biases”, Financial Analyst Journal 58-1, p. 22-34

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

Geltner, D. (1993) “Estimating Market Values from Appraised Values without Assuming and Efficient Market”, Journal of Real Estate Research, Vol. 8, p. 325-345

Glosten, L., and R. Jagannathan (1994) “A Contingent Claims Framework for Analyzing the Performance of Portfolio Managers” Journal of Empirical Finance 1, p. 133-160

Hamilton, J.D., (1994) “Modeling Time Series with Changes in Regime”, in Time Series Analysis Princeton, Princeton University Press, p. 677-703

Harvey, C., and A. Siddique (2000) “Conditional Skewness in Asset Pricing Tests”, Journal of Finance 55, p. 1263-1295

Henker, T. (1998) ‘Naïve Diversification for Hedge Funds’, Journal of Alternative Investments Winter 1998, Volume 1, Number 3, 33-38

Jagannathan, R., and T. Ma (2003) “Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps”, Journal of Finance 58, p. 1651-1683

Jorion, P. (1985) “International Portfolio Diversification with Estimation Risk”, Journal of Business 58, p. 259-278

Jorion, P. (1986) “Bayes-Stein Estimation for Portfolio Analysis”, Journal of Financial and Quantitative Analysis 21-3, p. 279-292

Kempf, A. and C. Memmel (2003) “On the Estimation of the Global Minimum Variance Portfolio”, discussion paper, Social Sciences Research Network

Lhabitant, F. (2001) “Hedge Funds Investing: A Quantitative Look inside the Black Box”, The Journal of Financial Transformation, Vol. 1(1), p. 82-90

Lhabitant, F. S. and M. Learned (2002) ‘Hedge Fund Diversification: How Much is Enough?’ Research paper 52, FAME – International Center for Financial Asset Management and Engineering

Ledoit, O., and M. Wolf (2003) “Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection” Journal of Empirical Finance 10, p. 603-621

Liang, B. (2000) “Hedge Funds: the Living and the Dead”, Journal of Financial and Quantitative Analysis 35, p. 309-326

Liang, B. (2001) “Hedge Fund Performance: 1990-1999”, Financial Analysts Journal 57, p. 11-18

Markowitz, H. (1952) “Portfolio Selection”, Journal of Finance 7, p. 77-91

Merton, R. (1980) “On Estimating the Expected Return on the Market: An Exploratory Investigation”, Journal of Financial Economics 8, p. 323-361

Michaud, R. (1989) “The Markowitz Optimization Enigma: Is “Optimized” Optimal?” Financial Analysts Journal 45, p. 31-42

Mitchell, M., and T. Pulvino (2001) “Characteristics of Risk and Return in Risk Arbitrage”, The Journal of Finance 56, no. 6, p. 2135–2175

Newey, W., and K. West (1987) “A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix”, Econometrica 55, p. 703-708

Plerou, V., P. Gopikrishnan, B. Rosenow, N. Amaral, T. Guhr and E. Stanley (2002) “Random Matrix Approach to Cross Correlations in Financial Data”, Physical Review E, Vol. 65, p. 1-18

Schneeweis, T. and R. Spurgin (1998) “Multifactor Analysis of Hedge Fund, Managed Futures, and Mutual Fund Return and Risk Characteristics” Journal of Alternative Investments Fall 1998, Volume 1, Number 2, p. 1-24

Schneeweis, T. and R. Spurgin (2000) “The Benefits of Index Option-Based Strategies for Institutional Portfolios” Journal of Alternative Investments, Spring 2001, p. 44 – 52

Sharpe, W. (1963) “A Simplified Model for Portfolio Analysis”, Management Science 9, p. 277-293

Sharpe, W. (1992) “Asset Allocation: Management Style and Performance Measurement,” The Journal of Portfolio Management, Winter, p. 7-19

Treynor , J. and K. Mazuy (1966) “Can Mutual Funds Outguess the Market?” Harvard Business Review 44, p. 131-136

## Appendix1

### HFR Strategy Definitions

**Convertible Arbitrage** involves purchasing a portfolio of convertible securities, generally convertible bonds, and hedging a portion of the equity risk by selling short the underlying common stock. Certain managers may also seek to hedge interest rate exposure under some circumstances. Most managers employ some degree of leverage, ranging from zero to 6:1. The equity hedge ratio may range from 30 to 100 percent. The average grade of bond in a typical portfolio is BB-, with individual ratings ranging from AA to CCC. However, as the default risk of the company is hedged by shorting the underlying common stock, the risk is considerably better than the rating of the unhedged bond indicates.

**Distressed Securities** strategies invest in, and may sell short, the securities of companies where the security's price has been, or is expected to be, affected by a distressed situation. This may involve reorganizations, bankruptcies, distressed sales and other corporate restructurings. Depending on the manager's style, investments may be made in bank debt, corporate debt, trade claims, common stock, preferred stock and warrants. Strategies may be sub-categorized as "high-yield" or "orphan equities." Leverage may be used by some managers. Fund managers may run a market hedge using S&P put options or put options spreads.

**Emerging Markets** funds invest in securities of companies or the sovereign debt of developing or "emerging" countries. Investments are primarily long. "Emerging Markets" include countries in Latin America, Eastern Europe, the former Soviet Union, Africa and parts of Asia. Emerging Markets - Global funds will shift their weightings among these regions according to market conditions and manager perspectives. In addition, some managers invest solely in individual regions. Emerging Markets - Asia involves investing in the emerging markets of Asia. Emerging Markets - Eastern Europe/CIS funds concentrate their investment activities in the nations of Eastern Europe and the CIS (the former Soviet Union). Emerging Markets - Latin America is a strategy that entails investing throughout Central and South America.

**Equity Hedge** investing consists of a core holding of long equities hedged at all times with short sales of stocks and/or stock index options. Some managers maintain a substantial portion of assets within a hedged structure and commonly employ leverage. Where short sales are used, hedged assets may be comprised of an equal dollar value of long and short stock positions. Other variations use short sales unrelated to long holdings and/or puts on the S&P 500 index and put spreads. Conservative funds mitigate market risk by maintaining market exposure from zero to 100 percent. Aggressive funds may magnify market risk by exceeding 100 percent exposure and, in some instances, maintain a short exposure. In addition to equities, some funds may have limited assets invested in other types of securities.

**Equity Market Neutral** investing seeks to profit by exploiting pricing inefficiencies between related equity securities, neutralizing exposure to market risk by combining long and short positions. One example of this strategy is to build portfolios made up of long positions in the strongest companies in several industries and taking corresponding short positions in those showing signs of weakness.

**Equity Market Neutral: Statistical Arbitrage** utilizes quantitative analysis of technical factors to exploit pricing inefficiencies between related equity securities, neutralizing exposure to market risk by combining long and short positions. The strategy is based on quantitative models for selecting specific stocks with equal dollar amounts comprising the long and short sides of the portfolio. Portfolios are typically structured to be market, industry, sector, and dollar neutral.

**Equity Non-Hedge** funds are predominately long equities although they have the ability to hedge with short sales of stocks and/or stock index options. These funds are commonly known as "stock-pickers." Some funds employ leverage to enhance returns. When market conditions warrant, managers may implement a hedge in the portfolio. Funds may also opportunistically short individual stocks. The important distinction between equity non-hedge funds and equity hedge funds is equity non-hedge funds do not always have a hedge in place. In addition to equities, some funds may have limited assets invested in other types of securities.

**Event-Driven** is also known as "corporate life cycle" investing. This involves investing in opportunities created by significant transactional events, such as spin-offs, mergers and acquisitions, bankruptcy reorganizations, recapitalizations and share buybacks. The portfolio of some Event-Driven managers may shift in majority weighting between Risk Arbitrage and Distressed Securities, while others may take a broader scope. Instruments include long and short common and preferred stocks, as well as debt securities and options. Leverage may be used by some managers. Fund managers may hedge against market risk by purchasing S&P put options or put option spreads.

