Rank Alpha Funds of Hedge Funds

Carol Alexander and Anca Dimitriu

Carol Alexander
is a professor of risk management and director of research at the ISMA Centre, University of Reading, UK.
c.alexander@ismacentre.rdg.ac.uk

Anca Dimitriu
is a visiting research fellow at the ISMA Centre.
a.dimitriu@ismacentre.rdg.ac.uk

During the bear stock market of the last few years institutional investors as well as high net worth individual investors have found a new source of returns in alternative investment strategies. A rough comparison illustrates the difference in performance between traditional and alternative investments: over the period January 2000 to December 2003, the MSCI World Index lost on average 5.9% per year, while the CSFB-Tremont hedge fund index gained on average 6.8% per year, with much lower volatility. Despite some high-profile losses, investors are committing more assets to the alternative investment industry. According to TASS Research, 2003 was a record-breaking year for the alternative investment industry, the net flow into hedge funds being estimated at 60 billion USD. The same source estimates that hedge fund assets currently total between 725 and 750 billion USD worldwide.

Portfolio optimization within a large universe of hedge funds has become a key area for academic research. This article develops a portfolio construction model that is specifically designed for investors in hedge funds, incorporating specific controls for data reliability issues and operational limitations. In terms of data, the fact that hedge funds report to commercial database providers on a voluntary basis creates a number of sampling issues. First, even if one combines the largest commercial databases available,[1] there is no indication of the size of the population of hedge funds and the degree to which the reporting funds are representative of this population. Beyond the small sample bias, the databases comprise performance estimates rather than liquid market prices and this generates both autocorrelation and a significant amount of noise in the data. Regarding the operational limitations, short sales are not possible, there are minimum investment limits, long lock-up periods, and advance notice, regular subscriptions/redemption as rare as once per year, and sales and early redemption fees.

Given the special features of hedge funds, portfolio management tools that were developed for traditional investments, being heavily reliant on the existence of liquid markets and efficient prices, need refining. High-quality data is an essential ingredient for any portfolio optimization model but this is not likely to be achieved for hedge funds in the near future. We use one of the three largest databases, with both dead and alive funds to diminish the impact of survivorship bias. We account for the instant history bias through the use of dummy variables in factor models and report performance on a relative basis, benchmarked against portfolios that are affected by the same biases. Operational limitations are addressed by imposing constraints on the optimization and by including an estimate of annual turnover as a key diagnostic of the portfolio's performance.

Traditional portfolio optimization models require forecasts of the portfolio expected returns and/or an estimate of their covariance matrix. Often expected returns are estimated using a factor model, but given the highly dynamic and heterogeneous styles used in alternative investments, the non-uniqueness of the factor model representation is one of the most important problems to address when optimizing portfolios of hedge funds. Amenc and Martellini [2003a] show that different factor models can generate very different estimates of a hedge fund's alpha and consequently argue that the hedge fund industry should promote its diversification potential rather than its very uncertain alpha benefits.

In this article we show that, while estimates of a fund's alpha do indeed vary widely according to the factor model used, there is still valuable information contained in the alpha estimate. First we use four different factor models to estimate the alpha of individual hedge funds: our "base case" model is the simplest representation of the fund returns as a function of the two most important underlying asset classes, equities and bonds; the "broad fundamental" factor model employs several indices to capture the performance of the main asset classes, and other factors representing specific types of non-linear strategies such as market timing, volatility trading, and equilibrium trading; the "multi-factor" model is based on hedge fund indices; and finally the "statistical" factor model is based on factors extracted from fund returns through principal component analysis. Then, for each of these four factor models, our hedge fund selection process is determined only by the rank of the fund's alpha and not by the actual value of its estimated alpha. We find significant agreement on the ranking of funds based on their estimated alphas from different factor models. Thus there is considerable scope for optimal hedge fund investment models where funds are selected according to their "rank alpha."

In contrast to other research on incorporating the uncertainty about parameter estimates into optimal portfolios, we require neither assumptions about investors' preferences nor subjective views on expected returns and/or managerial skill.[2] We deal with the uncertainty in alpha estimates by using classic minimum-variance optimization, which is based solely on the funds' covariance matrix. Having considered several methods of linking allocations to the ordinal rank of the fund's alpha estimate we find that the best out-of-sample performance is obtained by a simple "in-out" ranking of the funds, according to whether the alpha estimate is above a predetermined threshold, followed by minimum variance optimization.

