2-2005

Is stellar hedge fund performance for real?

Robert KOSOWSKI
INSEAD

Narayan Y. NAIK
London Business School

Melvyn TEO
Singapore Management University, melvynteo@smu.edu.sg

Is Stellar Hedge Fund Performance for Real?

Robert Kosowski

Narayan Y. Naik*

Melvyn Teo

Abstract

We apply a robust bootstrap to evaluate the performance of a large universe of hedge funds. Our bootstrap estimates indicate that the performance of the top hedge funds cannot be attributed to chance alone. This is true even after adjusting for back fill bias, serial correlation, and structural breaks. Also, we find that hedge fund alpha differences persist over three year horizons. However, an investment strategy designed around this will run into difficulties as the persistence is often confined to small funds that are effectively closed to new inflows. Moreover, Bayesian estimates suggest that standard alphas may be overestimated by 41% for the average top fund.

* Corresponding author: Narayan Y. Naik, London Business School, Sussex Place, Regent's Park, London NW1 4SA, United Kingdom. E-mail: nnaik@london.edu, Tel: +44-20-7262 5050, extension 3579 and Fax: +44-20-7724 3317. Robert Kosowski is Assistant Professor of Finance at INSEAD, Narayan Y. Naik is Associate Professor of Finance at the London Business School, and Melvyn Teo is Assistant Professor of Finance at the Singapore Management University. We would like to thank Vikas Agarwal, William Fung, Alan Timmerman, and Russ Wermers for many helpful comments and constructive suggestions. Naik is grateful for funding from Centre for Hedge Fund Research and Education at the London Business School. We are grateful to CISDM, HFR, MSCI, and TASS for providing us with the data. We are responsible for all errors.

1 Introduction

Were stellar hedge funds like George Soros' Quantum Fund just lucky and their existence to be expected with such a large sample of 5000 hedge funds in 2003? If not, does their abnormal performance persist and can it be exploited by means of trading strategies? What drives top performance? These questions are increasingly on the minds of institutional and retail investors who have recently raised their portfolio allocations to hedge funds.[1] The aim of this paper is to answer these questions by examining a comprehensive hedge fund database using a robust bootstrap method and a Bayesian framework.

The hedge fund industry has changed considerably over the last decade. The HFR 2003 report indicates that there were 530 hedge funds managing under US$39 billion in 1990, while there are over 5000 hedge funds managing over US$817 billion by end of 2003. The strategy mix of the hedge fund industry has also shifted notably. In 1990, the industry was dominated by funds following the Global Macro strategy. By the end of 2003, the largest number of funds belonged to equity based strategies like Equity Hedge, Hedge Long-Bias, Event Driven, etc.

Our understanding of the risk-return trade-offs for different hedge fund strategies has improved significantly in recent years, thanks to the pioneering work of Fung and Hsieh (1999, 2001, 2002), Mitchell and Pulvino (2001) and more recently, Agarwal and Naik (2004).[2] Specifically, we know now that hedge fund returns relate to conventional asset class returns in a linear as well as an option-like way. They also relate to the returns of the small cap minus large cap spread and the

[1] About-one third of institutional investors including CalPERS in America plan to increase their allocation to such funds. ("For the fortunate few," The Economist, print edition, 3rd July 2003.)

In Europe several countries including Luxembourg, Germany and Ireland have revised the rules and regulations governing the creation and distribution of hedge funds and followed the trend set in Asia by permitting the distribution of hedge funds to retail investors. ("Hedge Funds entering the 'Mass-Market' Arena" by Michael Ferguson, Ernst and Young, Luxembourg, February 2004.)

[2] Fung and Hsieh (1999, 2001) show that "Global/Macro" funds deliver "collar" like payoffs while "Trend Followers" exhibit a "look-back straddle" like payoff. Mitchell and Pulvino (2001) and Agarwal and Naik (2004) demonstrate that a number of equity-based hedge fund strategy payoffs resemble that obtained from writing an uncovered put option on the equity market.

credit spread. More importantly, a significant part of the variation in hedge fund returns over time can be explained by systematic risk factors. These insights enable us to segregate hedge fund returns into two components: one that can be explained by exposure to systematic risks (the market risk component), and other that cannot be explained by systematic risk factors (the manager specific component). The former represents the reward for bearing market risk, while the latter represents the reward attributable to manager skill.

Estimating the manager specific component, or alpha, has received considerable attention in the recent years, both in the popular press and in the financial economics literature.[3] In this paper, we use the insights from the pioneering work listed earlier, and estimate the non-systematic component of return on hedge funds. Specifically, we focus on two particular issues. Is the non-systematic component of return purely due to luck? If not, does it show persistence and is it possible to construct a trading rule that can capture it?

To answer these questions, one requires a methodology that fulfills at least three requirements. First, it must account for the fact that star funds are drawn from a large cross-section of hedge funds which increases the potential for some managers to do particularly well. Second, it should allow for the fact that hedge fund performance measures do not follow parametric normal distributions given the funds' dynamic trading strategies and holdings of options and derivatives securities. Third, it should be robust to possible misspecification of the factor model given that the complexity of hedge fund returns make their benchmarking more challenging than for mutual funds.

The bootstrap methodology presented by Kosowski et al (2004) satisfies these conditions. Therefore we follow their methodology and provide a comprehensive examination of hedge fund performance that explicitly controls for luck, while minimizing potential bias from mis-specification. In bootstrapping performance estimates, we explicitly model and control for the expected idiosyncratic variation in hedge fund returns. Specifically, we model the cross-sectional distribution of alpha estimates (across all funds) with the bootstrap, and then examine the significance of alpha

[3] See, for example, Treynor and Mazuy (1966), Jensen (1968), Treynor and Black (1972), Merton (1981), Henriksson and Merton (1981), Dybvig and Ingersoll (1982), Dybvig and Ross (1985), Admati et al (1986), Jagannathan and Korajczyk (1986), Connor and Korajczyk (1986), Lehmann and Modest (1987), Grinblatt and Titman (1989) and Glosten and Jagannathan (1994).

outliers. The rationale for the bootstrap methodology is threefold. First, evidence of non-normality in hedge fund returns and alpha estimates documented below proves that traditional parametric normal assumptions are inappropriate for more than 70% of hedge funds. Non-normal benchmark or individual security returns, co-skewed benchmark and security returns, holdings of options and derivatives with non-linear payoffs, dynamic factor loading strategies by managers, and time-series and cross-sectional correlations in the idiosyncratic return component of funds may all result in non-normal distributions of estimated alphas. Second, while the central limit theorem justifies regarding the normal distribution as a first-order approximation to the true distribution, the bootstrap can substantially improve on this approximation (see, for example, Bickel and Freedman (1984) and Hall (1986)). As emphasized by Horowitz (2003), the bootstrap has been shown in Monte Carlo experiments to spectacularly reduce differences between the true and nominal probabilities of correctly rejecting a given null hypothesis (in our case, that no superior hedge fund managers exist). Furthermore, given the difficulties in parametrically modelling the joint distribution of hedge fund performance across thousands of funds, most of which are very sparsely overlapping, the bootstrap offers a very attractive alternative approach.

Although we present several bootstrap results for the distribution of fund alphas, our tests also rely on bootstrapping the distribution of the t-statistic of the alphas, which has superior bootstrap properties, especially in the extreme tails of the performance distribution. Moreover, we present results for several extensions of the bootstrap methodology. In particular, we show that our results are not sensitive to the presence of an omitted factor or possible cross-sectional correlations in idiosyncratic returns among hedge funds. In addition, we carry out extensive robustness tests that take into account issues peculiar to hedge fund data like backfill bias (Ackermann, McEnally, and Ravenscraft (1999) and Posthuma and Van de Sluis (2003)) and short term serial correlation in returns (Getsmansky, Lo, and Makarov (2004)). Across all classes of measures, our bootstrap tests indicate that, controlling for sampling variability (luck), superior hedge funds that beat their benchmarks (net of expenses) by an economically and statistically significant amount do exist. We find this result both with our alpha bootstrap and our t-statistic of alpha bootstrap, as well as with the bootstrap extensions that we employ.

As a complement to the bootstrap approach, we also adopt a Bayesian framework to examine hedge fund performance. Our motivation is threefold. First, it has been shown that information from seemingly unrelated assets (henceforth SUR) can improve estimates of performance measures. See, for example, Stambaugh (1997) and Pastor and Stambaugh (2002) (henceforth PS). Second, hedge fund returns have shorter time series than mutual fund returns thus making the application of the Bayesian SUR approach even more relevant. Third, the SUR method provides additional robustness against the misspecification of the regression model that is being adopted, which is a concern, given that hedge fund strategies are more complex than those of mutual funds. Consistent with PS's results for mutual funds, we find that standard OLS measures of performance for hedge funds overestimate performance by 41% compared to the more accurate Bayesian estimates.

To test whether investors can take advantage of the abnormal performance of the top fund, we implement a simple routine to determine whether high-alpha funds persist. The results show that funds with high alphas tend to have high alphas in the future. This is true over evaluation and formation horizons of three years. This persistence not due to the imputation of fees. However, it is difficult for an investor to take advantage of this persistence as it is mostly driven by small funds who experience little inflows, suggesting that they are effectively closed to new investments. Our results are, in general, consistent with Berk and Green (2002), where any persistence is competed away by fund inflows (since managers may have decreasing returns to scale in their talents).

Our paper proceeds as follows. Section 2 describes the hedge fund data used in our study, while Section 3 presents the performance measures used in our bootstrapping procedure. Section 4 provides the empirical results based on the bootstrap and Bayesian SUR approaches. It includes bootstrap results broken down by investment category and the robustness checks on the bootstrap procedure. Section 5 presents a case study on the top hedge funds and persistence tests on fund alphas. We conclude the paper in Section 6.

2 Data

We evaluate the performance of hedge funds using monthly net-of-fee returns of live and dead hedge funds reported in the TASS, HFR, CISDM, and MSCI datasets over January 1990 to December 2002 period - a time period that covers both market upturns and downturns, as well as relatively calm and turbulent periods. In our fund universe, we have a total of 6,392 live hedge funds and 2,946 dead hedge funds. However, due to concerns that funds with assets under management below 20 million USD may be too small for many institutional investors, we exclude such funds from the analysis. This leaves us with a total of 4,300 live hedge funds and 1,233 dead hedge funds. The breakdown of funds by database is illustrated in Figure 1. The Venn diagram in Figure 1 reveals that the funds are roughly evenly split among TASS, HFR, and CISDM/MSCI.[4] While there are overlaps among the databases, there are many funds that belong to only one specific database. For example, there are 1,410 funds and 1,513 funds peculiar to the TASS and HFR databases respectively. This highlights the advantage of obtaining our funds from a variety of data vendors.

<Insert Figure 1 about here>

Although the term "hedge fund" originated from the equity long and short strategy employed by managers like Alfred Winslow Jones, the new definition of hedge funds covers a multitude of different strategies. Unlike the traditional investment arena, there does not exist a universally accepted norm to classify hedge funds into different strategy classes. We follow Agarwal, Daniel, and Naik (2004) and segregate them into five broad investment categories: Directional Traders, Relative Value, Security Selection, Multi-process, and Fund of Funds. Directional Trader funds usually bet on the direction of market prices of currencies, commodities, equities, and bonds in the futures and cash market. Relative Value funds take positions on spread relations between prices of financial assets and aim to minimize market exposure. Security Selection funds take long and short positions in undervalued and overvalued securities respectively and reduce systematic risks in the process. Usually they take positions in equity markets. Multi-process funds employ multiple strategies usually involving investments in opportunities created by significant transactional events,

[4]The CISDM and MSCI databases are combined in Figure 1 to facilitate illustration. A further breakdown of funds into each of the four databases is available upon request.

such as spin-offs, mergers and acquisitions, bankruptcy reorganizations, recapitalizations, and share buybacks. Fund of Funds invest in a pool of hedge funds and typically have lower minimum investment requirements. We also single out Equity Long/Short funds as their strategies are particularly easy to understand and their risks are easy to capture. Note that Equity Long/Short funds are a subset of Security Selection funds. As a prelude to the bootstrap analysis, we perform tests of normality, heteroskedasticity, and serial correlation to examine the behavior of fund returns in our sample broken down by investment category.

<Insert Table I about here>

The Jarque Bera test results reported in Table I, suggest that the majority of funds have returns that are not normally distributed. Over all classes of funds, about 70% of the funds fail the normality test, attesting to the need for non-parametric methods like the bootstrap when evaluating fund performance. Table I also reveals that fund returns are often serially correlated and fund residuals often heteroscedastic. Directional Traders fund returns appear to be most serially correlated while Multi-process fund residuals are most prone to heteroscedasticity.