**Fixed Income: Arbitrage** is a market neutral hedging strategy that seeks to profit by exploiting pricing inefficiencies between related fixed income securities while neutralizing exposure to interest rate risk. Fixed Income Arbitrage is a generic description of a variety of strategies involving investment in fixed income instruments, and weighted in an attempt to eliminate or reduce exposure to changes in the yield curve. Managers attempt to exploit relative mispricing between related sets of fixed income securities. The generic types of fixed income hedging trades include: yield-curve arbitrage, corporate versus Treasury yield spreads, municipal bond versus Treasury yield spreads and cash versus futures.

**Fixed Income: Convertible Bonds** funds are primarily long only convertible bonds. Convertible bonds have both fixed income and equity characteristics. If the underlying common stock appreciates, the convertible bond's value should rise to reflect this increased value. Downside protection is offered because if the underlying common stock declines, the convertible bond's value can decline only to the point where it behaves like a straight bond.

**Fixed Income: Diversified** funds may invest in a variety of fixed income strategies. While many invest in multiple strategies, others may focus on a single strategy less followed by most fixed income hedge funds. Areas of focus include municipal bonds, corporate bonds, and global fixed income securities.

**Fixed Income: High-Yield** managers invest in non-investment grade debt. Objectives may range from high current income to acquisition of undervalued instruments. Emphasis is placed on assessing credit risk of the issuer. Some of the available high-yield instruments include extendible/reset securities, increasing-rate notes, pay-in-kind securities, step-up coupon securities, split-coupon securities and usable bonds.

**Fixed Income: Mortgage-Backed** funds invest in mortgage-backed securities. Many funds focus solely on AAA-rated bonds. Instruments include: government agency, government-sponsored enterprise, private-label fixed- or adjustable-rate mortgage pass-through securities, fixed- or adjustable-rate collateralized mortgage obligations (CMOs), real estate mortgage investment conduits (REMICs) and stripped mortgage-backed securities (SMBSs). Funds may look to capitalize on security-specific mispricings. Hedging of prepayment risk and interest rate risk is common. Leverage may be used, as well as futures, short sales and options.

**Macro** involves investing by making leveraged bets on anticipated price movements of stock markets, interest rates, foreign exchange and physical commodities. Macro managers employ a "top-down" global approach, and may invest in any markets using any instruments to participate in expected market movements. These movements may result from forecasted shifts in world economies, political fortunes or global supply and demand for resources, both physical and financial. Exchange-traded and over-the-counter derivatives are often used to magnify these price movements.

**Market Timing** involves allocating assets among investments by switching into investments that appear to be beginning an uptrend, and switching out of investments that appear to be starting a downtrend. This primarily consists of switching between mutual funds and money markets. Typically, technical trend-following indicators are used to determine the direction of a fund and identify buy and sell signals. In an up move "buy signal," money is transferred from a money market fund into a mutual fund in an attempt to capture a capital gain. In a down move "sell signal," the assets in the mutual fund are sold and moved back into the money market for safe keeping until the next up move. The goal is to avoid being invested in mutual funds during a market decline.

**Merger Arbitrage**, sometimes called Risk Arbitrage, involves investment in event-driven situations such as leveraged buy-outs, mergers and hostile takeovers. Normally, the stock of an acquisition target appreciates while the acquiring company's stock decreases in value. These strategies generate returns by purchasing stock of the company being acquired, and in some instances, selling short the stock of the acquiring company. Managers may employ the use of equity options as a low-risk alternative to the outright purchase or sale of common stock. Most Merger Arbitrage funds hedge against market risk by purchasing S&P put options or put option spreads.

**Regulation D** Managers invest in Regulation D securities, sometimes referred to as structured discount convertibles. The securities are privately offered to the investment manager by companies in need of timely financing and the terms are negotiated. The terms of any particular deal are reflective of the negotiating strength of the issuing company. Once a deal is closed, there is a waiting period for the private share offering to be registered with the SEC. The manager can only convert into private shares and cannot trade them publicly during this period; therefore their investment is illiquid until it becomes registered. Managers will hedge with common stock until the registration becomes effective and then liquidate the position gradually.

**Relative Value Arbitrage** attempts to take advantage of relative pricing discrepancies between instruments including equities, debt, options and futures. Managers may use mathematical, fundamental, or technical analysis to determine misvaluations. Securities may be mispriced relative to the underlying security, related securities, groups of securities, or the overall market. Many funds use leverage and seek opportunities globally. Arbitrage strategies include dividend arbitrage, pairs trading, options arbitrage and yield curve trading.

**Sector: Energy** focuses on investment within the energy sector. Investments can be long and short in various instruments with funds either diversified across the entire sector or specializing within a sub-sector.

**Sector: Financial** is a strategy that invests in securities of bank holding companies, banks, thrifts, insurance companies, mortgage banks and various other financial services companies.

**Sector: Healthcare/Biotechnology** funds invest in companies involved in the healthcare, pharmaceutical, biotechnology, and medical device areas.

**Sector: Metals/Mining** funds invest in securities of companies primarily focused on mining, processing and dealing in precious metals and other commodities. Some funds may employ arbitrage strategies on a worldwide basis.

**Sector: Real Estate** involves investing in securities of real estate investment trusts (REITs) and other real estate companies. Some funds may also invest directly in real estate property.

**Sector: Technology** funds emphasize investment in securities of the technology arena. Some of the sub-sectors include multimedia, networking, PC producers, retailers, semiconductors, software, and telecommunications.

**Short Selling** involves the sale of a security not owned by the seller; a technique used to take advantage of an anticipated price decline. To effect a short sale, the seller borrows securities from a third party in order to make delivery to the purchaser. The seller returns the borrowed securities to the lender by purchasing the securities in the open market. If the seller can buy that stock back at a lower price, a profit results. If the price rises, however, a loss results. A short seller must generally pledge other securities or cash with the lender in an amount equal to the market price of the borrowed securities. This deposit may be increased or decreased in response to changes in the market price of the borrowed securities.

**HFRI FOF: Conservative** exhibit one or more of the following characteristics: seeks consistent returns by primarily investing in funds that generally engage in more "conservative" strategies such as Equity Market Neutral, Fixed Income Arbitrage, and Convertible Arbitrage; exhibits a lower historical annual standard deviation than the HFRI Fund of Funds Composite Index. A fund in the HFRI FOF Conservative Index shows generally consistent performance regardless of market conditions.

**HFRI FOF: Diversified** exhibit one or more of the following characteristics: invests in a variety of strategies among multiple managers; historical annual return and/or a standard deviation generally similar to the HFRI Fund of Fund Composite index; demonstrates generally close performance and returns distribution correlation to the HFRI Fund of Fund Composite Index. A fund in the HFRI FOF Diversified Index tends to show minimal loss in down markets while achieving superior returns in up markets.

**HFRI FOF: Market Defensive** exhibit one or more of the following characteristics: invests in funds that generally engage in short-biased strategies such as short selling and managed futures; shows a negative correlation to the general market benchmarks (S&P). A fund in the FOF Market Defensive Index exhibits higher returns during down markets than during up markets.

**HFRI FOF: Strategic** exhibit one or more of the following characteristics: seeks superior returns by primarily investing in funds that generally engage in more opportunistic strategies such as Emerging Markets, Sector specific, and Equity Hedge; exhibits a greater dispersion of returns and higher volatility compared to the HFRI Fund of Funds Composite Index. A fund in the HFRI FOF Strategic Index tends to outperform the HFRI Fund of Fund Composite Index in up markets and underperform the index in down markets.