Given the reporting practices of hedge funds, there is a high degree of randomness in the sample covariance matrix, which is likely to distort our minimum variance optimization. Some authors advocate "cleaning" the covariance matrix by imposing some factor structure before using it in the portfolio optimization. However, portfolio weight constraints are essential for hedge funds investing and as the sampling errors appear to already be significantly reduced through the weights' constraining we have found no further benefit from imposing a factor structure on the correlation matrix.

HEDGE FUND DATA AND BIASES

Hedge fund data are subject to several measurement biases caused by the data collection process and by the nature of the industry: *survivorship bias*, when a database does not include the performance of funds that ceased operating during the sample period; *selection or self-reporting bias*, when the hedge funds in the database are not representative of the population of hedge funds; *instant history bias*, when the funds entering the database are allowed to back-fill their results; and *multi-period sampling bias*, when the analysis is restricted to funds having a minimum amount of history available. Fung and Hsieh [2000] provide an extensive analysis of biases in the TASS hedge fund database. They estimate a survivorship bias of approximately 3% per annum. Regarding the instant history bias they found an average incubation (back-filled) period of one year with an associated bias of 1.4% p.a., while the multi-period sampling bias was negligible.

Our fund data comes from the Hedge Fund Research (HFR) dead and alive funds databases, from which we select the period January 1990 to May 2003. We restrict our analysis to U.S. domiciled funds reporting net of all fees in USD, having funds under management above 10 million USD and not using leverage. Additionally, to minimize the sample bias of alpha estimates, we require that each fund has at least five years of reporting history. After imposing these selection criteria our database comprises 282 funds.[3] The representation of

To determine the impact of the instant history bias in our database, for each fund we examine the difference between the monthly average of the excess return (over S&P 500) in the first year and the monthly average of the excess return in the first five years. The difference is equivalent to 3.97% and the standard deviation of the difference is 1.01% p.a., so there is a clear "first year" bias in the reported fund performance. In order to eliminate the instant history bias on alpha we use dummy variables for the first year of reporting in all factor models. The estimated multi-period bias is negligible, at –0.33% p.a.. Selection and survivorship biases are addressed by including "dead" funds that have sufficiently long reporting history. But this is not still sufficient to ensure that the portfolio performance is identical to the experience of an investor in these funds, because there is no information on the performance of individual funds after having ceased reporting. Statistics show that some funds stopped reporting to HFR because of extraordinarily good performance but some also because of negative performance. If some funds were liquidated, their investors probably recovered only part of the net asset value last reported. To deal with all these potential biases, we present all performance results on a relative basis, benchmarked against the equally weighted index of all funds in our sample. The relative performance can be interpreted as bias-free since both the portfolios and their benchmark are affected by the same biases.

FACTOR MODELS FOR HEDGE FUNDS

Assuming that investors are only willing to reward managers for superior performance that cannot be easily replicated, a fund's return may be decomposed into the part explained by a factor model, which can be replicated by standard asset baskets and common trading strategies, and the factor model alpha and residual behavior which are attributed to the fund manager's skill. Various fundamental and statistical multi-factor models for hedge funds have been analyzed by Fung and Hsieh [1997, 2001], Schneeweis and Spurgin [1998, 2000], Liang [2001], Agarwal and Naik [2000, 2004], Edwards and Caglayan [2001], Mitchell and Pulvino [2001], and Lhabitant [2001], among others. The wide range of models developed and the fact that no single model dominates the others is explained by the large diversity of the strategies employed by hedge funds and their highly dynamic nature.

The general factor model representation is:

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

where r_{it} is the net of fees excess return on fund i during month t; α_i is the risk-adjusted performance, i.e., the "alpha" of fund i over the estimation sample; F_{kt} is the excess return on the k^{th} risk factor over the month t;[4] β_{ik} is the loading of the fund i on k^{th} factor, i.e., the sensitivity of the fund i to the factor k over the estimation sample; and ε_{it} is the error term.