3 Generalized asset class factor model

3.1 Factor benchmarks and performance measure α

In order to examine the value added by hedge funds, we regress the net-of-fee monthly excess return (in excess of the risk free rate) of a hedge fund on the excess returns earned by passive option-based, traditional buy-and-hold, and primitive trend following strategies. That is, we use as performance benchmarks the combined factors from the Agarwal and Naik (2004) and Fung and Hsieh (2001) models. The Agarwal and Naik (2004) factors are Russell 3000 index (RUS3000), Fama and French size (SMB), book-to-market (HML) and momentum (MOM) factors, MSCI excluding US index (MXUS), MSCI emerging markets index (MEM), Salomon bond index (SBG), Salomon world government index (SBW), Lehman high yield bond index (LHY), Fed trade weighted dollar index (FRBI), Goldman Sachs commodity index (GSCI), Moody's BAA rated corporate bond index (BAA), OTM call option index (SPCX), and OTM put option index (SPPX).[5] The Fung and Hsieh (2001) factors are S&P 500 return (SP), Wilshire small cap minus large cap return (SML), change in the constant maturity yield of the 10 year treasury (TSY), change in the spread of Moody's Baa minus 10 year treasury (HYMTSY), bond PTFS (PTFSBD), commodities PTFS (PTFSCOM), currency PTFS (PTFSFX), short term interest rate straddle PTFS (PTFSSIR), and stock PTFS (PTFSSTK), where PTFS denotes primitive trend following strategy. This represents the most comprehensive collection of hedge fund factors used in a study to date and minimizes the possibility of an omitted factor. Agarwal and Naik (2004) and Fung and Hsieh (2001) both show that their respective factor model strongly explains variation in individual hedge fund returns. To conserve degrees of freedom and to mitigate potential multi-collinearity problems, we use a stepwise regression approach where the independent variables are entered into the discriminant function one at a time, based on their discriminating power. The single best variable is chosen first; the initial variable is then paired with each of the other independent variables, one at a time, and a second variable with maximum incremental discriminating power is chosen, and so on. We use this procedure to ascertain the factors that, ex post, explain the returns earned by hedge funds during our sample period. We take the statistical significance of the factors, which are computed via Newey-West (1987) standard errors, as the measure of discriminating power.[6]

Since we use excess returns on selected options on index portfolios as additional "factor excess returns", the intercept ($\widehat{\alpha}^i$) from the regression below represents the value added by the manager of hedge fund i after controlling for her linear and non-linear risk exposures. In particular, to evaluate the performance of hedge funds we run the following regression[7]

[5] To avoid multi-collinearity problems, we omit the Agarwal and Naik (2004) ATM call option index and ATM put option index factors from our analysis.

[6] Our base bootstrap results remain unchanged when we use the full set of Agarwal and Naik (2004) factors or when we use the full set of Fung and Hsieh (2001) factors in place of the stepwise regression approach.

[7] See Glosten and Jagnnathan's (1994) section 2 for the theoretical underpinnings of equation (1). In our regression the value added or the "alpha" is given by the intercept while in their regression (see their section 4) it is given by the sum of the intercept, the coefficient on the index and the value of the call option.

$$r_t^i = \widehat{\alpha}^i + \sum_{k=1}^{K} \widehat{\beta}_k^i F_{k,t} + \widehat{\epsilon}_t^i \tag{1}$$

where,

r_t^i = net-of-fees excess return (in excess of the risk free rate of interest) on an individual hedge fund i for month t,

$\widehat{\alpha}^i$ = *alpha* performance measure - value added by a hedge fund i over the regression time period,

$\widehat{\beta}_k^i$ = average factor loading of an individual hedge fund i on kth factor during the regression period,

$F_{k,t}$ = excess return (in excess of the risk free rate of interest) on kth factor for month t, $(k = 1, ..., K)$ where the factor could be a Trading Strategy factor (an option-based strategy) or a Location factor (a long position in an index), and a Trend Following factor

$\widehat{\epsilon}_t^i$ = error term.

In the next section we apply the non-parametric bootstrap to test whether the returns of hedge funds can be explained by luck alone. We evaluate the performance (alpha and its t-statistic) of the hedge funds relative to the Agarwal and Naik (2004) and Fung and Hsieh (2001) factors and then bootstrap the residuals. We first do so on the entire sample of hedge funds. Subsequent tests are done on subperiods to gauge the effects of a structural break, and on subsamples of funds broken down by style category to ascertain the robustness of the basic results. The advantage of using the Agarwal and Naik (2004) model is that it includes option-based risk factors (SPCX and SPPX). These factors capture the left tail risk in hedge funds which are ignored by the commonly used mean-variance framework.

3.2 The bootstrap approach

The bootstrap is a nonparametric approach to statistical inference.[8] There are several advantages in using the bootstrap to evaluate hedge fund performance. First, traditional parametric methods use *a priori* assumptions about the shape of the distribution from which individual fund alphas are drawn. As Table I shows the empirical distribution of residuals from multi-factor performance regressions is highly non-normal for the hedge funds in our data. 70 percent of our hedge funds exhibit non-normalities compared to about 48 percent of the mutual funds in the Kosowski et al (2004) study. Thus, the distribution of $\widehat{\alpha}$ may be poorly approximated by normality and its statistical significance should be evaluated by means of a non-parametric approach such as the bootstrap.

Second, although the central limit theorem justifies regarding the normal distribution as a first-order approximation to the true distribution, the bootstrap can substantially improve on this approximation (see, for example, Bickel and Freedman (1984) and Hall (1986)). As emphasized by Horowitz (2003), the bootstrap has been shown in Monte Carlo experiments to spectacularly reduce differences between the true and nominal probabilities of correctly rejecting a given null hypothesis (in our case, that no superior fund managers exist).

Third, the use of the bootstrap approach to assess the performance of the best and worst fund managers is motivated by dependence of the distribution of funds' performance estimates on the entire covariance matrix characterizing the joint distribution of individual funds. Even for funds with residuals that are adequately approximated by a normal distribution, it is infeasible to apply standard statistical methods to assess the significance of extreme alphas drawn from a large universe of funds. In this case, the best alpha is the maximum value drawn from a multivariate distribution whose dimension depends on the number of funds in existence. The distribution of this maximum alpha depends on the entire covariance matrix for the joint distribution of the individual fund alphas or, more generally, its copula - which is generally impossible to estimate with precision. For example, besides the very large dimension of this matrix with several hundred or even thousands

[8] Our approach is based on the bootstrap introduced by Efron (1979). For a detailed discussion of the properties of the bootstrap, see, for example, Efron and Tibshirani (1993) or Hall (1992).

of funds, the entry and exit of funds imply that many funds do not even have overlapping return records with which to estimate covariances.

In addition, refinements of the bootstrap (which we will implement) provide a general approach for dealing with unknown time-series dependencies that are due, for example, to heteroskedasticity or serial correlation in the residuals from performance regressions. These bootstrap refinements also address the estimation of cross-sectional correlations in regression residuals, thus avoiding the estimation of a very large covariance matrix for these residuals. See Kosowski et al (2004) and the Appendix for further details on the bootstrap approach.

The hypothesis that the manager of the very best fund among L funds cannot produce a positive alpha is:

$$H_0 : \max_{i=1,\dots,L} \alpha_i \leq 0 \text{ , and}$$
$$H_A : \max_{i=1,\dots,L} \alpha_i > 0.$$

We also examine whether the funds in the left and right tail generate statistically significant alphas. To illustrate, suppose that a group of hedge funds have been ranked by their alphas, and let i^* be the rank of a given fund. When testing whether managers of the best ranked funds cannot generate positive alphas, the null and alternative hypotheses are

$$H_0 : \alpha_{i^*} \leq 0 \text{ , and}$$
$$H_A : \alpha_{i^*} > 0.$$

We also test whether managers of the worst-ranked funds cannot generate negative alphas. Here, the hypothesis test takes the following form for a fund with rank i^*:

$$H_0 : \alpha_{i^*} \geq 0 \text{ , and}$$
$$H_A : \alpha_{i^*} < 0.$$

Kosowski et al (2004) propose the estimated t-statistic of $\widehat{\alpha}$, $\widehat{t}_{\widehat{\alpha}}$, as a second measure of performance and justify its use based on its superior statistical properties. Although $\widehat{\alpha}$ measures the economic size of abnormal performance, it has a relatively high coverage error in construction of confidence intervals.[9] Also, $\widehat{t}_{\widehat{\alpha}}$ has another attractive statistical property. Specifically, funds with a shorter history of monthly net returns will have an alpha estimated with less precision, and will tend to generate alphas that are outliers. The t-statistic provides a correction for these spurious outliers by normalizing the estimated alpha by the estimated precision of the alpha estimate–it is related to the well-known "information ratio" method of performance measurement of Treynor and Black (1973). For these reasons, the our bootstrap tests will typically also include testing $\widehat{t}_{\widehat{\alpha}}$.[10]

Using this performance measure, the null and alternative hypotheses for the highest ranked fund are

$$\begin{aligned} H_0 &: \max_{i=1,\ldots,L} t_i \leq 0, \text{ and} \\ H_A &: \max_{i=1,\ldots,L} t_i > 0. \end{aligned}$$

For the lowest-ranked fund, the null and alternative hypotheses are given by reversing the inequalities above.

3.3 Bootstrap implementation: The baseline bootstrap procedure (residual-only resampling)

In this section, we illustrate our bootstrapping procedure for generating fund alphas with the multi-factor model in equation (1). The basic intuition for the bootstrap is the comparison of the actually observed top fund performance to the performance of top funds in artificially generated samples of data where funds' performance is controlled to be the result of sampling variability or luck. To

[9] The coverage probability is the probability that the confidence interval includes the true parameter, and the coverage error is the difference between the true and nominal coverage. Alternatively, $\widehat{t}_{\widehat{\alpha}}$ is a pivotal statistic, and, thus, generates lower coverage errors. A pivotal statistic is one that is not a function of nuisance parameters, such as $Var(\varepsilon_{it})$. For further details, see the Kosowski et al (2004).

[10] We estimate $\widehat{t}_{\widehat{\alpha}}$ using a heteroskedasticity and autocorrelation adjusted estimate of the standard error.

prepare for our bootstrap procedure, we use the multi-factor model in equation (1) to compute the OLS-estimated alphas, factor loadings, and residuals using the time-series of monthly net returns for fund i in the following equation:

$$r_t^i = \widehat{\alpha}^i + \sum_{k=1}^{K} \widehat{\beta}_k^i F_{k,t} + \widehat{\epsilon}_t^i \tag{2}$$

where r_{it}, $\widehat{\alpha}^i, \widehat{\beta}_k^i$ and $F_{k,t}$ are defined as above. For fund i, the coefficient estimates, $\{\widehat{\alpha}_i,\ \widehat{\beta}_i\}$, are saved, as well as the time-series of estimated residuals, $\{\widehat{\epsilon}_t^i\ ,\ t = 1, T_i\}$. In addition, the t-statistic of alpha, $\widehat{t_{\widehat{\alpha}}}$, is computed.

In a portfolio context, residual-only resampling is used to help control for non-normal security returns, which can induce non-normal portfolio returns.[11] In addition, co-skewness of individual securities and the market as well as dynamic factor loading strategies by hedge funds may result in non-normal distributions of estimated alphas.

For residual (only) resampling, we draw a sample with replacement from the fund i residuals that are saved in the first step, creating a time-series of resampled residuals, $\{\widehat{\epsilon}_{i,t}^b\ ,\ t = s_1^b, s_2^b, ..., s_{T_i}^b\}$, where b=1 (for bootstrap resample number one), and, as indicated, where a sample is drawn having the same number of residuals (e.g., the same number of time periods, T_i) as the original sample for each fund i. This resampling procedure is repeated for the remaining bootstrap iterations, $b = 2, ..., B$.

Next, for each bootstrap iteration, b, a time-series of (bootstrapped) monthly net returns is constructed for this fund, imposing the null hypothesis of zero true performance ($\alpha_i = 0$, or, equivalently, $\widehat{t_{\widehat{\alpha}}} = 0$):

$$\{r_{i,t}^b = \sum_{k=1}^{K} \widehat{\beta}_k^i F_{k,t} + \widehat{\epsilon}_{i,t}^b\ ,\ \ t = s_1^b, s_2^b, ..., s_{T_i}^b\}, \tag{3}$$

where $s_1^b, s_2^b, ..., s_{T_i}^b$ is the time reordering imposed by resampling the residuals in bootstrap iteration b. As indicated by Equation (3), this sequence of artificial returns has a true alpha (and t-statistic

[11] Co-skewness of individual security returns may not diversify away in large portfolios, especially if an omitted factor is responsible. In Table I, we provide evidence that the majority of funds have non-normal distributions of returns.

of alpha) of zero, since the residuals are drawn from a sample that is mean zero by construction. However, when we next regress the returns for a given bootstrap sample, b, on the multi-factor model, a positive estimated alpha (and t-statistic) may result, since that bootstrap may have drawn an abnormally high number of positive residuals, or, conversely, a negative alpha (and t-statistic) may result if an abnormally high number of negative residuals are drawn.

Repeating these steps across funds, $i = 1, ..., N$, and bootstrap iterations, $b = 1, ..., B$, we then build the cross-sectional distribution of the alpha estimates, $\widehat{\alpha}_i^b$, or their t-statistics, $\widehat{t}_{\widehat{\alpha}_i}^b$, resulting purely from sampling variation, as we impose the null of no abnormal performance. Bootstrapping the distribution of the maximum $\widehat{t}_{\widehat{\alpha}}$ proceeds similarly. If we find that very few of the bootstrap iterations generate as large values of $\widehat{\alpha}$ or $\widehat{t}_{\widehat{\alpha}}$, as those observed in the actual data, this suggests that sampling variation (luck) is not the source of performance, but that genuine stock picking/asset selection skills may exist. In all of our bootstrap tests, we set $B = 1,000$.

Our baseline bootstrap approach–"residual-only resampling"–resamples only the residuals saved above–this procedure is described in the next section. An extension to this approach–"residual and factor resampling"–independently resamples both the factor returns and the residuals, and is described in detail in Appendix A.

4 Empirical results

4.1 Basic bootstrap results

Table II Panel A displays the results from our application of the bootstrap algorithm to assess the statistical significance of hedge funds.[12] The funds are ranked in two different ways.. The first two rows of Panel A rank funds according to their estimated alphas. The third and fourth rows rank funds based on the estimated alpha t-statistics. The rationale for the ranking by $t-$statistics is based on its superior statistical properties (discussed in Section 3.2 above) and its interpretation as

[12] As we require sufficient return data to estimate the factor loadings, only funds with at least 30 months of return data are included in the bootstrap sample. As a robustness check, we also perform the bootstrap on funds with at least 48 months of return data and find qualitatively similar results. These results are available from the authors upon request.

a measure that penalizes alpha for its inherent noise (or standard error). The results are displayed for the extreme top 5 and bottom 5 funds as well as funds at the 1st, 3rd, 5th, and 10th percentiles on both ends of the alpha and alpha *t*-statistic spectrums.