## Appendix 2

## Replicating Portfolios for the Statistical Factor Model

Given a set of *k* stationary random variables, $X_1$, $X_2$, ...$X_k$, the basic idea in principal component analysis is to find the linear combinations of the original variables so that (1) they explain, successively, the maximum amount of variance possible and (2) they are orthogonal. These linear combination are called principal components. By convention, the first principal component is the linear combination of $X_1$, $X_2$, ...$X_k$ that explains the most variation, the second principal component is the linear combination that (1) explains the most of the remaining variation and (2) is uncorrelated with the first principal component. Each subsequent principal component accounts for as much as possible from the remaining variation and is uncorrelated with the previous principal components. To reproduce the total variation of a k system of variables, one needs exactly k principal components. However, if the first few principal components account for a large part of the total variability, the dimensionality and much of the noise in the original data can be significantly reduced by modelling the system using only the first few principal components.

Our first challenge is to determine the optimal number of principal components to explain hedge fund returns. Principal component analysis heavily relies on the quality of the correlation matrix, so it is important to be able to separate 'true' correlation from noise or measurement errors. Especially for hedge fund returns, where the correlation matrix is computed on small samples the measurement risk is large and separating real information from noise becomes essential. To this end, we make use of random matrix theory and compare the properties of the correlation matrix of all the funds in our sample – the empirical correlation matrix – to the properties of a correlation matrix of an identical number of mutually uncorrelated returns series, following a method proposed by Plerou et al. (2002).[29] Uncorrelated series were constructed by randomly drawing observations from the empirical distribution of our fund returns. If the properties of the empirical correlation matrix are consistent with those of the random matrix, then the information content of the empirical matrix is entirely driven by randomness. However, deviations of the empirical matrix from the properties of the random matrix reveal information about 'true' correlation between hedge fund returns.

We perform the analysis over two sub-periods of 5 years, 1993-1997 and 1998-2002, for a robustness check. In the first period, the sample includes 92 funds and in the second period, 214 funds. We first examine two distributions: of the random correlations and the empirical correlations. As expected, the random correlations have a Gaussian distribution centred on zero. However, the distribution of the empirical correlations is centred on a positive value (0.16 in the first period and 0.20 in the second period). Considering the general market conditions during each period, this may indicate that correlation in hedge funds increases in more volatile periods. The left tail of the empirical correlation distribution shows good agreement with the random correlation distribution, but the observed positive correlations are less likely to be random (the 95% confidence intervals for the random cross correlations is (-.27, .27) in both sub-periods). In each period, we also find that only the four largest empirical eigenvalues significantly exceed the range of the random ones. This suggests that the 'true' correlation pattern can be captured with just the first four principal components and the variation in fund returns relating to the higher eigenvalues is simply uncorrelated noise.

After establishing the number of principal components to be used in our statistical factor model, the next target is to construct portfolios that replicate them. The principal components cannot be used directly in the factor model, as they do not represent investable portfolios – some funds will have negative loadings on the principal components. With no short sale restrictions, we require a selection criterion and an optimisation model which will produce feasible portfolios resembling the first few principal components. After investigating several alternatives, the best replicas were produced through the methodology proposed by Fung and Hsieh (1997): to construct a portfolio to replicate one principal component we selected the funds that were highly correlated solely with that component. Then, the portfolio was optimised to have maximum correlation with the principal component it is replicating, subject to a positive weights constraint. In order to construct replica portfolios for the period Jan-90 to May-03, we split the sample in three: 1990-1993, 1994-1997 and 1998-2003.

[29] We note that general results from RMT are not directly applicable to hedge fund returns, unless the number of datapoints is less than the number of funds included in the analysis, which is not our case.

## Appendix 3

### Dynamic Strategies and Market Timing: A Markov Switching Model

One way of testing for the presence of dynamic strategies is to estimate switching models for the relationship between the strategy returns and the relevant asset class returns. Regime switching models provide a systematic approach to modelling multiple breaks and regime shifts in the data generating process. In a regime-switching model the process is considered to be time-invariant conditional on a state variable that indicates the regime prevailing at the time. Regime shifts are considered to be stochastic, rather than deterministic events, which fits well our problem setting where we have no knowledge of the exact strategies and signals used by individual funds to switch their positions. These models allow us to infer from the pattern of returns the type of strategies followed by the funds, as well as the switching times.

In order to test the existence of switching relationships, we specify a simple Markov switching model with two states for the relationship between the returns on PC2-PC4 portfolios and the major asset classes. We estimate single factor switching models rather than multi-factor in order to avoid the assumption that the switching times are the same for strategies applied to different asset classes.

In the general form of the estimated model, the regression intercept, slope and the variance of the error term are all assumed to be state dependent. Following Hamilton (1994), if we let $s_t$ denote the latent state variable which can take one of K = 2 possible values (i.e. 1 or 2), then the regression model can be written as:

$$\mathbf{y}_t = \mathbf{z}_t' \, b_{S,\,t} + e_{S,\,t}$$

where $\mathbf{y}_t$ is the (T x 1) vector of the statistical factor returns; $\mathbf{z}_t = (\mathbf{1}\ \mathbf{x}_t)$ is the (T x 2) matrix of explanatory variables, with $\mathbf{x}_t$ denoting the fundamental factor returns, $b_{S,t} = (\alpha_{S,t}, \beta_{S,t})$ is the vector of state dependent regression coefficients; $e_{S,t}$ is the vector of state dependent disturbances, assumed normal with state dependent variance $\sigma_{S,t}^2$. The transition probabilities for the two states are assumed to follow a first-order Markov chain and to be constant over time. The 2x2 matrix of transition probabilities is denoted ($p_{ij}$). Now a standard maximum likelihood approach allows the estimation of two sets of coefficients for the regression and variance of the residual terms, together with a set of transition probabilities.

For each portfolio, we attempted to estimate switching models using each of the factors used in the fundamental model – however convergence was achieved for only a subset of them. For these we interpreted as evidence of dynamic strategies the presence of either (1) different signs of the slope coefficient in the two regimes, indicating a strategy that switches between long and short positions in that asset class, or (2) a significant slope coefficient in only one of the states, indicating a strategy that only trades on that asset class at certain times.

As shown by the results below we have been able to identify a number of asset classes for each of the PC2-PC4 portfolios on which switching strategies might have been in place. For example, PC4 portfolio has a positive and strong relationship with W5000 index in one regime and again, a strong but negative relationship with W5000 index returns in the second regime. Similar types of relationships have been identified also for the other two portfolios. The returns of all PC2-PC4 portfolios have regime switching relationships with the returns of GS Commodity Index and small caps SP index, and PC2 and PC4 portfolios also have switching relationships with the returns of W5000.

Therefore, in the statistical factor model where the first principal component captures the common trend in fund returns just as the traditional asset classes do so in the fundamental factor model, the other three principal components are capturing factors related to dynamic hedge fund strategies.