Since there is no consensus on the best model, we estimate four factor models:

- The "base case" model that only has two factors, U.S. equities (Wilshire 5000 index) and bonds (Lehman Government/Credit Intermediate index). Asness, Krail, and Liew [2001] show that hedge funds may have more market exposure than one expects due to stale prices or illiquidity of the securities they trade, so, following Dimson's [1979] arguments, we also include the lagged equity index excess returns as a factor.
- A broad fundamental model, including as factors: *equity* indices (Wilshire 5000, S&P 500 growth and value, S&P mid-cap, and small-cap to capture differences in equity investment styles, MSCI world index excluding U.S. to account for the investment opportunities outside U.S. and MSCI emerging markets index to capture the emerging markets investment opportunities as a separate asset class); *bond* indices (Lehman Government, Lehman Credit Bond, Lehman High Yield, and Lehman Mortgage Backed Securities); the Fed trade-weighted *foreign exchange* rate index as a proxy for foreign exchange risk; the Goldman Sachs Commodity index to capture *commodity*-related investment risk factors. It is common practice to go beyond static asset class mixes and analyze the performance of funds using simple trading strategies. As suggested by Treynor and Mazuy [1966], squared market returns can proxy for market timing abilities. Additionally, we include two factors capturing specific non-linear trading strategies: the prices' *dispersion* as a leading indicator of equilibrium trading strategies (Alexander and Dimitriu [2005]) and the change in the equity implied volatility index (VIX) to account for *volatility* trades (Schneeweis and Spurgin [1998]).
- The HFR model having the HFR indices as factors. Since they represent portfolios with non-linear exposures to traditional asset classes, this model should explain the returns on individual funds better than the two factor models above (Lhabitant [2001]). We do not reconstruct the indices from the funds in our sub-sample: the HFR indices are a better choice as factors since they are more representative of the entire population of funds and are also investable.
- The PCA factor model uses as factors investable portfolios replicating the first four orthogonal com-

EXHIBIT 1
Base Case Model

	All Funds	Convertible Arbitrage	Distressed Securities	Emerging Markets	Equity Hedge	Equity MN	Equity Non-Hedge	Event-Driven	Fixed Income	Funds of Funds	Market Timing	Managed Futures	Merger Arbitrage	Sectors
Alpha	0.62	0.64	0.44	1.11	0.75	0.38	0.65	0.6	0.32	0.48	0.7	0.93	0.35	0.8
	60%	*100%*	*47%*	*17%*	*62%*	*46%*	*53%*	*65%*	*35%*	*77%*	*73%*	*56%*	*75%*	*53%*
W5000	0.38	0.04	**0.22**	**0.78**	**0.61**	**0.21**	**0.76**	**0.39**	**0.23**	**0.19**	**0.07**	**-0.23**	**0.1**	**0.6**
	80%	*50%*	***76%***	***100%***	***83%***	***65%***	***95%***	***95%***	***65%***	***81%***	***80%***	***61%***	***50%***	***100%***
W5000 (Lagged)	0.12	0.05	**0.15**	**0.44**	0.17	0	**0.28**	**0.1**	0.07	**0.06**	NA	-0.15	**0.06**	0.35
	38%	*40%*	***76%***	***50%***	*30%*	*15%*	***42%***	***45%***	*35%*	***62%***	*0%*	*28%*	***75%***	*5%*
LEH Bond Index	-0.25	**0.13**	-0.4	NA	-0.68	-0.15	0.45	-0.2	0.09	-0.46	-0.52	1.15	-0.17	0.14
	20%	***60%***	*12%*	*0%*	*25%*	*19%*	*5%*	*5%*	*18%*	*29%*	*20%*	*17%*	*25%*	*16%*
Dummy (1st yr)	0.32	**1.42**	-0.16	-3.68	-0.18	2.25	-1.39	**-0.53**	0.79	-0.1	1.61	2.36	NA	-2.15
	24%	***60%***	*12%*	*8%*	*25%*	*19%*	*11%*	***40%***	*35%*	*29%*	*20%*	*28%*	*0%*	*11%*
R^2	0.27	0.2	0.16	0.23	0.32	0.19	0.39	0.32	0.3	0.28	0.24	0.13	0.11	0.29

ponents from principal components analysis (PCA) of the system of all funds' returns. The intuition behind statistical factor analysis is that if a group of funds use similar strategies in the same markets, their returns should be correlated. Through PCA, the major common styles can be extracted from fund returns. We use four principal component factors, and the portfolios replicating these are denoted by PC1 to PC4.[5] The PC1 portfolio is well diversified across all strategies (with particular emphasis on funds of funds and equity funds because most funds are of these types) and, capturing the common trend in fund returns, PC1 has very similar behavior to an equally-weighted index of all funds, to which it is also highly correlated. The PC2 portfolio is clearly dominated by managed futures, which stand out as an investment style with returns uncorrelated to the common trend but still representative of a significant part of the hedge fund population. PC3 comprises mainly equity market neutral and funds of funds, while in PC4 there are technology funds and again equity market neutral funds. Fung and Hsieh [1997] show that the first principal component is explained to a large extent by a linear representation of the traditional asset classes, but this is not the case for the other principal components. Indeed, we find that the first principal component has strong linear relationships with our broad fundamental factors ($R^2 = 0.79$) and hedge fund indices ($R^2 = 0.89$), while the other three principal components have no obvious significant linear relationship with fundamental factors (maximum $R^2 = 0.20$), and a rather weak one with the hedge fund indices (maximum $R^2 = 0.35$). This provides clear evidence that the portfolios replicating the higher principal components are in fact capturing dynamic trading strategies and derivatives, as well as style switching.[6]