<Insert Table II about here>

The results displayed in Table II Panel A suggest that whether we rank funds by abnormal performance (alpha) or alpha *t*-statistic, the abnormal returns of funds in the extreme deciles cannot be explained by luck alone. The top alpha fund registers a spectacular abnormal return of 7.35% per month while the top alpha *t*-statistic fund registers an impressive *t*-statistic of 42.23. The bootstrap p-values suggest that such abnormal returns are unlikely (probability < 0.01) the result of random chance. The same may be said of the other top five funds and the other funds in the top 90th percentile. While these funds have alphas and alpha *t*-statistics that are nowhere as impressive as the top fund, none of their alphas or alpha *t*-statistics can be explained by luck alone. A similar situation prevails with the funds in the left tail of the return distribution. However, here the bootstrapped p-values reveal that the fund alphas and alpha *t*-statistics are more noisy than those of the funds in the right tail. For the left tail funds displayed in Panel A, the hypothesis that the alphas of these funds are due to random chance cannot be rejected for two of the ten displayed funds at the 5% level of significance. Also, the hypothesis that the alpha *t*-statistics of these funds are the result of random chance cannot be rejected for two of the ten displayed funds at the 5% level of significance.

To further illustrate the baseline results of the bootstrap, we plot the kernel density estimate of the bootstrapped and actual alpha distributions for all funds in Figure 2. The two densities are rather different. The distribution of the alphas have much more mass on the tails than the bootstrapped alpha distribution. This is especially true of the right tail and suggests that we are likely to find funds in the right tail with superior alphas that cannot be explained by random resampling alone.

<Insert Figure 2 about here>

4.1.1 Adjusting for backfill bias

Some hedge fund studies make the case that the back fill bias artificially inflates the returns of hedge funds (Ackermann, McEnally, and Ravenscraft, 1999; Posthuma and Van de Sluis, 2003). This is because hedge funds often include returns for their entire history in the fund databases. This includes back dated returns for the period prior to a fund's listing on the database. Since hedge fund inclusion in databases is done on a voluntary basis, hedge funds with poor track records will have a disincentive to report to databases while hedge funds with good track records will have an incentive to report to databases. Posthuma and Van de Sluis (2003) claim that correcting for backfill bias in the TASS dataset reduces average returns by 4% per annum. In response to these concerns, we control for backfill bias in our bootstrap analysis in a similar fashion to Posthuma and Van de Sluis (2003). That is, we only include funds in the TASS dataset (since only TASS reports the date a fund first lists on the database) and only include a fund's return after it starts reporting to TASS. Hence we remove all the backfilled returns from our data sample. Then, we bootstrap the residuals of the funds from the backfill bias free subsample. The results displayed in Table II Panel B indicate that while the right tail alphas in the backfill bias free sample are uniformly smaller than the right tail alphas in the full sample, luck still cannot explain the over performance of these extreme hedge funds. The p-values of the alphas and alpha t-statistics of the right tail funds in Table II Panel B are all below 0.05.

4.1.2 Adjusting for serial correlation

Other hedge fund studies have documented that hedge fund returns are often highly serially correlated, in contrast to mutual fund returns, for example. According to Getsmansky, Lo, and Makarov (2004), the most likely reason for this serial correlation is funds' illiquidity exposure: hedge funds trade in securities which are not actively traded and whose market prices are not readily available. To remove the effect of artificial serial correlation induced by illiquidity exposure, we adopt the methodology pioneered in Okunev and White (2003) to smooth the hedge fund returns and eliminate serial correlation of up to order two. Then, we bootstrap the residuals from the smoothed sample. The results displayed in Table II Panel C are similar to those for the unsmoothed sample.

The alphas of all the right tail funds cannot be explained by chance. Also the p-values of the alpha t-statistic of all the right tail funds easily survive the adjustment for serial correlation. We note however that the alphas and alpha t-statistics of the funds in the right tail are much lower in the smoothed sample than in the unsmoothed sample. Nonetheless, the results in Panel C suggest that short term serial correlation cannot explain away the abnormal performance of the top funds. Unlike the funds in the right tail, some of the funds in the left tail of the alpha and t-statistic spectrums lose their statistical significance with the smoothing. In general the results for the funds in the left tail are not as robust as the results for the funds in the right tail. With the serial correlation adjustment, we see that the t-statistics are somewhat more robust than the alphas in the face of perturbations to the estimation procedure.

4.1.3 Adjusting for structural breaks

If the regression coefficients such as alpha and beta in hedge fund performance regressions exhibit structural breaks over time then constant coefficient regressions and bootstrap results are likely to be misspecified. Therefore we examine hedge fund returns for evidence of structural breaks.

Theoretically, structural instability can be expected for several reasons. First, changes in the parameters that relate hedge fund returns to state variables could arise from disappearing market inefficiencies, major changes in market sentiments, institutional changes, or large macroeconomic shocks that lead to changes in economic growth or affect risk premia (Paye and Timmermann (2003)).

Empirically there are reasons to expect instability in financial time series. Stock and Watson (1996) examine a large set of financial and macroeconomic time series, many of which are commonly used as state variables in financial models. Andreou and Ghysels (2002a,b) report evidence of breaks in the comovements of foreign exchange returns and the volatility dynamics of asset returns related to the Asian and Russian financial crises. Pastor and Stambaugh (2001) find evidence of structural breaks in the US equity premium in the context of a Bayesian framework.

Several recent papers have documented time-variation in the return characteristics of hedge funds (Agarwal, Daniel and Naik (2004) and Fung and Hsieh (2003)). There is evidence that hedge funds change their loadings on different risk factors over time and that inflows into funds affect subsequent returns. Anecdotal evidence suggests that hedge funds suffer from sudden shocks such as the Asian and Russian crisis which can be expected to lead to structural breaks in their return series.

We therefore wish to examine whether the hedge fund return series exhibit structural breaks and determine both the number of breaks, K, and estimate the time of their occurrence, $(T_1, T_2, ..., T_K)$. Bai and Perron (1998) provide a least-squares method for optimally determining the unknown breakpoints as well as the resulting size of shifts in parameters. In results available upon request, we apply the Bai and Perron (1998) test to HFR hedge fund return indices and find that most categories have a common structural break in December 2000.[13] See Appendix B for details of the Bai and Perron (1998) procedure. The structural break in December 2000 nicely coincides with the height of the technology bubble in the late 90s.

Based on the structural break evidence in December 2000, we repeat our bootstrap procedure allowing for a break in the beta slope coefficients using a dummy regression. Panel D of Table II reports the bootstrap results adjusted for the structural break. The top fund alphas in Panel D are less extreme than in Panel A of Table II, but they remain statistically significant. The bottom funds tend to lose their significance after adjusting for the structural break. To be fair, the sample of funds used in the structural break tests are much smaller than those used in the baseline tests since we require a minimum of 24 months of data before and after the break. Since loser funds tend to be short-lived, this further reduces the number of loser funds in the test sample.

Overall, the results in this section has shown that, the performance of the extreme top funds cannot be explained by luck alone. The top funds appear to be overperforming their benchmarks on a consistent basis. Backfill bias, serial correlation of returns, and structural breaks cannot explain our results. The bottom funds also appear to be underperforming their benchmarks but these results are not robust to adjustments for a structural break in December 2000.

[13] HFR hedge fund returns indices are regressed on factor benchmarks. Factor benchmarks are determined for each hedge fund index category by means of a step-wise regression approach described in Section 3 above. Both the intercept and the slope coefficients are allowed to vary.

4.2 Bootstrap results by fund category

In this section we group funds by broad investment category and "dead" funds by their reasons for leaving the databases, and then perform the bootstrap test on each subsample. We do so for two reasons. First, we wish to gauge the robustness of the results from the previous section. One possibility could be that overperformance in a particular style category is driving our results. If we find that most of the right tail funds belong to a particular investment style category then we are open to the critique that the Agarwal and Naik (2004) and Fung and Hsieh (2001) factors may not adequately capture the risks for that style. Such concerns will be moot if we find instead that right tail funds in all styles possess alphas and alpha t-statistics that cannot be explained by chance. Second, it will be interesting to compare the abnormal performance of funds who left the databases for different reasons. Such an exercise will have implications on survivorship bias related issues. Ackermann, McEnally, and Ravenscraft (1999) find that while terminated funds tend to perform worse than the average live fund, funds that stopped reporting tend to perform better than the average live fund. By excluding terminated and stopped reporting funds from the analysis, the average performance of the funds in the database may not be very different due to the countervailing forces exerted by the exclusion of these two groups of funds. Hence, it will be useful to characterize the left and right distributions of these groups of funds using the bootstrap and discuss the implications on survivorship bias.

To group funds by broad investment category, we use the five broad investment categories used in Agarwal, Daniel, and Naik (2004) and discussed in Section 2. We also single out Equity Long/Short funds as their strategies are particularly easy to understand and as they form the majority of the hedge funds in 2003. The results displayed in Table III suggests that we can find funds that consistently outperform or underperform their benchmarks in every investment category. Two other observations may be made from the bootstrap results in Table III. First, the best and worst funds in the sample are directional trader funds. These funds make aggressive bets on the direction of the market and it is no surprise that the top alpha, top alpha t-statistic, bottom alpha, and bottom alpha t-statistic funds are all directional trader funds. Second, the extreme positive alphas for fund of funds are smaller in magnitude than those for the other investment categories.

One interpretation is that this reflects the significant amount of transactions costs that these funds incur in moving money into and out of various hedge funds.

<Insert Table III about here>

To group funds that disappeared, we use the death indicator reported in the databases. The death indicator reveals whether a fund terminated, stopped reporting, or closed. The results of the bootstrap for the various funds grouped by death reason are reported in Table IV. Consistent with the results in Ackermann, McEnally, and Ravenscraft (1999), we find that funds that stopped reporting may have either performed well or performed poorly. The former group of funds might find that they have garnered so much inflows that they no longer need the marketing benefits from reporting to the databases. The latter group of funds may find that reporting their histories of poor returns hurt their inflows and hence opt to stop reporting. We also find that consistent with a priori intuition, the performance of the closed funds as a group seems very strong. While the closed funds in the right tail do not have spectacular alphas and alpha t-statistics as compared to some of the funds that stopped reporting, they are nonetheless significant according to their bootstrapped p-values. More importantly, more of the left tail closed funds have insignificantly negative alphas. Finally, we also find that there exist some funds with consistently positive alphas who terminate suggesting that hedge funds terminate for reasons other than poor performance.

<Insert Table IV about here>

4.3 Sensitivity analysis of the bootstrap procedure

The results of the previous sections have relied on the simple residual only resampling bootstrap. Given the complexity of hedge fund strategies, their return series may not fulfill the assumptions underlying the simple bootstrap approach. It remains to see whether our basic inferences change when we adopt more sophisticated bootstrap techniques that allow for deviations from the above assumptions. It is to this sensitivity analysis that we now turn.

In this section, we trace out the effects of four alternative bootstrap procedures on our basic results. They follow closely the bootstrap extensions used in Kosowski et al (2004). Each of these four procedures are motivated by different concerns regarding the appropriateness of the basic residual only based bootstrap. The first concern is that there is correlation between the factor returns and regression residuals. This correlation could occur if the hedge fund managers trade in securities that have a return co-skewness with the factor returns. Such correlation may also stem from managerial market- or factor-timing abilities that are not captured by the performance model. It is important to break any such correlation since the correlation may cloud our inferences on the significance of the alpha and alpha t-statistic. Hence, in addition to resampling the residuals, we also resample the factors. This is known as the factor-residual resampling bootstrap. The factor resampling is done so across all funds at the same time and is independent of the residual resampling. See Appendix A for further details on the factor-residual resampling bootstrap.

The second concern is that the residuals are cross-sectionally dependent. This could arise from funds holding the same securities or from funds following the same trading strategies. Hence we refine our residual-only bootstrap procedure to accommodate cross-sectional dependence in residuals. Rather than drawing sequences of time periods that are unique to each fund, for each bootstrap iteration, we draw T time periods from the set $\{t = 1, ..., T\}$, then resample the residuals from this reindexed time sequence across all funds, thereby preserving any cross-sectional correlation in the residuals. Since some funds may as a result be allocated bootstrap index entries from periods they did not exist, we omit a fund if it does not have at least 30 observations after applying the bootstrap index.

The third concern is that there may exist a persistent factor that is not accounted for by the Agarwal and Naik (2004) and Fung and Hsieh (2001) multi-factor models. If the reason why funds have high alphas is because they load on this omitted factor then we may be lead erroneously to conclude that fund managers have security selection ability if we do not include the omitted factor in our performance measurement model. This may be more of a concern for hedge funds than for mutual funds since hedge funds follow a wider range of trading strategies. One way to rectify this is to use a Monte Carlo simulation to simulate an omitted factor. The basic idea is that a persistent and common source of variation in fund residuals attributable to a hypothetical missing factor may be modelled as a slow mean-reverting process such as an AR(1) process. Hence we simulate such a factor and redo the residual resampling bootstrap on the performance model augmented with the simulated omitted factor. Details of the construction of the omitted factor are available in Appendix C.