**Markov Switching Models for PC2-PC4 Portfolios**

| Model | | $\alpha_1$ | $\alpha_2$ | $\beta_1$ | $\beta_2$ | $p_{11}$ | $p_{22}$ | $\sigma_1$ | $\sigma_2$ |
|---|---|---|---|---|---|---|---|---|---|
| **PC2/GSCI** | Coefficient | 0.3589 | 1.0970 | **-6.6810** | **3.0220** | 0.7394 | 0.8467 | 0.7034 | 1.9718 |
| | Std error | 0.1332 | 0.2387 | 2.6801 | 4.8432 | 0.4317 | 0.3563 | 0.4055 | 0.2712 |
| | Z-stat | 2.6937 | 4.5953 | -2.4928 | 0.62 | 1.71 | 2.37 | 1.73 | 7.27 |
| | P-value | 0.0071 | 0 | **0.0127** | **0.5326** | 0.0868 | 0.0175 | 0.0828 | 0 |
| **PC2/SC600** | Coefficient | 0.4401 | 1.2313 | **-2.3525** | **-12.36** | 0.7994 | 0.8257 | 0.9032 | 1.9845 |
| | Std error | 0.1543 | 0.3232 | 3.2286 | 5.4138 | 0.3933 | 0.4301 | 1.6491 | 0.3401 |
| | Z-stat | 2.85 | 3.81 | -0.73 | -2.28 | 2.03 | 1.92 | 0.55 | 5.83 |
| | P-value | 0.0043 | 0.0001 | **0.4662** | **0.0224** | 0.0421 | 0.0549 | 0.5839 | 0.0000 |
| **PC2/W5000** | Coefficient | 0.4684 | 1.2472 | **-5.3136** | **-12.85** | 0.7686 | 0.8055 | 0.8688 | 1.9682 |
| | Std error | 0.1556 | 0.3516 | 3.4508 | 5.7183 | 0.4633 | 0.4677 | 1.1965 | 0.3283 |
| | Z-stat | 3.01 | 3.55 | -1.54 | -2.25 | 1.66 | 1.72 | 0.73 | 6.00 |
| | P-value | 0.0026 | 0.0004 | **0.1236** | **0.0246** | 0.0972 | 0.0850 | 0.4678 | 0.0000 |
| **PC3/GSCI** | Coefficient | 0.4742 | 1.2345 | **0.4623** | **4.6786** | 0.9387 | 0.9859 | 0.6223 | 1.4258 |
| | Std error | 0.1089 | 0.1369 | 2.1775 | 2.2343 | 0.5143 | 0.3099 | 0.2243 | 0.2323 |
| | Z-stat | 4.36 | 9.02 | 0.21 | 2.09 | 1.83 | 3.18 | 2.77 | 6.14 |
| | P-value | 0.0000 | 0.0000 | **0.8319** | **0.0363** | 0.0680 | 0.0015 | 0.0055 | 0.0000 |
| **PC3/SC600** | Coefficient | 0.8394 | 1.2777 | **1.4383** | **31.3582** | 0.9465 | 0.8991 | 1.0072 | 1.3529 |
| | Std error | 0.1150 | 0.2566 | 2.1758 | 7.7152 | 0.3195 | 0.3452 | 13.9138 | 0.4904 |
| | Z-stat | 7.30 | 4.98 | 0.66 | 4.06 | 2.96 | 2.60 | 0.07 | 2.76 |
| | P-value | 0.0000 | 0.0000 | **0.5086** | **0.0000** | 0.0031 | 0.0092 | 0.9423 | 0.0058 |
| **PC4/GSCI** | Coefficient | 0.9112 | 1.2406 | **-7.6743** | **6.1992** | 0.8843 | 0.8610 | 0.8179 | 1.7135 |
| | Std error | 0.1180 | 0.2760 | 2.2306 | 5.9722 | 0.3285 | 0.3619 | 0.5302 | 0.3649 |
| | Z-stat | 7.72 | 4.49 | -3.44 | 1.04 | 2.69 | 2.38 | 1.54 | 4.70 |
| | P-value | 0.0000 | 0.0000 | **0.0006** | **0.2993** | 0.0071 | 0.0173 | 0.1229 | 0.0000 |
| **PC4/SC600** | Coefficient | 0.7488 | 3.8130 | **11.4778** | **-37.684** | 0.9409 | 0.4120 | 0.9578 | 0.9389 |
| | Std error | 0.0854 | 0.5134 | 1.9607 | 16.172 | 0.1817 | 0.9257 | 1.6571 | 4.5243 |
| | Z-stat | 8.77 | 7.43 | 5.85 | -2.33 | 5.18 | 0.45 | 0.58 | 0.21 |
| | P-value | 0.0000 | 0.0000 | **0.0000** | **0.0198** | 0.0000 | 0.6563 | 0.5633 | 0.8356 |
| **PC4/W5000** | Coefficient | 1.3653 | 0.9038 | **-15.4912** | **17.293** | 0.9583 | 0.9893 | 0.8197 | 1.2662 |
| | Std error | 0.1413 | 0.1498 | 3.1131 | 4.1038 | 0.2946 | 0.4800 | 0.5420 | 0.2984 |
| | Z-stat | 9.66 | 6.03 | -4.98 | 4.21 | 3.25 | 2.06 | 1.51 | 4.24 |
| | P-value | 0.0000 | 0.0000 | **0.0000** | **0.0000** | 0.0011 | 0.0393 | 0.1304 | 0.0000 |

**Table 1 Base case factor model estimation results**

| | All Funds | Convert Arb | Distress Sec | Emerge Markets | Equity Hedge | Equity MN | Equity Non-Hedge | Event-Driven | Fixed Income | FoF | Market Timing | Manag Futures | Merger Arb | Sectors |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Alpha | 0.62 | 0.64 | 0.44 | 1.11 | 0.75 | 0.38 | 0.65 | 0.60 | 0.32 | 0.48 | 0.70 | 0.93 | 0.35 | 0.80 |
| | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% |
| W5000 | 0.38 | 0.04 | **0.22** | **0.78** | **0.61** | **0.21** | **0.76** | **0.39** | **0.23** | **0.19** | **0.07** | **-0.23** | **0.10** | **0.60** |
| | 80% | 50% | **76%** | **100%** | **83%** | **65%** | **95%** | **95%** | **65%** | **81%** | **80%** | **61%** | **50%** | **100%** |
| Lagged W5000 | 0.12 | 0.05 | **0.15** | **0.44** | 0.17 | 0.00 | **0.28** | **0.10** | 0.07 | **0.06** | NA | -0.15 | **0.06** | 0.35 |
| | 38% | 40% | **76%** | **50%** | 30% | 15% | **42%** | **45%** | 35% | **62%** | 0% | 28% | **75%** | 5% |
| Lehman Bond | -0.25 | **0.13** | -0.40 | NA | -0.68 | -0.15 | 0.45 | -0.20 | 0.09 | -0.46 | -0.52 | 1.15 | -0.17 | 0.14 |
| | 20% | **60%** | 12% | 0% | 25% | 19% | 5% | 5% | 18% | 29% | 20% | 17% | 25% | 16% |
| 1st yr rep dummy | 0.32 | **1.42** | -0.16 | -3.68 | -0.18 | 2.25 | -1.39 | **-0.53** | 0.79 | -0.10 | 1.61 | 2.36 | NA | -2.15 |
| | 24% | **60%** | 12% | 8% | 25% | 19% | 11% | **40%** | 35% | 29% | 20% | 28% | 0% | 11% |
| $R^2$ | 0.27 | 0.20 | 0.16 | 0.23 | 0.32 | 0.19 | 0.39 | 0.32 | 0.30 | 0.28 | 0.24 | 0.13 | 0.11 | 0.29 |