MODEL ESTIMATION

Each of the four models was estimated by least square regressions over the period January 1990 to May 2003 on each of the 282 fund returns in excess of the three-month U.S. T-bill rate. For each fund we used the entire data sample available: some funds entered the database after January 1990 or ceased reporting before May 2003. To select the significant factors for each fund, a backward stepwise regression method was applied: starting with the most complete model which passed a multicollinearity filter,[7] one-by-one the non-significant factors were removed until a parsimonious model was obtained. The results of estimating the four factor models in this way are grouped for presentation, combining results for all funds within the same HFR strategy type. The cells in Exhibits 1 to 4 report two figures: the average coefficient estimate over all funds in that strategy (above) and the percentage of these funds for which the coefficient was statistically significant at 10% (below). The coefficient

EXHIBIT 2
Broad Fundamental Model

	All Funds	Convertible Arbitrage	Distressed Securities	Emerging Markets	Equity Hedge	Equity MN	Equity Non-Hedge	Event-Driven	Fixed Income	Funds of Funds	Market Timing	Managed Futures	Merger Arbitrage	Sector
Alpha	0.55	0.81	0.63	1.35	0.58	0.4	0.4	0.57	0.36	0.51	0.62	0.34	0.47	0.57
	60%	*100%*	*59%*	*50%*	*47%*	*54%*	*47%*	*65%*	*47%*	*79%*	*67%*	*39%*	*100%*	*63%*
W5000	0.62	0.04	0.18	**1.14**	0.82	NA	**0.84**	0.64	0.26	0.28	0.36	0.69	NA	0.63
	18%	*10%*	*6%*	***50%***	*21%*	*0%*	***42%***	*5%*	*29%*	*21%*	*20%*	*6%*	*0%*	*21%*
SP500g	0.23	0.12	NA	NA	0.29	**0.29**	NA	NA	-0.08	0.01	**0.23**	NA	NA	0.52
	7%	*20%*	*0%*	*0%*	*2%*	***31%***	*0%*	*0%*	*6%*	*2%*	***33%***	*0%*	*0%*	*5%*
SP500v	0.22	-0.01	0.29	0.21	0.32	0.26	0.47	0.09	0.08	-0.31	NA	NA	NA	0.45
	7%	*20%*	*12%*	*8%*	*2%*	*15%*	*5%*	*5%*	*6%*	*4%*	*0%*	*0%*	*0%*	*26%*
MD400	0.4	NA	0.4	0.24	0.46	0.1	0.58	0.31	NA	0.15	0.15	NA	0.04	1.01
	10%	*0%*	*6%*	*25%*	*21%*	*4%*	*11%*	*15%*	*0%*	*6%*	*7%*	*0%*	*25%*	*11%*
SC600	**0.39**	0.09	**0.15**	0.45	**0.55**	0.34	0.81	**0.37**	0.29	**0.22**	0.33	NA	**0.07**	0.76
	38%	*10%*	***47%***	*8%*	***47%***	*15%*	*37%*	***70%***	*24%*	***62%***	*7%*	*0%*	***75%***	*37%*
MSCIW	0	0.09	-0.08	0.13	0.19	-0.01	0.41	NaN	0.04	0.05	NaN	**-0.25**	NaN	0.16
	9%	*10%*	*6%*	*8%*	*6%*	*8%*	*16%*	*0%*	*6%*	*4%*	*0%*	***50%***	*0%*	*5%*
SCI (EMF)	0.18	0.07	0.17	**0.56**	0.18	0.04	0.21	0.15	0.14	**0.09**	0.06	-0.22	NaN	0.1