The fourth and last concern is that the residual only resampling bootstrap requires the working assumption that conditional on factor realizations, the residuals are independently and identically distributed (henceforth I.I.D.). This may not seem like such as strong assumption given that it allows for conditional time series dependence of returns through the time series behavior of the factors, and that the simple bootstrap has some robustness properties that apply even if the I.I.D. assumption is violated (Hall (1992)). Nonetheless, to allow explicitly for the time series dependence in the residuals, we follow the stationary bootstrap procedure suggested by Politis and Romano (1994). Basically, the stationary bootstrap requires resampling of data blocks of random length to form pseudo time series. The data block length are generated from a geometric distribution whose mean is 1/q where q is a parameter specified by the econometrician. We choose q from the set {0.1, 0.5, 1} and repeat the stationary bootstrap procedure three times for blocks of average length = 10, 2, and 1. The results from these four extensions of the simple residual resampling bootstrap are presented in Table V.

<Insert Table V about here>

We see that the bootstrapped p-values for the extreme funds are not strongly affected by the different extensions. The p-values for the best and worst alpha appears to be more sensitive to variation in the bootstrap methodology than the p-values for the alpha *t*-statistics. Nonetheless, all the bootstrapped p-values for the best and worst alpha fund are still below 0.01. The results in Table V clearly suggest that our conclusion of significant security selection ability amongst hedge fund managers is robust to the choice of the bootstrap methodology.

4.4 An alternative Bayesian approach

Compared to return series on equity indices such as the S&P 500, the average hedge fund has a relatively short time series. This necessarily reduces the precision with which performance measures such as alpha can be estimated. However, as Pastor and Stambaugh (2002) (henceforth PS) point out, it is possible to substantially improve the alpha estimates by using historical returns on "seemingly-unrelated" assets not used in the definition of the alpha performance measure. These so-called non-benchmark passive assets have longer time series than the benchmark series and are correlated with hedge fund returns. The correlation between the hedge fund and non-benchmark passive returns can be exploited to improve our alpha estimates independent of whether these passive returns are priced by the benchmarks. PS use seemingly unrelated asset returns to improve the precision of alpha estimates of mutual funds.[14] Given that the average hedge fund has a much shorter returns series than the average mutual fund, the application of this methodology to hedge funds is even more relevant than to mutual funds. By using information on passive non-return benchmarks, we can in fact double the length of the time series used for estimating alphas.

Stambaugh (1997) shows how assets with longer historical time series provide information about the moments of assets with shorter histories. We follow the PS methodology and regress non-benchmark passive returns on benchmark returns. Let $F^N_{m,t}$ denote the $m \times 1$ vector of non-benchmark passive asset returns in month t on the k benchmark returns $F^B_{k,t}$:

$$F^N_t = \widehat{\alpha}^N + \sum_{k=1}^{K} \widehat{\beta}^N_k F^B_{k,t} + \widehat{\epsilon}^N_t \tag{4}$$

Importantly $\widehat{\epsilon}^N_t$ in Equation 1 and $\widehat{\epsilon}^N_t$ in Equation 4 are allowed to be correlated. PS show that the improvement in the estimation of alpha performance measure does not depend on whether the benchmarks $F^B_{k,t}$ perfectly price the non-benchmark passive assets $F^N_{m,t}$ or not. We also define the regression of a fund i's return on $p(= m + k)$ benchmark and non-benchmark assets:

$$R^i_t = \widehat{\delta}^i + \sum_{m=1}^{M} \widehat{c}'^{iN}_m F_{m,t} + \sum_{k=1}^{K} \widehat{c}'^{iB}_k F^B_{k,t} + \widehat{u}^i_t, \text{ for } i = 1, ..., L \tag{5}$$

Using Equation 1 and the assumption that $F^B_{k,t}$ is uncorrelated with both $\widehat{\epsilon}^N_t$ and $\widehat{u}^i_t$, PS show that

$$\alpha_i = \delta_i + c'_{iN}\alpha_N. \tag{6}$$

PS also show how to derive the posterior estimate $\widetilde{\alpha}_i$ of α_i in Equation 6 from the posterior moments of δ_i, c'_{iN}, and α_N. PS provide analytical expressions for the posterior moments $\widetilde{\alpha}_i$, $\widetilde{\delta}_i$, $\widetilde{c}'_{iN}$, and

[14] They find a median difference between their Bayesian posterior alphas and the OLS alphas of 2.3% per annum for all funds and of 8.1% for small-company growth funds.

$\widetilde{\alpha}_N$. As we show below, the posterior alpha estimate $\widetilde{\alpha}_i$ is below the OLS alpha $\widehat{\alpha}^i$ from equation 1 for 19 out of the top 20 Equity Long/Short hedge funds.

In the application of the Bayesian framework, several choices regarding the choice of fund returns, benchmark series, and non-benchmark series need to be made. We follow PS and apply an "empirical" Bayesian approach to estimate the prior distribution of various variables. The prior distribution of the covariance matrix of $\widehat{\epsilon}_t^N$ in equation 4. denoted by Σ is specified as an inverted Wishart distribution,

$$\Sigma^{-1} \sim W(H^{-1}, \nu).$$

We follow PS in specifying the empirical estimates of the priors. We set the degrees of freedom $\nu = m + 3$ which implies that the prior contains very little information about Σ. Moreover, we specify $H = s^2(\nu - m - 1)I_m$, and $E(\Sigma) = s^2 I_m$. The value of s^2 is set equal to the average of the diagonal elements of the sample estimates of Σ obtained using OLS regressions in equation 4. The parameters in equation 5 are specified as follows. The prior for σ_u^2, the variance of $u_{i,t}$, is an inverted gamma distribution or

$$\sigma_u^2 \sim \frac{\nu_0 s_0^2}{\chi_{\nu_0}^2}, \tag{7}$$

where $\chi_{\nu_0}^2$ represents a chi-square variate with ν_0 degrees of freedom. We define $c_L = (c_{LN}' c_{LB}')'$. Conditional on σ_u^2, the priors for δ_L and c_L are set to be normal distributions, independent of each other:

$$\delta_L | \sigma_u^2 \sim N\left(\delta_0, \left(\frac{\sigma_u^2}{E\left(\sigma_u^2\right)}\right)\sigma_\delta^2\right), \tag{8}$$

and

$$c_L | \sigma_u^2 \sim N\left(c_0, \left(\frac{\sigma_u^2}{E\left(\sigma_u^2\right)}\right)\Phi_c\right). \tag{9}$$

where σ_δ^2 represents the marginal prior variance of δ_L and Φ_c , the marginal prior covariance of c_L. We set $\sigma_\delta^2 = \infty$ (implemented by using a very large value computationally). As PS point out, this implies that the prior for α_A is diffuse and the prior mean of δ_0 does not play a role. We specify the values for s_0, ν_0, c_0, and Φ_c in equations 7, 8, and 9 using an empirical Bayes approach explained in detail in PS. The intution for this approach is that for any particular fund i in question, we determine the prior distributions of the parameters of interest by examining funds j that are similar to fund i. PS define funds to be similar when they have the same investment objective. However, since our hedge funds have different benchmarks and the Equity Long/Short funds that we are interested in belong to the same investment objective we require an alternative definition of the set of similar funds. Thus we define funds $j = 1, ..., J$ as similar when they share the same benchmark factor $F^B_{k,t}$ from the step-wise regression which chooses a fund's benchmarks $F^B_{k,t}$.[15] We use cross-sectional averages from the regression results for funds $j = 1, ..., J$ to define the prior uncertainty about a parameter for fund i. The prior mean and covariance matrix of c_L are denoted by c_0 and Φ_c respectively.

Following PS, all of our estimates are based on diffuse or completely non-informative priors.[16] As non-benchmark passive assets we use the following four time-series: the returns on the top three excluded factors from the stepwise regression for each fund in addition to the HFR Long/Short Equity index. The limited number of benchmark factors is motivated by the observation of PS that if the number of non-benchmark assets increases without a sufficient increase in R^2 then the posterior alpha estimate may be less precise. We follow PS's suggestion that a different set of non-benchmark assets could be used for each fund i. As PS point out, the non-benchmark passive factor should be highly correlated with fund returns, a condition fulfilled by the HFR Long/Short Equity index.

<Insert Table VI about here>

Table VI reports the estimates of posterior alpha estimate $\widetilde{\alpha}_i$ and the OLS alpha $\widehat{\alpha}^i$ (from equation 1). The first column reports the OLS alpha estimate. Columns two to four report the posterior alpha estimate (from the seemingly unrelated assets approach) for different values of the prior of σ_{α_N}. Columns five to seven report the differences between the OLS alpha and the posterior

[15] We require similar funds j to share the first benchmark factor $F^B_{k,t}$ with fund i if fund i has only one ($k = 1$) benchmark factor and to share the first two benchmark factors if fund i has more than one benchmark factor ($k > 1$).

[16] As PS point out the Bayesian framework can also accomodate informative beliefs.

alpha. Column eight reports the number of monthly observations for each fund and columns nine and ten report the start and end dates of the fund. As can be seen the alpha is reduced in 19 out of 20 cases and the median of the percentage of the alpha reduction is about 41% when compared to the OLS alpha. This shows that using longer time-series provides important information and that the short-sample period OLS alphas are overestimating actual performance. Although there is no evidence refuting the statistical significance of the top fund alphas from the bootstrap analysis, the Bayesian results suggest that caution must be exercised when evaluating those alphas.

5 Taking advantage of the abnormal performance of hedge funds

5.1 A case study: The top twenty Equity Long/Short funds

Thus far we have presented evidence which suggest that the performance of the top hedge funds cannot be attributed to luck. These funds appear to be able to deliver superior performance on a consistent basis. In this section, we take a closer look at the top twenty funds based on alpha. Our ultimate aim is to see whether one can form an investment strategy that takes advantage of the funds' superior returns. If the top funds are small, closed to new investments, and their high alphas driven by a few stellar return data points, then it does not bode well for potential investors of hedge funds hoping to profit from the documented abnormal performance. For the case study, we confine our analysis to Equity Long/Short funds as these funds form the largest group of funds and their returns are most easily captured by the factor models of Agarwal and Naik (2004) and Fung and Hsieh (2001).

The summary statistics for the top twenty equity long short funds are showcased on Table VII. We report the following for each fund: alpha, bootstrap $p-$value, R^2, mean return, standard deviation of returns, maximum monthly return, minimum monthly return, adjusted alpha/alpha, fund factors selected using the stepwise regression algorithm, fund status, start assets under management (henceforth AUM), end AUM, and mean fund inflow. Adjusted alpha is the alpha of the fund after removing the top three return observations. Fund status indicates whether a fund is still "live" in the databases, or whether it has terminated, stopped reporting, or closed. Because a fund may be effectively closed (i.e., it no long accepts any new inflows) but still reporting to the databases, we examine the start and end AUMs to gauge the size of the funds as well as analyze the mean fund inflow to ascertain whether a fund has effectively stopped accepting new inflows

<Insert Table VII about here>

Three salient observations may be made from the summary statistics in Table VII. First, fund alpha is driven mostly by the extreme return observations. On average, the top three return observations contribute close to 40% of fund alpha. There are six funds for whom adjusted alpha makes up less than 50% of alpha. In fact, for the top fourteenth fund, the adjusted alpha is negative. The low adjusted alpha/alpha numbers are corroborated by the high standard deviations and the large extreme return values reported in Table VII. Second, some of the funds have stopped reporting to the databases. Of the twenty funds, six funds have stopped reporting while two funds have been terminated. Third, the inflows for these funds are very small. The median monthly inflow is about 4% for these funds. If we took funds with inflow below 2% of AUM as effectively closed, then there are nine funds which are effectively closed. Further, seven of these nine funds are still reporting to the databases. The ending assets under management column of Table VII reveals that many of these funds remained small (AUM $<$ 100 million) over their sample period despite their stellar returns, further suggesting that many of them are effectively closed. As a robustness check, we also examine the average pairwise alpha correlations between the top funds. If there was evidence that their rolling alphas are correlated and funds outperform at the same time, this might point to a missing factor in the regression model. The average pairwise correlation between the top funds' rolling alphas is 0.07. This, in addition to our omitted factor robustness checks above, provides further evidence against the possibility of a missing factor. Overall, the results in Table VII suggests that it will not be straight forward to take advantage of the superior returns of the top funds as a large part of their alpha is driven by extreme data points. Even if one can identify the good funds early on these funds are small and may no longer accept new fund inflows making any investment strategy difficult.

5.2 Persistence tests of fund alpha

To complete our analysis of whether investors can profit from the observed abnormal performance, it remains to test whether the alphas of hedge funds persist. The bootstrap results suggest that managers with superior alphas are not just lucky. This implies that these managers possess a certain element of skill, which in turn implies that fund alphas should persist at some horizon, assuming that managerial skill persists and that managerial tenure is long enough. Yet, the results from the case study in the previous subsection suggest that those superior alphas are driven by a few spectacular return data points (which may not persist) and belong to funds that are typically small and potentially closed to new inflows.

Performance persistence is an age old topic in the mutual fund literature. Hendricks, Patel, and Zeckhauser (1993) among others show that mutual fund returns persist in the medium term (one to three years). However Carhart (1997) shows that this is either due to managers adopting momentum strategies or to the persistence of fund expenses. The latter result stems from the fact that this year's winners tend to be winners because they charge low expenses. Conversely, this year's losers tend to be losers because they charge high expenses. However, newer studies like Kosowski et al (2004), and Mamaysky, Spiegel, and Zhang (2004) show that more sophisticated econometric methods allow one to pick out funds whose returns cannot be explained by four factor covariation or expense ratios. With hedge funds though, previous studies have found little evidence of persistence in returns. Agarwal and Naik (2000) show that hedge fund returns only persist in the short term (one to three months). Getsmansky, Lo and Makarov (2004) credit that to illiquidity induced by the assets that hedge funds trade. Like Agarwal and Naik (2000), Brown, Goetzmann, and Ibbotson (1999) find no evidence of persistence in hedge fund returns at annual horizons. What of hedge fund alphas?