**Table 2 Fundamental factor model estimation results**

| | All Funds | Convert Arb | Distress Sec | Emerge Markets | Equity Hedge | Equity MN | Equity Non-Hedge | Event-Driven | Fixed Income | FoF | Market Timing | Manag Futures | Merger Arb | Sectors |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Alpha | 0.55 | 0.81 | 0.63 | 1.35 | 0.58 | 0.40 | 0.40 | 0.57 | 0.36 | 0.51 | 0.62 | 0.34 | 0.47 | 0.57 |
| | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% |
| W5000 | 0.62 | 0.04 | 0.18 | **1.14** | 0.82 | NA | **0.84** | 0.64 | 0.26 | 0.28 | 0.36 | 0.69 | NA | 0.63 |
| | 18% | 10% | 6% | **50%** | 21% | 0% | **42%** | 5% | 29% | 21% | 20% | 6% | 0% | 21% |
| SP500g | 0.23 | 0.12 | NA | NA | 0.29 | **0.29** | NA | NA | -0.08 | 0.01 | **0.23** | NA | NA | 0.52 |
| | 7% | 20% | 0% | 0% | 2% | **31%** | 0% | 0% | 6% | 2% | **33%** | 0% | 0% | 5% |
| SP500v | 0.22 | -0.01 | 0.29 | 0.21 | 0.32 | 0.26 | 0.47 | 0.09 | 0.08 | -0.31 | NA | NA | NA | 0.45 |
| | 7% | 20% | 12% | 8% | 2% | 15% | 5% | 5% | 6% | 4% | 0% | 0% | 0% | 26% |
| MD400 | 0.40 | NA | 0.40 | 0.24 | 0.46 | 0.10 | 0.58 | 0.31 | NA | 0.15 | 0.15 | NA | 0.04 | 1.01 |
| | 10% | 0% | 6% | 25% | 21% | 4% | 11% | 15% | 0% | 6% | 7% | 0% | 25% | 11% |
| SC600 | **0.39** | 0.09 | **0.15** | 0.45 | **0.55** | 0.34 | 0.81 | **0.37** | 0.29 | **0.22** | 0.33 | NA | **0.07** | 0.76 |
| | **38%** | 10% | **47%** | 8% | **47%** | 15% | 37% | **70%** | 24% | **62%** | 7% | 0% | **75%** | 37% |
| MSCIW EXUS | 0.00 | 0.09 | -0.08 | 0.13 | 0.19 | -0.01 | 0.41 | NaN | 0.04 | 0.05 | NaN | **-0.25** | NaN | 0.16 |
| | 9% | 10% | 6% | 8% | 6% | 8% | 16% | 0% | 6% | 4% | 0% | **50%** | 0% | 5% |
| MSCI EMF | 0.18 | 0.07 | 0.17 | **0.56** | 0.18 | 0.04 | 0.21 | 0.15 | 0.14 | **0.09** | 0.06 | -0.22 | NaN | 0.10 |
| | 20% | 20% | 24% | **67%** | 21% | 12% | 16% | 5% | 12% | **31%** | 20% | 6% | 0% | 11% |
| LEH GOV | 0.25 | 0.88 | NaN | NaN | 0.30 | 0.41 | 0.66 | NaN | 0.11 | -0.22 | -0.69 | 1.12 | NaN | 0.41 |
| | 6% | 10% | 0% | 0% | 4% | 12% | 5% | 0% | 6% | 4% | 13% | 6% | 0% | 16% |
| LEH CREDIT | -0.21 | 0.09 | -0.23 | NaN | -0.33 | NaN | NaN | -0.47 | 0.03 | -0.22 | NaN | NaN | -0.10 | NaN |
| | 6% | 20% | 24% | 0% | 6% | 0% | 0% | 10% | 6% | 6% | 0% | 0% | 25% | 0% |
| LEH HY | 0.09 | **0.20** | **0.27** | -0.46 | -0.11 | -0.06 | 0.10 | **0.26** | 0.17 | 0.08 | 0.12 | 0.40 | **0.03** | **-0.02** |
| | 24% | **30%** | **65%** | 8% | 23% | 12% | 21% | **30%** | 6% | 17% | 27% | 22% | **50%** | **37%** |
| LEH MBKD | -0.20 | NA | -0.65 | NA | NA | NA | NA | 0.19 | 0.27 | -0.39 | NA | 0.05 | NA | -1.11 |
| | 5% | 0% | 6% | 0% | 0% | 0% | 0% | 5% | 18% | 10% | 0% | 11% | 0% | 5% |
| FX | -0.45 | 0.08 | NA | NA | -0.48 | -0.18 | -0.77 | -0.51 | -0.23 | 0.02 | NA | **-0.65** | NA | -0.41 |
| | 11% | 10% | 0% | 0% | 8% | 8% | 16% | 5% | 12% | 8% | 0% | **67%** | 0% | 16% |
| GSCI Com | 0.02 | -0.03 | -0.13 | **0.17** | -0.01 | -0.06 | -0.03 | 0.08 | -0.01 | 0.05 | 0.08 | **-0.18** | -0.07 | 0.27 |
| | 19% | 10% | 12% | **42%** | 8% | 12% | 26% | 15% | 12% | 23% | 20% | **39%** | 25% | 26% |
| L W5000 | 0.13 | 0.03 | 0.11 | 0.19 | 0.14 | 0.14 | 0.25 | 0.09 | 0.08 | 0.07 | -0.27 | 0.38 | **0.05** | 0.21 |
| | 18% | 10% | 18% | 8% | 28% | 4% | 26% | 25% | 18% | 17% | 7% | 6% | **50%** | 21% |
| SC600^2 | 0.00 | **-0.01** | **-0.01** | **-0.02** | **0.01** | **0.00** | **0.00** | -0.01 | **0.00** | **0.00** | **0.01** | **0.01** | 0.00 | 0.01 |
| | 39% | **60%** | **53%** | **33%** | **36%** | **50%** | **42%** | 25% | **41%** | **35%** | **33%** | **50%** | 25% | 26% |
| LEH HY^2 | 0.01 | **-0.04** | **0.02** | **0.00** | -0.02 | 0.02 | **0.13** | 0.00 | **0.01** | -0.02 | **-0.01** | **0.04** | **-0.01** | **0.01** |
| | 28% | **40%** | **47%** | **33%** | 26% | 19% | **37%** | 5% | **35%** | 19% | **40%** | **33%** | **50%** | **37%** |
| VIX | 0.00 | 0.00 | 0.00 | -0.01 | **0.00** | **0.00** | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | **0.00** | 0.00 | **0.00** |
| | 27% | 20% | 18% | 25% | **30%** | **38%** | 11% | 15% | 6% | 27% | 27% | **61%** | 25% | **32%** |
| DISP | 0.56 | NA | 0.63 | -0.95 | **1.97** | -1.09 | 1.59 | -0.88 | 0.40 | **0.44** | -6.18 | -2.95 | -1.37 | **1.41** |
| | 29% | 0% | 18% | 25% | **45%** | 23% | 26% | 20% | 12% | **33%** | 13% | 17% | 25% | **58%** |
| 1st yr rep dummy | 0.43 | **0.92** | 0.23 | -1.39 | 0.05 | 1.55 | -1.33 | **0.48** | **0.75** | 0.15 | 1.41 | 3.52 | NA | -0.44 |
| | 25% | **50%** | 24% | 17% | 28% | 23% | 11% | **40%** | **35%** | 29% | 20% | 11% | 0% | 16% |
| $R^2$ | 0.36 | 0.27 | 0.29 | 0.36 | 0.42 | 0.24 | 0.49 | 0.43 | 0.38 | 0.39 | 0.26 | 0.22 | 0.23 | 0.42 |

**Table 3 HFRI factor model estimation results**

| | All Funds | Convert Arb | Distressed Securities | Emerge Markets | Equity Hedge | Equity MN | Equity Non-Hedge | Event-Driven | Fixed Income | FoF | Market timing | Manag Futures | Merger Arb | Sectors |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Alpha | -0.12 | 0.18 | -0.34 | 0.34 | -0.24 | -0.12 | -0.58 | -0.56 | 0.02 | -0.01 | 0.02 | 0.56 | -0.07 | -0.31 |
| | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% |
| Conv Arb | 0.27 | **0.89** | -0.25 | 1.84 | 0.3 | 0.12 | -0.32 | 0.39 | **0.42** | 0.11 | 0.07 | -0.27 | **-0.06** | -0.74 |
| | 25% | **50%** | 6% | 33% | 25% | 12% | 11% | 25% | **47%** | 29% | 33% | 28% | **75%** | 11% |
| Regul D | -0.03 | **-0.01** | 0.00 | **-0.42** | 0.18 | -0.08 | -0.32 | -0.04 | **-0.06** | 0.10 | **0.02** | -0.69 | 0.06 | 0.05 |
| | 27% | **50%** | 29% | **42%** | 25% | 12% | 26% | 15% | **53%** | 27% | **40%** | 11% | 25% | 26% |
| Rel Val | -0.34 | NA | -1.48 | **-1.64** | -0.60 | -0.24 | -0.39 | -0.04 | **0.09** | -0.42 | 0.50 | -0.76 | -0.86 | 0.59 |
| | 28% | 0% | 6% | **50%** | 23% | 35% | 21% | 35% | **41%** | 27% | 33% | 33% | 25% | 37% |
| Distress Sec | 0.39 | 0.23 | **1.30** | 3.39 | -0.40 | -0.12 | 0.12 | **1.01** | -0.65 | 0.17 | 0.83 | -0.61 | 0.11 | -0.95 |
| | 30% | 30% | **100%** | 17% | 26% | 15% | 26% | **60%** | 18% | 33% | 7% | 28% | 25% | 11% |
| Em Mk (Total) | 0.33 | -0.03 | -0.20 | **1.44** | -0.02 | 0.19 | 0.15 | 0.29 | 0.04 | 0.21 | -0.10 | -0.26 | -0.13 | -0.10 |
| | 22% | 30% | 18% | **100%** | 8% | 15% | 21% | 35% | 29% | 17% | 33% | 11% | 25% | 16% |
| Eq Hedge | 0.68 | -0.25 | -0.12 | NA | **1.19** | 0.98 | 1.10 | 0.46 | NA | 0.47 | 0.39 | -0.58 | -0.13 | 1.03 |
| | 21% | 20% | 12% | 0% | **40%** | 12% | 11% | 10% | 0% | 31% | 13% | 22% | 25% | 21% |
| Eq MN | 0.29 | -0.36 | -0.43 | 1.56 | 0.05 | **1.08** | 0.81 | 0.06 | -0.27 | 0.18 | -0.25 | NA | NA | -0.06 |
| | 24% | 20% | 18% | 8% | 30% | **58%** | 21% | 10% | 18% | 21% | 27% | 0% | 0% | 37% |
| Eq Non-Hedge | 0.75 | -0.20 | 0.00 | -0.25 | 1.13 | 0.51 | **1.14** | 0.42 | 0.37 | 0.19 | 0.34 | 0.71 | 0.08 | 0.78 |
| | 16% | 10% | 12% | 8% | 19% | 15% | **68%** | 15% | 12% | 4% | 13% | 6% | 25% | 11% |
| Ev Driven | -0.13 | NA | -0.65 | **-1.65** | 0.34 | -0.04 | 0.18 | 0.56 | -0.02 | 0.17 | -0.69 | -1.31 | **0.00** | 0.23 |
| | 23% | 0% | 18% | **42%** | 11% | 15% | 16% | 50% | 12% | 27% | 20% | 33% | **75%** | 26% |
| Fixed Inc (Total) | -0.18 | **0.61** | **-0.54** | NA | 0.49 | -1.68 | -0.15 | -0.08 | **-0.45** | -0.63 | 0.08 | -2.44 | **0.15** | **0.95** |
| | 34% | **40%** | **47%** | 0% | 34% | 19% | 32% | 35% | **53%** | 38% | 33% | 17% | **50%** | **47%** |
| Fixed Inc Conv Arb | -0.01 | -0.19 | 0.20 | 0.86 | -0.20 | 0.59 | -0.20 | 0.13 | 0.19 | 0.13 | 0.08 | -0.43 | NA | -0.3 |
| | 20% | 20% | 35% | 8% | 17% | 4% | 26% | 20% | 35% | 23% | 7% | 33% | 0% | 11% |
| Fixed Inc HY | 0.03 | NA | 0.07 | **-0.26** | -0.13 | 0.24 | 0.96 | -0.75 | **0.78** | -0.05 | -0.27 | 0.05 | **0.02** | 0.09 |
| | 26% | 0% | 35% | **42%** | 19% | 15% | 16% | 30% | **47%** | 23% | 40% | 28% | **50%** | 26% |
| Fixed Inc Arbitrage | -0.09 | 0.26 | **-0.22** | -0.57 | 0.13 | -0.30 | -0.49 | -0.41 | 0.35 | 0.16 | 0.41 | -1.02 | 0.53 | -0.37 |
| | 25% | 10% | **41%** | 25% | 25% | 19% | 26% | 35% | 35% | 19% | 33% | 11% | 25% | 32% |
| Fixed Inc Diversified | 0.27 | -0.40 | 0.66 | 0.76 | -0.59 | 0.37 | 0.78 | 0.30 | **0.46** | **0.17** | -0.80 | **1.55** | **-0.19** | -0.48 |
| | 34% | 30% | 35% | 25% | 23% | 19% | 16% | 15% | **41%** | **48%** | 33% | **83%** | **50%** | 37% |
| Fixed Inc MBKD | -0.09 | -0.51 | 0.02 | -0.42 | -0.25 | 0.37 | 0.14 | 0.08 | 0.19 | 0.19 | -0.36 | -0.22 | **-0.20** | -0.67 |
| | 28% | 20% | 29% | 33% | 34% | 19% | 11% | 20% | 47% | 29% | 33% | 28% | **50%** | 26% |
| FoF | 0.22 | 0.07 | 0.04 | -1.34 | -0.21 | -0.55 | -0.66 | -0.49 | 0.33 | **0.63** | 0.08 | **1.00** | NA | **-0.36** |
| | 30% | 20% | 18% | 8% | 23% | 8% | 26% | 20% | 18% | **65%** | 13% | **50%** | 0% | **42%** |
| Mk Timing | 0.08 | **-0.05** | **-0.14** | 0.24 | 0.14 | 0.05 | **-0.08** | 0.12 | 0.07 | 0.05 | **0.41** | -0.87 | NA | 0.46 |
| | 34% | **40%** | **41%** | 17% | 38% | 31% | **42%** | 25% | 29% | 37% | **67%** | 17% | 0% | 26% |
| Macro | 0.31 | NA | 0.12 | -0.60 | 0.19 | 0.23 | 0.34 | 0.32 | 0.00 | 0.12 | 0.47 | **1.00** | 0.04 | 0.23 |
| | 28% | 0% | 18% | 25% | 26% | 27% | 21% | 25% | 24% | 23% | 33% | **72%** | 25% | 37% |
| Short Sell | 0.04 | 0.22 | -0.04 | -0.69 | -0.03 | 0.09 | 0.08 | 0.03 | -0.12 | 0.11 | 0.24 | 0.23 | NA | -0.23 |
| | 31% | 10% | 35% | 8% | 36% | 15% | 37% | 35% | 18% | 37% | 47% | 39% | 0% | 32% |
| Merger Arb | 0.60 | 0.25 | 0.51 | **2.24** | 0.89 | 0.57 | 1.29 | **0.24** | -0.33 | -0.06 | 0.89 | 0.52 | 0.80 | 0.66 |
| | 30% | 10% | 18% | **42%** | 32% | 23% | 32% | **50%** | 29% | 31% | 20% | 6% | 100% | 37% |
| Sector (Total) | -0.02 | 0.10 | -0.05 | -0.55 | 0.30 | -0.13 | -0.03 | 0.24 | -0.21 | -0.05 | NA | 0.76 | NA | -0.54 |
| | 13% | 30% | 35% | 25% | 13% | 8% | 11% | 15% | 12% | 8% | 0% | 6% | 0% | 16% |
| Energy | 0.05 | **0.00** | 0.06 | 0.12 | 0.04 | -0.07 | 0.07 | **0.02** | -0.02 | 0.01 | 0.02 | **0.27** | NA | **0.11** |
| | 30% | **40%** | 6% | 8% | 32% | 35% | 32% | **45%** | 35% | 23% | 13% | **44%** | 0% | **58%** |
| Financial | -0.03 | -0.04 | -0.14 | -0.13 | 0.02 | 0.10 | **-0.16** | 0.06 | NA | -0.03 | -0.07 | **-0.20** | -0.06 | 0.20 |
| | 31% | 10% | 24% | 25% | 36% | 15% | **47%** | 35% | 0% | 37% | 33% | **50%** | 25% | 32% |
| HC/Bio | -0.02 | 0.01 | -0.11 | -0.04 | 0.02 | **-0.10** | -0.10 | 0.01 | 0.07 | 0.03 | -0.19 | -0.32 | **0.02** | 0.39 |
| | 24% | 30% | 12% | 25% | 13% | **54%** | 21% | 30% | 12% | 23% | 13% | 28% | **50%** | 32% |
| Real Estate | 0.10 | **-0.01** | -0.51 | 0.44 | -0.16 | 0.41 | 0.33 | 0.13 | 0.11 | 0.07 | 0.56 | -0.61 | **-0.11** | 0.15 |
| | 20% | **40%** | 12% | 17% | 17% | 12% | 32% | 5% | 29% | 21% | 27% | 6% | **50%** | 32% |
| Techn | 0.37 | NA | 0.17 | NA | 0.41 | -0.05 | 0.53 | NA | -0.03 | 0.07 | 0.38 | NA | NA | 0.67 |
| | 7% | 0% | 12% | 0% | 9% | 8% | 5% | 0% | 6% | 2% | 13% | 0% | 0% | 26% |
| Sect misc | 0.06 | 0.01 | -0.10 | 0.81 | 0.16 | -0.11 | 0.00 | -0.06 | NA | -0.03 | -0.16 | 0.32 | **0.09** | **0.01** |
| | 22% | 30% | 12% | 8% | 23% | 15% | 16% | 30% | 0% | 23% | 13% | 39% | **50%** | **42%** |
| 1st yr rep dummy | 0.46 | **0.90** | 0.42 | NA | 0.52 | 0.81 | -1.15 | 0.30 | 1.71 | -0.02 | 1.14 | 1.99 | NA | -0.69 |
| | 28% | **40%** | 35% | 0% | 34% | 27% | 21% | 30% | 12% | 27% | 27% | 39% | 0% | 32% |
| $R^2$ | 0.58 | 0.42 | 0.56 | 0.58 | 0.60 | 0.37 | 0.62 | 0.66 | 0.59 | 0.67 | 0.51 | 0.47 | 0.54 | 0.68 |