To test the persistence properties of hedge fund alphas, we first estimate individual fund rolling alphas using the stepwise regression approach and the factor models of Agarwal and Naik (2004) and Fung and Hsieh (2001). That is, we estimate individual fund alphas from past 36 months of return data (or a minimum of 30 months of return data) in the spirit of Carhart (1997). Then, we sort the funds into decile portfolios each month based on the alphas calculated over the past

36 months and evaluate average fund alpha estimated over the next 36 months for each of the portfolios. The spread between the mean fund alpha of decile 1 (highest alpha funds) and the mean fund alpha of decile 10 (lowest alpha funds) is presented in the first column and top two rows of Table VIII. Since we use overlapping three year periods to estimate the formation and evaluation period alphas (to maximize data), we compute the t-statistics from Newey-West (1987) standard errors to correct for the possible serial correlation.[17] We analyze average decile fund alpha as opposed to decile portfolio alpha in the evaluation period to allow the factors and factor loadings to vary across the diverse set of hedge funds within each decile portfolio.

<Insert Table VIII about here>

We find, from the results in Table VIII, that the alpha differences persist over the formation and evaluation period of three years. The alpha spread between the best alpha decile and the worst alpha decile is a modest 0.22% per month or 2.7% per year. Moreover, this alpha spread is statistically significant at the 1% level. Can investors take advantage of this persistence? We tackle this question in three parts. First, we examine the persistence of the value-weighted fund alpha spread. Instead of taking the average fund alpha, we weight fund alpha by fund assets under management (AUM). Table VIII reports that the value-weighted alpha spread is insignificant from zero suggesting that small funds drive the bulk of the alpha persistence. Second, we split the sample into small and large funds based on AUM at the start of the evaluation period. Small funds are funds with AUM less than the median AUM. Large funds are funds with AUM greater than the median AUM. Then we redo the persistence test on small and large funds separately. The results in Table VIII again suggest that small funds are responsible for much of the performance persistence. Third, we split the sample into closed and open funds based on mean fund inflow over the evaluation period. Closed funds are funds with mean inflow less than the median inflow over the evaluation period. Open funds are funds with mean inflow greater than the median inflow. We then redo the persistence analysis on each subset of funds. The results strongly indicate that the persistence is concentrated in funds that experience very little inflows and who may be effectively

[17] Note that there is no overlap between formation and evaluation period but only overlap between successive evaluation periods and successive formation periods.

closed. For those funds, the alpha spread is an economically and statistically significant 4.35% per year which is more than 2.5 times the alpha spread of 1.66% per year generated with open funds. Overall, the persistence tests suggest that because abnormal performance / persistence is driven mostly by small and effectively closed funds, it will not be easy for investors to take advantage of the abnormal performance and performance persistence. This dovetails with the insights obtained from the case study in the previous subsection.

One concern with the persistence tests thus far may be that the imputation of fees is clouding the analysis. It could be that the funds in the top decile persist because they charge low incentive and managerial fees, while the funds in the bottom decile continue to do poorly post-fee wise because they charge high incentive and managerial fees. In response to these concerns, we add back the incentive and managerial fees to obtain pre-fee returns, calculate the pre-fee alphas, and redo the sorts. The results with these pre-fee alphas in the bottom two rows of Table VIII are qualitatively similar to those with the post-fee alphas.

6 Conclusion

The results in this study has done much to forward the view that managers are responsible for the superior performance of some hedge funds. Using a non-parametric bootstrap methodology, we find that extreme fund alphas cannot be explained by luck alone. This is more the case for extreme right tail funds than for extreme left tail funds. That is, it is easier to attribute poor performance to a stroke of bad fortune. Our bootstrap methodology features numerous extensions, including extensions for an omitted factor, for cross-sectional dependence, and for independent factor resampling. We obtain similar results despite implementing controls for back fill bias, serial correlation, and structural breaks. It is therefore not surprising that we find persistence in fund alpha over horizons of three years.

However, two caveats are in order. First, the performance persistence is driven mostly by small funds who experience very little inflows and hence may be closed to new investments. Hence any investment strategy designed to take advantage of the persistence will encounter difficulties.

Second, Bayesian tests suggest that the average abnormal performance of the top funds may be overestimated by 41% with the conventional OLS approach.

7 Appendix

7.1 Appendix A: Bootstrap extension - Residual and factor resampling

Kosowski et al (2004) argue that independent resampling of regression residuals and factor returns breaks any correlation between these two components. Such a correlation would occur, for example, if a fund manager holds stocks having a return co-skewness with the market, or with other factor returns. Co-skewness may also occur if the manager has market- or factor-timing abilities that are not properly specified in the performance model. Breaking any such correlation may significantly change inferences about the significance of alpha, thus, following Kosowski et al (2004) we employ independent residual and factor resampling as an alternative approach.

For residual and factor resampling, we augment the residual resampling procedure with factor returns that are resampled independently of the residuals. When resampling these factor returns, the same draw is used across all funds, giving the following data for bootstrap iteration b for fund i:

$$\{\sum_{k=1}^{K}\widehat{\beta}_k^i F_{k,t}^b,\ t=\tau_1^b,\tau_2^b,...,\tau_{T_i}^b\} \text{ and } \{\widehat{\epsilon}_{i,t}^b,\ t=s_1^b,s_2^b,...,s_{T_i}^b\}$$

Resampling factor returns as well as residuals allows for sampling variation in the coefficient estimates, $\{\sum_{k=1}^{K}\widehat{\beta}_k^i\}$, that results from using a particular draw of factor realizations, as well as residuals, over the sample period.

Next, for each bootstrap iteration, b, a time-series of (bootstrapped) monthly net returns is constructed for fund i, again imposing the null hypothesis of zero true performance ($\alpha_i = 0$):

$$\begin{aligned} \{R_{i,t}^b &= \sum_{k=1}^{K}\widehat{\beta}_k^i F_{k,t}^b + \widehat{\epsilon}_{i,t_\epsilon}^b\ ,\ t_F=\tau_1^b,\tau_2^b,...,\tau_{T_i}^b \\ \text{and } t_\epsilon &= s_1^b,s_2^b,...,s_{T_i}^b\}, \end{aligned}$$

where $\tau_1^b, \tau_2^b, ..., \tau_{T_i}^b$ and $s_1^b, s_2^b, ..., s_{T_i}^b$ are the (matched) time reorderings imposed by resampling the factor returns and residuals, respectively, in bootstrap iteration b.

Repeating these steps across funds, $i = 1, ..., L$, and bootstrap iterations, $b = 1, ..., B$, we then build the cross-sectional distribution of the alpha estimates, $\widehat{\alpha}_i^b$, or their t-statistics, $\widehat{t}_{\widehat{\alpha}_i}^b$, resulting purely from sampling variation, as we impose the null of no abnormal performance. Bootstrapping the distribution of the maximum $\widehat{t_{\widehat{\alpha}}}$ proceeds similarly. If we find that very few of the bootstrap iterations generate a maximum $\widehat{\alpha}$, or $\widehat{t_{\widehat{\alpha}}}$, as high as that observed in the actual (unmodified) data, this suggests that sampling variation (luck) is not the source of performance, but that genuine skills actually exist. In all of our bootstrap tests to follow, we execute 1,000 bootstrap iterations ($B = 1,000$).

7.2 Appendix B: Structural break test

Bai and Perron (1998) provide a least-squares method for optimally determining the unknown breakpoints as well as the resulting size of shifts in parameters. Their approach is based on searching over the possible $K-$partitions $(T_1, T_2, ..., T_K)$ of the data to compute the minimizer of the sume of squared residuals. For a set of K breakpoints, $(T_1, T_2, ..., T_K)$ =$\{T_j\}$, the coefficients $\beta_{k,\{T_j\}}$ are chosen to minimize the sum of squared residuals

$$S_T(\{T_j\}) = \sum_{k=1}^{K+1} \sum_{T=T_{k-1}+1}^{T_k} (r_t - \widehat{\beta}'_{k,\{T_j\}} F_{k,t})^2$$

The resulting break dates$(\widehat{T}_1, \widehat{T}_2, ..., \widehat{T}_K)$ are selected so as to satisfy

$$(\widehat{T}_1, \widehat{T}_2, ..., \widehat{T}_K) = \arg \min_{T_2, T_2, .., T_K} S_T(T_1, ..., T_K),$$

where the minimization is over all partitions such as $T_k - T_{k-1} \geq \pi \times T$. The trimming percentage parameter π imposes a minimum length for the time between breaks, $\pi \times T$.Choosing π in practice involves a trade-off between the ability to detect regimes of relatively short length and the desire to avoid overfitting the data and simply identifying 'outliers'. Bai and Perron (2000) discuss computational and practical aspects of determining these design parameters.

In results available from the authors upon request we examine structural break tests for different trimming percentages. In the presence of autocorrelation and heteroskedasticity Bai and Perron (1998) advocate the use of a trimming percentage of 15%. As the diagnostics in Table 1 show there is overwhelming evidence of serial correlation and heteroskedasticity in hedge fund returns. We therefore choose the trimming percentage of 15% as our preferred mode.

Several recent papers have contributed to the structural break literature but their application is beyond the scope of this paper. These papers include Andrews (2003) which addresses the issue of identifying breaks at the end of the sample period, and Elliot and Mueller (2004), which contribute to the determination of confidence intervals for structural breaks. Hansen (2000) derives the large sample distributions of several test statistics for structural breaks allowing for structural change in the marginal distribution of the regressors and proposes a 'fixed regressor bootstrap' for determining the critical values under the null hypothesis. However this test does not permit multiple breaks in the regression coefficients.

7.3 Appendix C: Monte carlo simulation of persistent sources of variation in fund residuals

Kosowski et al (2004) modify the baseline bootstrap procedure to account for persistent sources of variation in fund residuals. Their procedure is based on the fact that a persistent and common source of variation in fund residuals driven by a hypothetical missing factor can be modelled as a slowly mean-reverting process such as an AR(1) process. Their bootstrap variation is based on a Monte Carlo simulation of the effect that such a factor would have if it was present in the residuals as a result of being omitted from the factor model.

Kosowski et al (2004) show that the presence of an omitted factor OF_t in residuals in a multi-factor model (with 4 factors, for example) can be modelled as follows

$$r_{it} = \alpha_i + \beta_{1i} f_{1t} + ... + \beta_{3i} f_{3t} + \beta_{4i} f_{4t} + \beta_{5i} OF_t + \varepsilon_{it}, \tag{10}$$

so that the residuals are the sum of the common omitted factor, OF_t, and the fund-specific (idiosyncratic) shocks ε_{it} :

$$u_{it} = \beta_{5i} OF_t + \varepsilon_{it}. \tag{11}$$

A missing factor such as an oil factor would have to be persistent to affect funds differently over time. Kosowski et al (2004) show that its dynamics can be captured through the following AR(1) process:

$$OF_t = \rho \times OF_{t-1} + e_t, \tag{12}$$

where ρ is the persistence parameter. This would make the omitted factor slowly mean-reverting. Funds with positive exposure to OF_t ($\beta_{5i} > 0$) operating during a period where its mean was positive would see their $\alpha-$estimates artificially increased, while funds with a positive exposure to OF_t operating at times when OF_t is negative would see a negative contribution to their $\alpha-$estimates. We following the Monte Carlo procedure proposed by Kosowski et al (2004) to show the robustness of our results to potential persistent source of variation in fund residuals. For details see the appendix of Kosowski et al (2004).

8 References

Ackermann, C., McEnally, R., Ravenscraft, D., 1999. The performance of hedge funds: risk, return, and incentive. Journal of Finance 54, 833-874.

Admati, A.R., Bhattacharya, S., Pfleiderer, P., Stephen, R., 1986. On timing and selectivity. Journal of Finance 41, 715-730.

Agarwal, V., Naik, N., 2000. Multi-period performance persistence analysis of hedge funds. Journal of Financial and Quantitative Analysis 35, 327-342.

Agarwal, V., Daniel, N., Naik, N., 2004. Flows, performance, and managerial incentives in the hedge fund industry. Unpublished working paper. London Business School.

Agarwal, V., Naik, N., 2004. Risk and portfolio decisions involving hedge funds. Review of Financial Studies 17, 63-98.

Andreou, E., Ghysels, E., 2002a. Tests for breaks in the conditional co-movements of asset returns. Statistica Sinica, forthcoming.

Andreou, E., Ghysels, E., 2002b. Detecting multiple breaks in financial market volatility dynamics. Journal of Applied Econometrics 17, 579-600.

Andrews, D.W.K., Kim, J.Y., 2003. End-of-sample cointegration breakdown tests. Cowles Foundation Discussion Papers 1404, Cowles Foundation, Yale University.

Bai, J., Perron, P., 1998. Estimating and testing linear models with multiple structural changes. Econometrica 66, 47-78.

Bai, J., Perron, P., 2000. Computation and analysis of multiple structural change models. Unpublished working paper, Boston University.

Berk, J., Green, R., 2002. Mutual fund flows and performance in rational markets. NBER working paper, No 9275.

Bickel, P.J., Freedman, D.A., 1984. Some asymptotics on the bootstrap. Annals of Statistics 9, 1196-1271.

Brown, S., Goetzmann, W., Ibbotson, R., 1999. Offshore hedge funds: Survival and performance. Journal of Business 72, 91-117.

Carhart, M., 1997. On persistence of mutual fund performance. Journal of Finance 52, 57-82.

Connor, G., Korajczyk, R.A., 1986. Performance measurement with the arbitrage pricing theory: A new framework for analysis. Journal of Financial Economics 15, 373-394.