**Table 4 Statistical factor model estimation results**

| | All Funds | Convert Arb | Distress Sec | Emerge Markets | Equity Hedge | Equity MN | Equity Non-Hedge | Event-Driven | Fixed Income | FoF | Market timing | Manag Futures | Merger Arb | Sectors |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Alpha | -0.21 | 0.56 | 0.36 | 0.34 | -0.57 | 0.01 | -1.09 | -0.15 | 0.30 | -0.07 | 0.27 | -0.92 | 0.31 | -0.65 |
| | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% |
| PC1 | 0.60 | 0.08 | 0.51 | 1.80 | 0.93 | 0.16 | 1.17 | 0.80 | 0.30 | 0.39 | 0.06 | -0.18 | 0.13 | 1.12 |
| | 79% | 50% | 88% | 100% | 85% | 46% | 95% | 100% | 65% | 92% | 87% | 28% | 75% | 89% |
| PC2 | 0.19 | 0.03 | 0.06 | 0.27 | -0.07 | 0.15 | 0.05 | 0.15 | 0.02 | 0.17 | 0.09 | 1.86 | 0.00 | 0.04 |
| | 39% | 20% | 18% | 33% | 34% | 46% | 42% | 30% | 35% | 44% | 33% | 94% | 0% | 37% |
| PC3 | -0.02 | -0.07 | -0.34 | -1.37 | 0.28 | 0.03 | 0.06 | -0.24 | -0.19 | 0.06 | -0.06 | 0.18 | -0.04 | 0.16 |
| | 44% | 20% | 47% | 58% | 45% | 42% | 37% | 65% | 35% | 50% | 33% | 33% | 25% | 37% |
| PC4 | 0.16 | 0.02 | -0.17 | 0.21 | 0.24 | 0.16 | 0.59 | 0.14 | -0.07 | 0.02 | 0.37 | 0.24 | 0.04 | 0.23 |
| | 29% | 10% | 18% | 17% | 36% | 38% | 53% | 35% | 18% | 25% | 20% | 22% | 25% | 37% |
| 1st yr rep | 0.21 | 0.85 | 0.49 | 0.39 | 0.33 | 0.29 | 0.27 | -0.16 | 0.24 | 0.00 | 0.51 | 0.51 | -0.15 | -0.45 |
| dummy | 23% | 60% | 12% | 8% | 30% | 31% | 16% | 25% | 35% | 21% | 13% | 11% | 0% | 16% |
| $R^2$ | 0.39 | 0.17 | 0.23 | 0.33 | 0.47 | 0.22 | 0.49 | 0.46 | 0.32 | 0.46 | 0.33 | 0.52 | 0.12 | 0.46 |

**Table 5 Average fund alphas**

| | Alpha | P-value |
|---|---|---|
| Base case model | 0.6890 | 0.0000 |
| Broad fundamental factor model | 0.5103 | 0.0000 |
| Hedge fund index model | 0.0057 | 0.3967 |
| Statistical factor model | -0.2787 | 0.0031 |

**Table 6 Correlations and probability of agreement between different models' alpha**

| **Correlation** | Two index model | Fundamental factor model | Multi factor HFRI model | Statistical factor model |
|---|---|---|---|---|
| Two index model | 1 | 0.7382 | 0.3993 | 0.2710 |
| Fundamental factor model | | 1 | 0.3927 | 0.5067 |
| Multi factor HFRI model | | | 1 | 0.4500 |
| Statistical factor model | | | | 1 |

| **Rank correlation** | Two index model | Fundamental factor model | Multi factor HFRI model | Statistical factor model |
|---|---|---|---|---|
| Two index model | 1 | 0.6804 | 0.3466 | 0.1941 |
| Fundamental factor model | | 1 | 0.3175 | 0.4560 |
| Multi factor HFRI model | | | 1 | 0.4338 |
| Statistical factor model | | | | 1 |

| **Prob of agreement** | Two index model | Fundamental factor model | Multi factor HFRI model | Statistical factor model |
|---|---|---|---|---|
| Two index model | 1 | 0.7621 | 0.6564 | 0.5881 |
| Fundamental factor model | | 1 | 0.6266 | 0.6605 |
| Multi factor HFRI model | | | 1 | 0.6884 |
| Statistical factor model | | | | 1 |

**Table 7 Average performance of randomly selected minimum variance portfolios (out-of-sample)**

| | Bounded | | Unbounded | |
|---|---|---|---|---|
| | Sample matrix | Cleaned matrix | Sample matrix | Cleaned matrix |
| Annual volatility | 2.32 | 2.30 | 1.97 | 2.54 |
| Annual returns | 7.49 | 7.20 | 7.92 | 7.69 |
| Skewness | 0.03 | -0.26 | -0.15 | -0.40 |
| Excess kurtosis | 5.47 | 5.73 | 4.14 | 5.28 |
| Info ratio | 3.22 | 3.12 | 4.01 | 3.02 |