Cohen, R., Coval, J., Pastor, L., 2002. Judging fund managers by the company they keep. NBER working paper series no.9359;

Dybvig, P.H., Ingersoll, J.E.Jr., 1982. Mean-variance theory in complete markets. Journal of Business 55, 233-251.

Dybvig, P.H., Ross, S., 1985. Differential information and performance measurement using a security market line. Journal of Finance 40, 383-399.

Elliott, G., Muller, U., 2002. Optimally testing general breaking processes in linear time series models. Working paper, Department of Economics, University of California, San Diego.

Fung, W., Hsieh, D.A., 1999. A primer on hedge funds. Journal of Empirical Finance 6. 309-331.

Fung, W., Hsieh, D.A., 2000. Performance characteristics of hedge funds and CTA funds: Natural versus spurious biases. Journal of Financial and Quantitative Analysis 35, 291-307.

Fung, W., Hsieh, D.A., 2001. The risk in hedge fund strategies: Theory and evidence from trend followers. Review of Financial Studies 14, 313-341.

Fung, W., Hsieh, D.A., 2003. Hedge fund benchmarks: A risk based approach. Discussion Paper 014, Centre for Hedge Fund Education and Research, London Business School.

Getsmansky, M., Lo, A., Makarov, I., 2004. An econometric model of serial correlation and illiquidity of hedge fund returns. Journal of Financial Economics, forthcoming.

Glosten, L.R., Jagannathan, R., 1994. A contingent claim approach to performance evaluation. Journal of Empirical Finance 1, 133-160.

Grinblatt, M., Titman, S., 1989. Portfolio performance evaluation: Old issues and new insights. Review of Financial Studies 2, 393-421.

Hansen, B. 2000. Testing for structural change in conditional models. Journal of Econometrics 97, 93-115.

Hall, P., 1986. On the bootstrap and confidence intervals. Annals of Statistics 14, 1431-1452.

Hall, P., 1992. The bootstrap and the edgeworth expansion. Springer-Verlag, New York.

Hendricks, D., Patel, J., Zeckhauser, R., 1993. Hot hands in mutual funds: Short-run persistence of performance, 1974-88. Journal of Finance 48, 93-130.

Henriksson, R.D., Merton, R.C., 1981. On market timing and investment performance II: Statistical procedures for evaluating forecasting skills. Journal of Business 54, 513-533.

Horowitz, J.L, Härdle, W., Kreiss, J. P., 2003. Bootstrap methods for time series. International Statistical Review 71, 435-459.

Jagannathan, R., Korajczyk, R., 1986. Assessing the market timing performance of managed portfolios. Journal of Business 59, 217-235.

Jensen, M.C., 1968. The performance of mutual funds in the period 1945-1964. Journal of Finance 23, 389-416.

Kosowski, R., Timmermann, A., White, H., Wermers, R., 2004. Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis. Unpublished working paper. INSEAD.

Lehmann, B.N., Modest, D.M., 1987. Mutual fund performance evaluation: A comparison of benchmarks and benchmark comparisons. Journal of Finance 42, 233-265.

Mamaysky, H., Spiegel, M., Zhang, H., 2003. Estimating the dynamics of mutual fund alphas and betas. Yale ICF working paper No. 03-03. Yale School of Management.

Merton, R.C., 1981. On market timing and investment performance I: An equilibrium theory of value for market forecasts. Journal of Business 54, 363-406.

Mitchell, M., Pulvino, T., 2001. Characteristics of risk in risk arbitrage. Journal of Finance 56, 2135-2175.

Newey, W., West, K., 1987. A simple positive-definite heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703-708.

Okunev, J., White, D., 2003. Hedge fund risk factors and value at risk of credit trading strategies. Unpublished working paper. University of New South Wales.

Pastor, L., Stambaugh, R., 2002. Mutual fund performance and seemingly unrelated assets. Journal of Financial Economics 63, 315-349.

Politis, D.N., Romano, J.P., 1994. The stationary bootstrap. Journal of American Statistical Association 89, 1303-1313.

Posthuma, N., Van De Sluis, P.J., 2003. A reality check on hedge fund returns. Unpublished working paper. Free University Amsterdam.

Sharpe, W., 1992. Asset allocation: Management style and performance measurement. Journal of Portfolio Management 18, 7-19.

Stambaugh, R., 1997. Analyzing investments whose histories differ in length. Journal of Financial Economics 45, 581-622.

Treynor, J., Black, F., 1973. How to use security analysis to improve portfolio selection. Journal of Business 46, 66-86.

Treynor, J., Black, F., 1972. Portfolio selection using special information, under the assumptions of the diagonal model, with mean-variance portfolio objectives, and without constraints. in G.P. Szego and K. Shell, eds., Mathematical methods in investment and Finance, (North-Holland: Amsterdam), 367-384.

Treynor, J., Mazuy, K., 1966. Can mutual funds outguess the market? Harvard Business Review 44, 131-136.

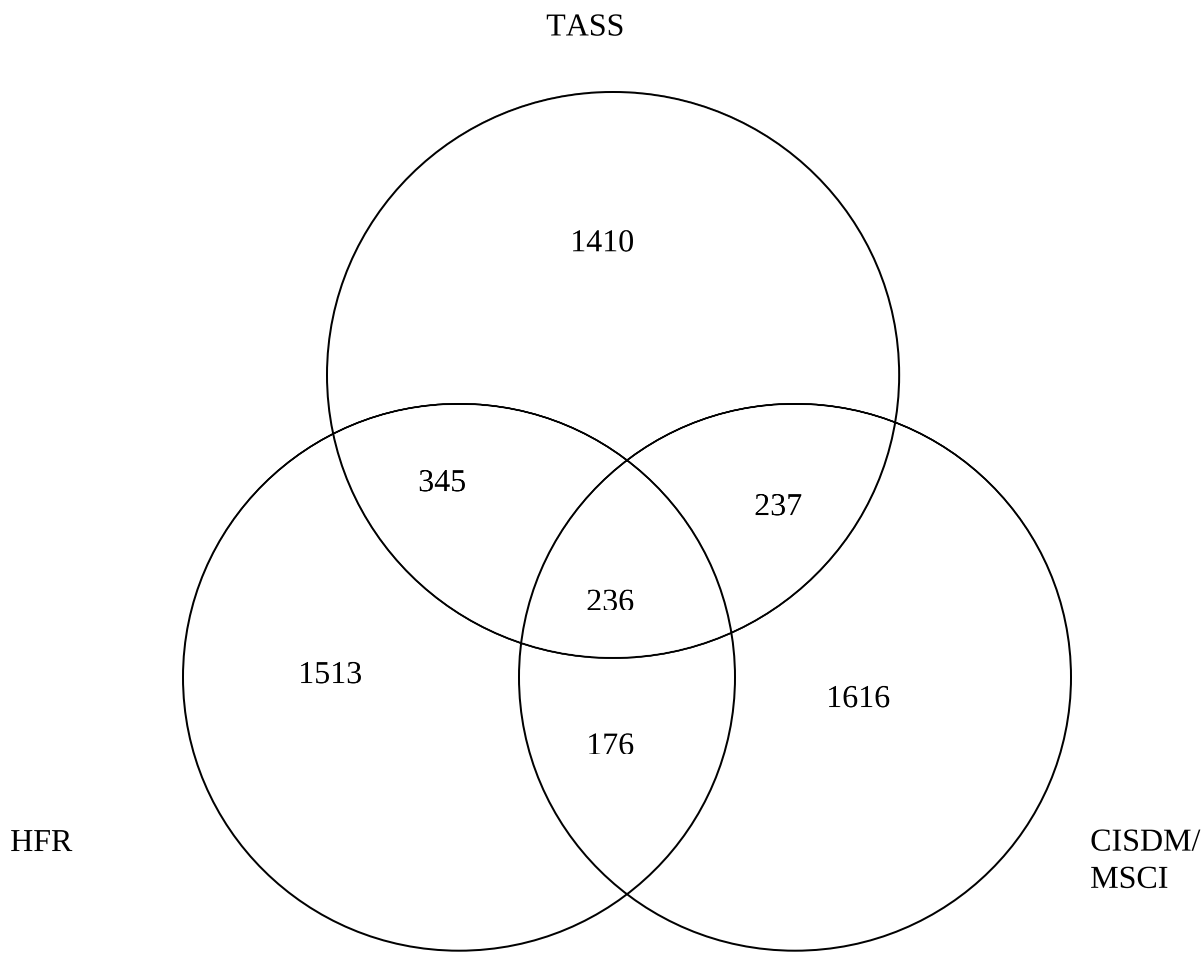

Figure 1: Composition of combined hedge fund database. The sample period is from January 1990 to December 2002. Funds with assets under management > US$20 million.

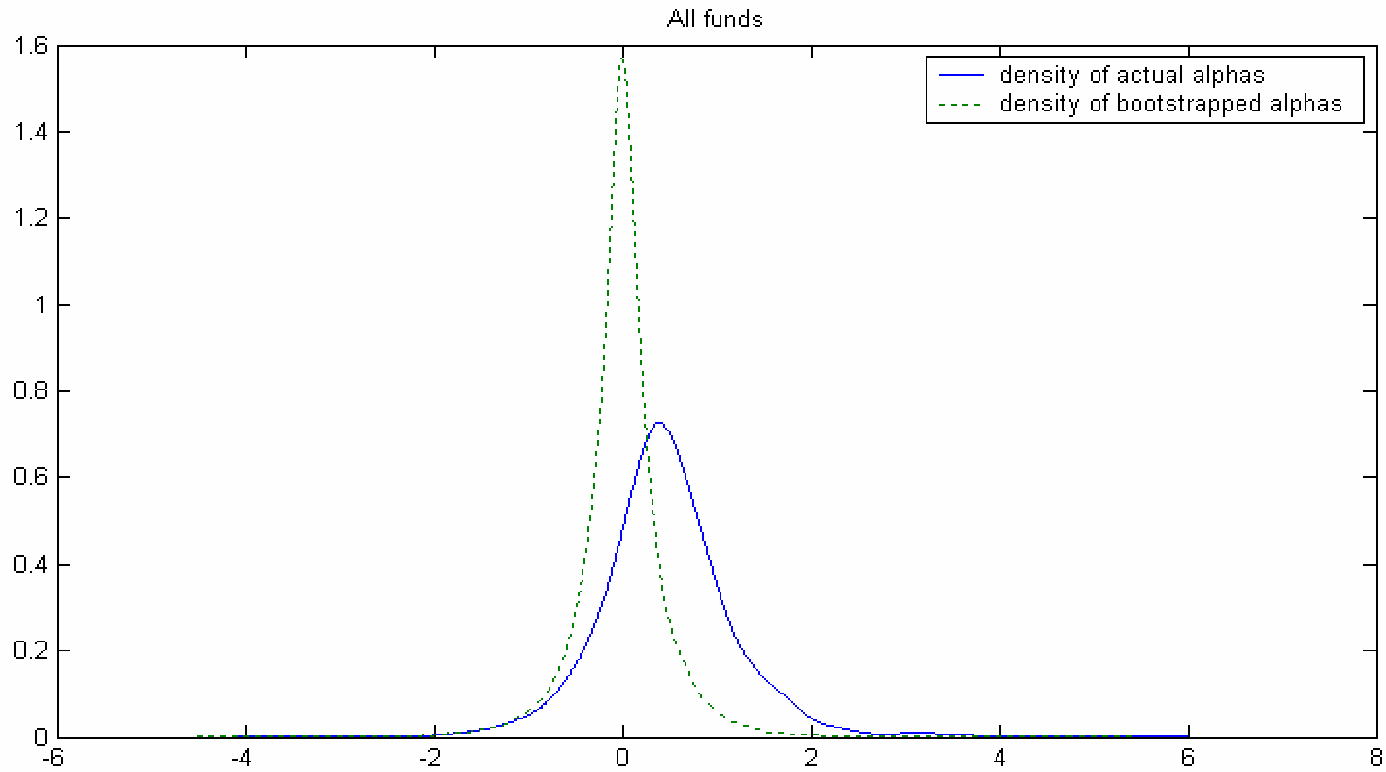

Figure 2. Kernel density estimate of the bootstrapped (dotted line) and the actual (solid line) alpha distributions for all funds. The sample period is from January 1990 to December 2002.

Table I

Summary Statistics and Tests of Normality, Heteroskedasticity, and Serial Correlation on Hedge Fund Residuals

The sample period is from January 1990 to December 2002. The distributional properties of hedge fund return residuals are reported via summary statistics on kurtosis and skewness. Tests of normality are conducted using the Jarque Bera test. Tests of heteroscedasticity are conducted with the Breusch Pagan test. And tests of serial correlation are conducted with the Ljung Box test. All tests are conducted on fund residuals. The percentage of funds with *p* -values which reject the null at the 10% confidence level is reported in each test. Fund residuals are obtained by regressing fund returns in excess of the risk free rate on the Agarwal and Naik (2004) option-based factors and the Fung and Hsieh (2001) factors.

	summary statistics		test of normality	test of heteroscedasticity	test of serial correlation (AR1)
	median kurtosis	median skewness	% with Jarque Bera *p*- value <0.1	% with Breusch Pagan *p*- value <0.1	% with Ljung Box *p*- value <0.1
Long/short Equity Funds	3.57	-0.97	75.86	34.48	37.93
Directional Trader Funds	3.54	-0.45	65.28	15.28	62.50
Multi-process Funds	3.91	0.13	63.64	36.36	27.27
Relative Value Funds	4.34	0.07	30.00	14.00	14.00
Security Selection Funds	3.57	-0.06	48.78	13.41	32.93
Fund of funds	3.91	-0.04	52.58	27.84	36.08

Table II

Statistical significance of the best and worst hedge funds' performance

The sample period is from January 1990 to December 2002. The first and second rows report the OLS estimate of alpha in percent per month and the bootstrapped *p* -value of alpha. Alpha is estimated relative to the Agarwal and Naik (2004) and Fung and Hsieh (2001) factors. The third and fourth rows report the *t* -statistic of alpha based on heteroskedasticity and autocorrelation consistent standard errors as well as the bootstrapped *p* -value of the *t* -statistic. The first column on the left (right) reports funds with the lowest (highest) alpha and *t* -statistic followed by the results for the funds with the second lowest (highest) alpha and *t* -statistic and marginal funds at different percentiles in the left (right) tail of the distribution. The *p* -value is based on the distribution of the best (worst) funds in 1000 bootstrap resamples. The results in Panel A are for the full sample of hedge funds returns. The results in Panel B are for the sample of fund returns which have been corrected for serial correlation of up to order 2 using the method of Okunev and White. (2003). Panel C reports bootstrap results after removing the backfilled returns. Panel D reports results from a dummy regression model that allows for structural breaks in the slope coefficients. Structural breaks were identified using the Bai and Perron (1998) test.