**Table 8 Performance of portfolios selected using alphas from different factor models and optimised to have minimum variance. Equally weighted portfolio of all funds included for comparison.**

| | Base case model | Fundamental factors model | HFR indices model | Statistical factor model | All factor models | Equally weighted |
|---|---|---|---|---|---|---|
| Annual returns | 8.28 | 8.15 | 9.44 | 8.94 | 9.06 | 10.44 |
| Annual volatility | 1.35 | 1.34 | 1.70 | 1.29 | 1.51 | 6.89 |
| Skewness | 0.30 | 0.06 | 0.10 | 0.22 | 0.49 | -0.07 |
| Excess kurtosis | -0.08 | -0.03 | 0.24 | -0.34 | 0.35 | 1.91 |
| Info ratio | 6.15 | 6.06 | 5.56 | 6.91 | 5.99 | 1.51 |
| Turnover | 6.30 | 7.13 | 7.66 | 4.94 | 7.02 | 4.92 |

**Table 9 Performance of portfolios selected using alphas from different factor models and optimised to have maximum information ratio. Equally weighted portfolio of all funds included for comparison.**

| | Base case model | Fundamental factors model | HFR indices model | Statistical factors model | All factor models | Equally weighted |
|---|---|---|---|---|---|---|
| Annual returns | 9.24 | 8.55 | 10.39 | 8.98 | NA | 10.44 |
| Annual volatility | 1.86 | 1.81 | 1.97 | 1.44 | NA | 6.89 |
| Skewness | 0.71 | 0.01 | 0.57 | 0.50 | NA | -0.07 |
| Excess kurtosis | 0.57 | 1.92 | 0.46 | 0.10 | NA | 1.91 |
| Info ratio | 4.97 | 4.73 | 5.28 | 6.22 | NA | 1.51 |
| Turnover | 8.99 | 9.77 | 7.05 | 7.59 | NA | 4.92 |

**Figure 1 Percentage of funds added to the database vs. percentage of funds that ceased reporting during the previous 12 months**

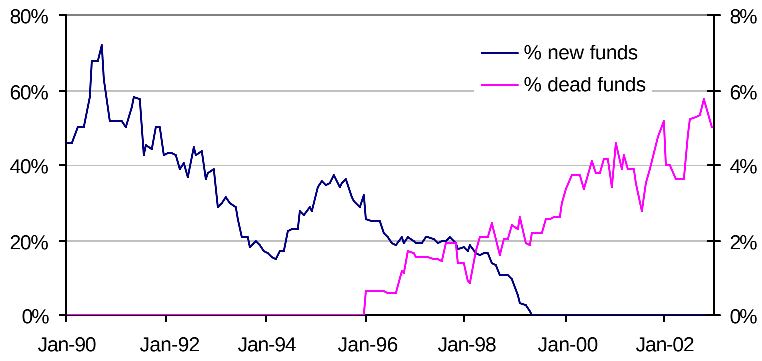


**Figure 2 Average monthly excess return over the first k months of reporting**

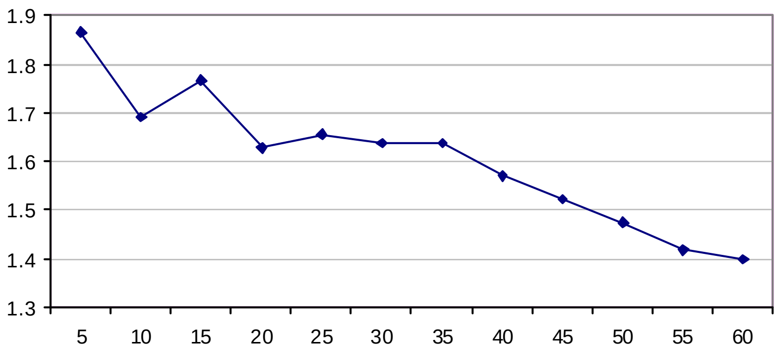


**Figure 3 Distribution of the difference between the monthly average excess return in the last year and the monthly average excess return in the last five years for the funds exiting the database**

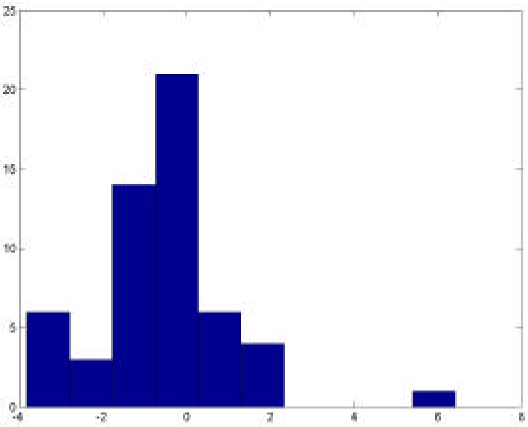

**Figure 4 Average information ratio of randomly selected, equally weighted portfolios**

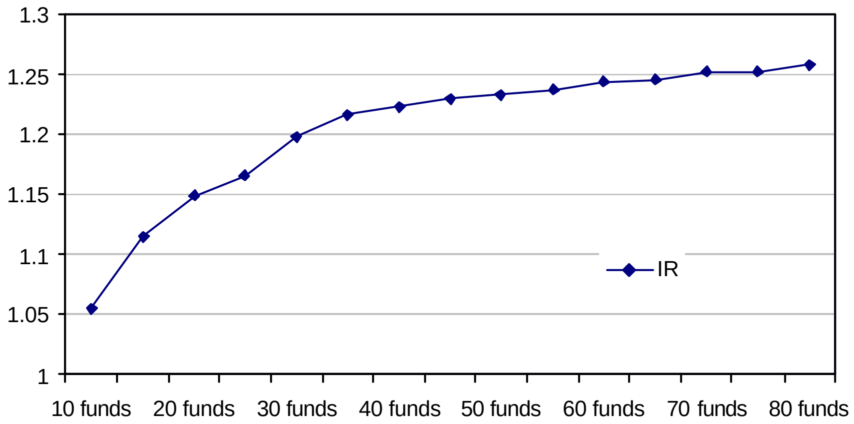


**Figure 5 Average skewness and kurtosis of randomly selected, equally weighted por tfolios**

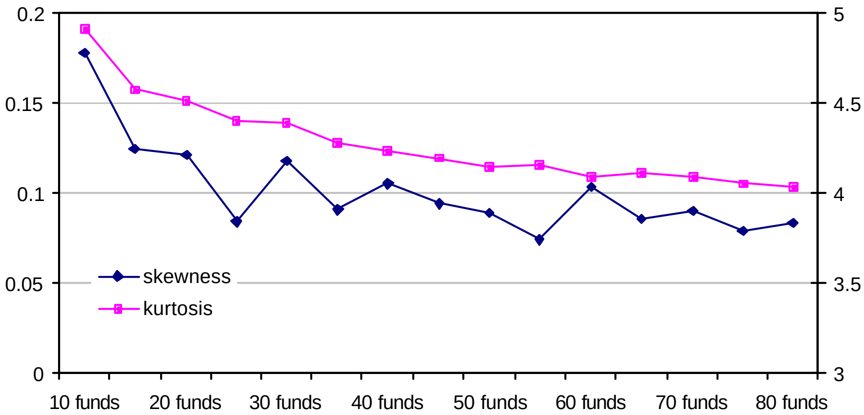


**Figure 6 Dollar value of portfolios selected using alphas from different factor models and optimised to have minimum variance. Equally weighted portfolio included for comparison**