Panel A: Basic model on full sample of hedge funds

	Bottom	2.	3.	4.	5.	1%	3%	4%	10%	10%	5%	3%	1%	5.	4.	3.	2.	Top
alpha (%)	-5.39	-4.93	-3.62	-3.16	-3.13	-1.86	-0.99	-0.67	-0.31	1.59	2.10	2.57	3.59	4.70	5.20	5.65	5.68	7.35
p -value (bootstrapped)	0.00	0.00	0.05	0.00	0.00	0.05	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00
t -alpha	-7.63	-7.60	-7.43	-6.99	-6.53	-3.51	-2.45	-1.76	-0.82	4.67	5.96	6.84	9.11	16.46	20.45	23.01	25.47	42.23
p -value (bootstrapped)	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.07	0.26	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Panel B: Basic model after adjusting for serial correlation (Okunev and White, 2003)

	Bottom	2.	3.	4.	5.	1%	3%	4%	10%	10%	5%	3%	1%	5.	4.	3.	2.	Top
alpha (%)	-2.28	-1.23	-0.90	-0.86	-0.69	-0.90	-0.49	-0.41	-0.15	1.14	1.46	1.78	3.05	2.54	3.05	3.18	3.39	3.49
p -value (bootstrapped)	0.03	0.03	0.22	0.01	0.09	0.22	0.04	0.17	0.27	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00
t -alpha	-2.70	-2.43	-2.30	-2.03	-1.78	-2.30	-1.51	-0.96	-0.36	4.27	5.25	5.87	8.57	7.15	8.57	8.95	9.71	9.83
p -value (bootstrapped)	0.00	0.01	0.03	0.02	0.05	0.03	0.08	0.17	0.36	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Panel C: Basic model after adjusting for backfill bias

	Bottom	2.	3.	4.	5.	1%	3%	4%	10%	10%	5%	3%	1%	5.	4.	3.	2.	Top
alpha (%)	-2.47	-2.34	-1.59	-1.53	-1.40	-1.05	-0.56	-0.37	-0.09	1.39	1.85	2.07	2.68	4.02	4.11	4.16	4.24	4.48
p -value (bootstrapped)	0.00	0.02	0.03	0.00	0.01	0.06	0.10	0.17	0.43	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
t -alpha	-3.74	-3.54	-3.38	-3.29	-3.04	-2.28	-1.39	-0.98	-0.29	4.67	5.89	6.84	9.97	13.31	13.36	15.29	19.41	27.43
p -value (bootstrapped)	0.01	0.00	0.01	0.01	0.02	0.02	0.17	0.18	0.35	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table II (continued)

Panel D: Basic model with structural break in Dec 2000 identified using the Bai and Perron (1998) test

	Bottom	2.	3.	4.	5.	1%	3%	4%	10%	10%	5%	3%	1%	5.	4.	3.	2.	Top
alpha (%)	-1.38	-0.80	-0.76	-0.76	-0.68	-0.49	-0.06	0.11	0.32	1.68	2.06	2.32	2.78	3.24	3.24	3.29	3.31	3.52
p -value (bootstrapped)	0.12	0.25	0.13	0.14	0.13	0.24	0.46	0.38	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
t -alpha	-1.86	-1.56	-1.55	-1.34	-1.33	-0.81	-0.11	0.21	0.85	7.42	8.91	10.16	13.10	15.68	15.68	16.23	16.84	19.47
p -value (bootstrapped)	0.08	0.10	0.11	0.12	0.11	0.26	0.48	0.42	0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table III

Statistical significance of the best and worst hedge funds' performance grouped by investment types

The sample period is from January 1990 to December 2002. The first and second rows report the OLS estimate of alpha in percent per month and the bootstrapped *p* -value of alpha. Alpha is estimated relative to the Agarwal and Naik (2004) and Fung and Hsieh (2001) factors. The third and fourth rows report the *t* -statistic of alpha based on heteroskedasticity and autocorrelation consistent standard errors as well as the bootstrapped *p* -value of the *t* -statistic. The first column on the left (right) reports funds with the lowest (highest) alpha and *t* -statistic followed by the results for the funds with the second lowest (highest) alpha and *t* -statistic and marginal funds at different percentiles in the left (right) tail of the distribution. The *p* -value is based on the distribution of the best (worst) funds in 1000 bootstrap resamples. Each panel showcases the results for the subsample of funds in a specific investment style category.

Panel A: Long Short Equity Funds

	Bottom	2.	3.	4.	5.	1%	3%	4%	10%	10%	5%	3%	1%	5.	4.	3.	2.	Top
alpha (%)	-3.16	-2.85	-2.64	-2.54	-2.46	-2.01	-1.27	-0.75	-0.36	1.87	2.34	3.00	3.80	4.29	4.32	4.44	4.70	5.68
p -value (bootstrapped)	0.00	0.00	0.00	0.06	0.01	0.00	0.01	0.05	0.20	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.01
t -alpha	-4.21	-3.95	-3.73	-3.57	-3.19	-3.14	-2.52	-1.99	-0.83	4.09	5.06	5.91	9.11	11.47	11.68	14.15	14.44	16.46
p -value (bootstrapped)	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.03	0.23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Panel B: Directional Trader Funds

	Bottom	2.	3.	4.	5.	1%	3%	4%	10%	10%	5%	3%	1%	5.	4.	3.	2.	Top
alpha (%)	-5.39	-3.07	-2.94	-2.61	-2.11	-2.11	-1.36	-0.88	-0.39	2.11	2.72	2.91	3.63	3.95	4.33	4.34	4.49	7.35
p -value (bootstrapped)	0.00	0.01	0.02	0.00	0.04	0.04	0.01	0.11	0.04	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
t -alpha	-7.63	-7.60	-4.54	-3.62	-3.22	-3.22	-2.50	-1.94	-0.98	4.26	5.03	5.69	6.69	6.76	7.23	7.44	7.54	42.23
p -value (bootstrapped)	0.00	0.00	0.00	0.00	0.01	0.01	0.05	0.06	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Panel C: Multi-process Funds

	Bottom	2.	3.	4.	5.	1%	3%	4%	10%	10%	5%	3%	1%	5.	4.	3.	2.	Top
alpha (%)	-1.93	-1.89	-1.35	-0.82	-0.80	-1.89	-0.77	-0.67	-0.24	1.04	1.52	2.00	2.48	2.18	2.42	2.48	2.58	2.66
p -value (bootstrapped)	0.01	0.03	0.04	0.00	0.01	0.03	0.03	0.00	0.43	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
t -alpha	-5.27	-3.65	-3.62	-3.30	-3.06	-3.65	-2.61	-2.08	-0.75	5.39	6.77	7.34	9.70	8.89	9.07	9.70	9.82	10.80
p -value (bootstrapped)	0.00	0.00	0.01	0.02	0.05	0.00	0.02	0.04	0.26	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Panel D: Relative Value Funds

	Bottom	2.	3.	4.	5.	1%	3%	4%	10%	10%	5%	3%	1%	5.	4.	3.	2.	Top
alpha (%)	-2.07	-1.97	-1.54	-1.35	-1.06	-1.35	-0.71	-0.53	-0.24	1.32	1.58	1.93	2.95	2.95	3.09	3.36	3.37	4.61
p -value (bootstrapped)	0.01	0.00	0.05	0.00	0.05	0.00	0.11	0.04	0.34	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
t -alpha	-5.26	-5.22	-3.84	-3.56	-3.29	-3.56	-2.40	-1.61	-0.71	6.08	7.07	8.64	13.19	13.19	14.79	20.45	23.01	25.47
p -value (bootstrapped)	0.00	0.00	0.01	0.02	0.01	0.02	0.01	0.10	0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table III (continued)

Panel E: Security Selection Funds

	Bottom	2.	3.	4.	5.	1%	3%	4%	10%	10%	5%	3%	1%	5.	4.	3.	2.	Top
alpha (%)	-4.93	-3.16	-2.85	-2.64	-2.54	-2.33	-1.24	-0.75	-0.36	1.84	2.33	2.99	3.80	4.43	4.44	4.70	5.65	5.68
p -value (bootstrapped)	0.00	0.00	0.00	0.00	0.06	0.00	0.07	0.05	0.20	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.01
t -alpha	-7.43	-6.99	-4.21	-3.95	-3.73	-3.45	-2.46	-1.87	-0.81	4.05	4.95	5.61	8.53	11.47	11.68	14.15	14.44	16.46
p -value (bootstrapped)	0.00	0.00	0.01	0.01	0.01	0.02	0.02	0.08	0.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Panel F: Fund of Funds

	Bottom	2.	3.	4.	5.	1%	3%	4%	10%	10%	5%	3%	1%	5.	4.	3.	2.	Top
alpha (%)	-2.45	-1.94	-1.37	-1.17	-1.15	-1.17	-0.75	-0.58	-0.22	0.86	1.04	1.28	1.67	1.67	1.72	1.87	2.04	2.49
p -value (bootstrapped)	0.00	0.01	0.00	0.04	0.00	0.04	0.08	0.09	0.24	0.00	0.00	0.00	0.05	0.05	0.00	0.01	0.00	0.00
t -alpha	-6.53	-4.18	-4.13	-4.12	-3.87	-4.12	-3.35	-2.10	-1.21	4.80	6.32	7.00	8.82	8.82	8.93	9.39	10.72	12.37
p -value (bootstrapped)	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.02	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table IV

Statistical significance of the best and worst dead hedge funds' performance grouped by death reason

The sample period is from January 1990 to December 2002. The first and second rows report the OLS estimate of alpha in percent per month and the bootstrapped *p* -value of alpha. Alpha is estimated relative to the Agarwal and Naik (2004) and Fung and Hsieh (2001) factors. The third and fourth rows report the *t* -statistic of alpha based on heteroskedasticity and autocorrelation consistent standard errors as well as the bootstrapped *p* -value of the *t* -statistic. The first column on the left (right) reports funds with the lowest (highest) alpha and *t* -statistic followed by the results for the funds with the second lowest (highest) alpha and *t* -statistic and marginal funds at different percentiles in the left (right) tail of the distribution. The *p* -value is based on the distribution of the best (worst) funds in 1000 bootstrap resamples. Each panel showcases the results for the subsample of dead funds who left the databases for a specific reason.

Panel A: Funds that were terminated

	Bottom	2.	3.	4.	5.	1%	3%	4%	10%	10%	5%	3%	1%	5.	4.	3.	2.	Top
alpha (%)	-5.39	-2.94	-1.89	-1.64	-1.57	-2.94	-1.57	-1.34	-0.69	1.63	1.93	2.64	2.91	2.70	2.75	2.91	4.18	5.65
p -value (bootstrapped)	0.00	0.02	0.03	0.01	0.09	0.02	0.09	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
t -alpha	-4.54	-4.18	-3.65	-3.35	-3.29	-4.18	-3.29	-2.74	-2.23	4.17	5.65	7.14	8.22	7.22	7.25	8.22	11.47	13.19
p -value (bootstrapped)	0.00	0.00	0.00	0.02	0.01	0.00	0.01	0.02	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Panel B: Funds that stopped reporting

	Bottom	2.	3.	4.	5.	1%	3%	4%	10%	10%	5%	3%	1%	5.	4.	3.	2.	Top
alpha (%)	-3.07	-2.42	-2.33	-2.11	-1.35	-2.42	-1.35	-1.02	-0.53	1.85	2.61	3.05	4.32	3.52	3.59	4.32	4.44	4.70
p -value (bootstrapped)	0.01	0.00	0.00	0.04	0.00	0.00	0.00	0.03	0.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
t -alpha	-6.99	-3.84	-3.45	-3.00	-2.80	-3.84	-2.52	-2.34	-1.39	5.60	6.77	8.07	12.37	10.11	10.80	12.37	14.79	25.47
p -value (bootstrapped)	0.00	0.01	0.02	0.02	0.01	0.01	0.02	0.04	0.18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Panel C: Funds that closed

	Bottom	2.	3.	4.	5.	1%	3%	4%	10%	10%	5%	3%	1%	5.	4.	3.	2.	Top
alpha (%)	-2.64	-0.99	-0.30	-0.24	-0.07	-2.64	-2.64	-2.64	-0.99	1.87	2.76	2.76	2.83	1.67	1.83	1.87	2.76	2.83
p -value (bootstrapped)	0.00	0.00	0.29	0.43	0.40	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.05	0.03	0.00	0.01	0.00
t -alpha	-3.95	-2.68	-0.73	-0.23	-0.15	-3.95	-3.95	-3.95	-2.68	6.16	6.76	6.76	42.23	5.36	5.52	6.16	6.76	42.23
p -value (bootstrapped)	0.01	0.02	0.25	0.43	0.44	0.01	0.01	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table V

Bootstrap Procedure Sensitivity Analysis

This table reports sensitivity analyses of the statistical significance of the worst and the best fund's performance to various bootstrap procedures. Column 1 reports results from factor-residual resampling instead of residual resampling. Column 2 reports results from a cross-sectional boostrap where the bootstrap index is the same for all funds for each boostrap. Column 3 reports results from a bootstrap that allows for the presence of a persistent omitted factor. Columns 4 to 6 reports results for the Politis and Romano (1994) stationary bootstrap using various smoothing parameters q. The p-values are based on the distribution of the best and worst fund's in the 1000 bootstrapped samples. Performance is measured relative to the Agarwal and Naik (2004) and Fung and Hsieh (2001) factors.

Panel A: Sensitivity Analysis of the Best Alpha Fund's Performance

	Factor-Residual Resampling	Cross-sectional Bootstrap	Omitted Factor Bootstrap	Politis and Romano (1994) Stationary Bootstrap		
				q=0.1	q=0.5	q=1
alpha (%)	7.352	7.352	7.352	7.352	7.352	7.352
p-value (boostrapped)	0.007	<0.001	<0.001	<0.001	<0.001	<0.001

Panel B: Sensitivity Analysis of the Best Alpha t-statistic Fund's Performance

	Factor-Residual Resampling	Cross-sectional Bootstrap	Omitted Factor Bootstrap	Politis and Romano (1994) Stationary Bootstrap		
				q=0.1	q=0.5	q=1
alpha t-statistic (%)	42.233	42.233	42.233	42.233	42.233	42.233
p-value (boostrapped)	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

Panel C: Sensitivity Analysis of the Worst Alpha Fund's Performance

	Factor-Residual Resampling	Cross-sectional Bootstrap	Omitted Factor Bootstrap	Politis and Romano (1994) Stationary Bootstrap		
				q=0.1	q=0.5	q=1
alpha (%)	-5.386	-5.386	-5.386	-5.386	-5.386	-5.386
p-value (boostrapped)	0.002	<0.001	0.002	<0.001	<0.001	<0.001

Panel D: Sensitivity Analysis of the Worst Alpha t-statistic Fund's Performance

	Factor-Residual Resampling	Cross-sectional Bootstrap	Omitted Factor Bootstrap	Politis and Romano (1994) Stationary Bootstrap		
				q=0.1	q=0.5	q=1
alpha t-statistic (%)	-7.629	-7.629	-7.629	-7.629	-7.629	-7.629
p-value (boostrapped)	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

Table VI

Top 20 hedge fund alpha estimates using seemingly unrelated assets approach of Pastor and Stambaugh (2002)

This table reports the monthly Bayesian posterior alphas and the OLS alphas (from equation (1)) for the top 20 equity long/short funds. The second column reports the OLS alpha estimate (α_{OLS}). Columns three to five report the posterior alpha estimate (α_{SURA}) from the seemingly unrelated assets approach of Pastor and Stambaugh (2002) for different values of the prior of $\sigma_{\alpha N}$. Columns six to eight report the differences between the OLS alpha and the posterior alpha. Column nine reports the number of monthly observations for each fund and columns 10 and 11 report the start and end dates of the fund. The sample period is from January 1990 to December 2002.

		α_{SURA} for $\sigma_{\alpha N}$ of			(α_{OLS}-α_{SURA}) for $\sigma_{\alpha N}$ of			No. of	start	end
fund	α_{OLS} (% p.m.)	0.00	0.02	∞	0.00	0.02	∞	obs.	date	date
1	5.68	4.20	4.21	4.32	1.48	1.47	1.36	33	Mar 00	Dec 02
2	4.70	2.40	2.40	2.41	2.30	2.30	2.30	40	Jul 97	Oct 00
3	4.45	2.73	2.74	2.78	1.72	1.71	1.67	50	Feb 95	Jan 01
4	4.32	0.97	1.10	1.55	3.35	3.22	2.77	30	Jul 98	Dec 00
5	4.30	2.34	2.37	3.14	1.95	1.93	1.16	43	Mar 99	Dec 02
6	4.18	1.40	1.55	1.73	2.78	2.64	2.46	34	Jan 97	Oct 99
7	3.80	2.13	2.11	0.68	1.67	1.68	3.12	34	May 99	Dec 02
8	3.75	2.60	2.58	2.51	1.15	1.17	1.24	80	Jan 96	Dec 02
9	3.64	1.61	1.62	2.02	2.02	2.02	1.62	47	Jan 99	Dec 02
10	3.60	0.27	0.28	0.32	3.33	3.32	3.28	72	Jan 97	Dec 02
11	3.60	3.06	3.07	3.09	0.53	0.53	0.51	39	May 98	Dec 02
12	3.52	3.21	3.21	3.21	0.31	0.31	0.31	35	May 98	Apr 01
13	3.18	-0.12	-0.16	-0.98	3.30	3.34	4.16	61	Mar 97	Dec 02
14	3.11	2.36	2.36	2.36	0.75	0.75	0.75	37	Dec 92	Dec 02
15	3.05	3.33	3.36	4.36	-0.28	-0.31	-1.32	35	Jun 98	Apr 01
16	3.00	1.77	1.76	1.58	1.23	1.24	1.42	56	Mar 98	Dec 02
17	2.98	2.60	2.62	2.73	0.37	0.35	0.25	34	Mar 00	Dec 02
18	2.86	1.94	1.94	1.91	0.93	0.93	0.96	35	Jan 96	Nov 98
19	2.78	1.79	1.79	1.80	0.99	0.99	0.98	42	Dec 98	Dec 02
20	2.64	1.66	1.66	1.71	0.98	0.97	0.93	30	Mar 98	Aug 00
Mean reduction of alpha in percent [α_{OLS} / (α_{OLS} -α_{SURA})]					42%	41%	41%			

Table VII

Summary statistics of top 20 equity long short funds

Summary statistics are presented for the top twenty long short equity hedge funds. Adjusted alpha ('adj alpha') is the estimated alpha of a fund after removing the top three return observations. Factors are the Agarwal and Naik (2004) and Fung and Hsieh (2001) factors chosen from a step wise regression. The Agarwal and Naik (2004) factors are Russell 3000 index (RUS3000), Fama and French size (SMB), book-to-market (HML) and momentum (MOM) factors, MSCI excluding US index (MXUS), MSCI emerging markets index (MEM), Salomon bond index (SBG), Salomon world government index (SBW), Lehman high yield bond index (LHY), Fed trade weighted dollar index (FRBI), Goldman Sachs commodity index (GSCI), Moody's BAA rated corporate bond index (BAA), OTM call option index (SPCX), and OTM put option index (SPPX). The Fung and Hsieh (2001) factors are S&P 500 return (SP), Wilshire small cap minus large cap return (SML), change in the constant maturity yield of the 10 year treasury (TSY), change in the spread of Moody's Baa - 10 year treasury (HYMTSY) bond PTFS (PTFSBD), commodities PTFS (PTFSCOM), currency PTFS (PTFSFX), short term interest rate straddle PTFS (PTFSSIR), and stock PTFS (PTFSSTK). PTFS is primitive trend following strategy. Fund inflow is net dollar flow scaled by last month's assets under management after adjustin Aum is assets under management in USD (millions). The sample period is from January 1990 to December 2002.

fund	alpha	bootstrap p value	rsqr	mean return	std dev	max return	min return	adj alpha /alpha	start date	end date	top three factors [beta]			fund status	start aum	end aum	mean inflow
1	5.680	0.004	0.44	1.51	7.61	18.26	-11.50	0.018	Mar 00	Dec 02	SPPX [0.4]	GSCI [-0.37]	SP [0.75]	live	22.2	64.3	0.035
2	4.700	0.000	0.56	4.20	4.27	17.38	-3.16	0.844	Jul 97	Oct 00	SP [-0.34]	RUS3000 [0.6]	PTFSSTK [0.03]	stop rep	23.6	211.8	0.161
3	4.445	0.000	0.54	2.60	5.66	18.64	-7.45	0.426	Feb 95	Jan 01	SP [0.81]	PTFSSTK [0.12]	PTFSBD [-0.07]	stop rep	21.6	354.0	0.101
4	4.319	0.000	0.70	1.12	6.70	17.30	-10.70	0.616	Jul 98	Dec 00	SP [3.18]	PTFSBD [-0.08]	HML [0.63]	stop rep	23.0	82.0	0.318
5	4.295	0.010	0.49	2.23	8.86	39.82	-21.94	0.886	Mar 99	Dec 02	SPCX [4.39]	BAA [0.63]	SPPX [-3.53]	live	23.3	127.1	0.030
6	4.182	0.000	0.91	1.92	6.59	16.54	-10.59	1.051	Jan 97	Oct 99	GSCI [0.7]	PTFSCOM [-0.12]	SML [0.56]	terminated	38.3	80.5	0.045
7	3.798	0.000	0.92	0.48	11.52	42.60	-31.84	1.616	May 99	Dec 02	MEM [-0.06]	HML [-0.8]	MOM [0.09]	live	21.6	21.7	0.014
8	3.748	0.000	0.64	1.46	8.54	19.60	-19.58	0.672	Jan 96	Dec 02	HML [-0.64]	SPCX [4.66]	SMB [0.49]	live	47.3	23.0	-0.014
9	3.636	0.020	0.49	2.14	8.79	39.82	-22.15	0.913	Jan 99	Dec 02	SPCX [4.36]	MOM [0.18]	SPPX [-4.85]	live	36.6	45.2	-0.017
10	3.596	0.000	0.60	1.21	9.75	49.55	-34.06	0.681	Jan 97	Dec 02	HML [-1.04]	MEM [0.21]	SBG [5.13]	live	42.4	42.7	0.133
11	3.596	0.000	0.44	3.65	11.09	97.61	-23.55	0.424	May 98	Dec 02	SMB [-1.34]	PTFSSIR [-0.24]	na	live	20.7	34.7	0.007
12	3.522	0.000	0.76	2.82	11.94	52.29	-50.30	0.510	May 98	Apr 01	RUS3000 [1.08]	SBW [1.05]	FRBI [-3.66]	stop rep	35.1	23.0	0.014
13	3.180	0.000	0.62	0.90	10.02	49.55	-34.06	1.097	Mar 97	Dec 02	SPCX [4.78]	HML [-1.02]	SML [0.5]	live	41.4	61.0	-0.014
14	3.110	0.009	0.86	0.37	9.00	43.41	-17.29	-0.509	Dec 92	Dec 02	HYMTSY [0.68]	HML [-0.81]	PTFSBD [0.18]	live	20.4	167.5	0.014
15	3.050	0.000	0.53	2.72	4.47	11.50	-5.32	0.936	Jun 98	Apr 01	PTFSSTK [0.12]	MOM [0.27]	HML [0.26]	stop rep	20.4	280.6	0.057
16	2.999	0.000	0.46	1.70	5.88	24.39	-12.72	0.633	Mar 98	Dec 02	PTFSSTK [0.14]	FRBI [-0.67]	MXUS [0.42]	live	20.8	347.3	0.048
17	2.975	0.000	0.79	1.71	9.36	43.86	-21.53	0.680	Mar 00	Dec 02	SP [1.34]	PTFSBD [-0.15]	SBG [2.02]	live	24.9	166.7	0.059
18	2.863	0.000	0.58	3.10	4.84	16.58	-5.30	0.362	Jan 96	Nov 98	MXUS [0.25]	PTFSSTK [0.11]	SPPX [-1.64]	stop rep	113.0	360.0	0.105
19	2.781	0.000	0.83	2.21	10.96	23.33	-26.35	0.043	Dec 98	Dec 02	MEM [0.59]	PTFSCOM[0.13]	SPPX [-2.62]	live	22.0	23.4	0.001
20	2.638	0.000	0.54	2.58	5.29	13.54	-8.32	0.625	Mar 98	Aug 00	HYMTSY [0.57]	SML[0.31]	PTFSSIR [0.07]	terminated	20.7	102.4	0.114

Table VIII

Persistence of Fund Alpha

Hedge funds are sorted each month from January 1993 to December 1999 into decile portfolios based on their alphas estimated over the prior 3 years. We require a minimum of 30 monthly return observations for this estimate. The mean alpha of the funds in each decile is calculated using the next three years of data. The spread is the difference in mean alpha between the funds in decile 1 (highest alpha) and decile 10 (lowest alpha). The factors used to estimate alpha are the Agarwal and Naik (2004) and Fung and Hsieh (2001) factors chosen from a step-wise regression. Closed funds are funds whose average inflows fall below the median average inflow over the evaluation period. Open funds are funds whose average inflows lie above the median average inflow over the evaluation period. Small funds are funds whose assets under management (aum) fall below the median aum at the start of the evaluation period. Large funds are funds whose aum lie above the median aum at the start of the evaluation period. All portfolios are equal-weighted unless stated otherwise. The *t*-statistics, derived from Newey and West (1987) standard errors, are in parentheses. * Significant at the 5% level; ** Significant at the 1% level.

Portfolio	Equal-weighted	Value-weighted	Closed funds	Open funds	Small funds	Large funds
Post-fee alpha spread	0.225**	0.118	0.363**	0.139	0.207**	0.173
	(3.49)	(0.70)	(5.09)	(1.73)	(2.83)	(1.86)
Pre-fee alpha spread	0.397**	0.261	0.537**	0.323**	0.417**	0.365**
	(5.88)	(1.49)	(6.38)	(4.05)	(5.04)	(4.33)