	20%	*20%*	*24%*	***67%***	*21%*	*12%*	*16%*	*5%*	*12%*	***31%***	*20%*	*6%*	*0%*	*11%*
EH (GOV)	0.25	0.88	NaN	NaN	0.3	0.41	0.66	NaN	0.11	-0.22	-0.69	1.12	NaN	0.41
	6%	*10%*	*0%*	*0%*	*4%*	*12%*	*5%*	*0%*	*6%*	*4%*	*13%*	*6%*	*0%*	*16%*
LEH	-0.21	0.09	-0.23	NaN	-0.33	NaN	NaN	-0.47	0.03	-0.22	NaN	NaN	-0.1	NaN
	6%	*20%*	*24%*	*0%*	*6%*	*0%*	*0%*	*10%*	*6%*	*6%*	*0%*	*0%*	*25%*	*0%*
LEH (HY)	0.09	**0.2**	**0.27**	-0.46	-0.11	-0.06	0.1	**0.26**	0.17	0.08	0.12	0.4	**0.03**	**-0.02**
	24%	***30%***	***65%***	*8%*	*23%*	*12%*	*21%*	***30%***	*6%*	*17%*	*27%*	*22%*	***50%***	***37%***
LEH (MB)	-0.2	NA	-0.65	NA	NA	NA	NA	0.19	0.27	-0.39	NA	0.05	NA	-1.11
	5%	*0%*	*6%*	*0%*	*0%*	*0%*	*0%*	*5%*	*18%*	*10%*	*0%*	*11%*	*0%*	*5%*
FX	-0.45	0.08	NA	NA	-0.48	-0.18	-0.77	-0.51	-0.23	0.02	NA	**-0.65**	NA	-0.41
	11%	*10%*	*0%*	*0%*	*8%*	*8%*	*16%*	*5%*	*12%*	*8%*	*0%*	***67%***	*0%*	*16%*
GSCI	0.02	-0.03	-0.13	**0.17**	-0.01	-0.06	-0.03	0.08	-0.01	0.05	0.08	**-0.18**	-0.07	0.27
	19%	*10%*	*12%*	***42%***	*8%*	*12%*	*26%*	*15%*	*12%*	*23%*	*20%*	***39%***	*25%*	*26%*
W5000	0.13	0.03	0.11	0.19	0.14	0.14	0.25	0.09	0.08	0.07	-0.27	0.38	**0.05**	0.21
	18%	*10%*	*18%*	*8%*	*28%*	*4%*	*26%*	*25%*	*18%*	*17%*	*7%*	*6%*	***50%***	*21%*
SC600^2	0	**-0.01**	**-0.01**	**-0.02**	**0.01**	**0**	**0**	-0.01	**0**	**0**	**0.01**	**0.01**	0	0.01
	39%	***60%***	***53%***	***33%***	***36%***	***50%***	***42%***	*25%*	***41%***	***35%***	***33%***	***50%***	*25%*	*26%*
EH (HY)^2	0.01	**-0.04**	**0.02**	**0**	-0.02	0.02	**0.13**	0	**0.01**	-0.02	**-0.01**	**0.04**	**-0.01**	**0.01**
	28%	***40%***	***47%***	***33%***	*26%*	*19%*	***37%***	*5%*	***35%***	*19%*	***40%***	***33%***	***50%***	***37%***
VIX	0	0	0	-0.01	**0**	**0**	0	0	0	0	0	**0**	0	**0**
	27%	*20%*	*18%*	*25%*	***30%***	***38%***	*11%*	*15%*	*6%*	*27%*	*27%*	***61%***	*25%*	***32%***
DISP	0.56	NA	0.63	-0.95	**1.97**	-1.09	1.59	-0.88	0.4	**0.44**	-6.18	-2.95	-1.37	**1.41**
	29%	*0%*	*18%*	*25%*	***45%***	*23%*	*26%*	*20%*	*12%*	***33%***	*13%*	*17%*	*25%*	***58%***
Dummy	0.43	**0.92**	0.23	-1.39	0.05	1.55	-1.33	**0.48**	**0.75**	0.15	1.41	3.52	NA	-0.44
(1st yr)	*25%*	***50%***	*24%*	*17%*	*28%*	*23%*	*11%*	***40%***	***35%***	*29%*	*20%*	*11%*	*0%*	*16%*
R^2	0.36	0.27	0.29	0.36	0.42	0.24	0.49	0.43	0.38	0.39	0.26	0.22	0.23	0.42

Exhibit 3

Multi-Factor HFRI Model

	All Funds	Convertible Arbitrage	Distressed Securities	Emerging Markets	Equity Hedge	Equity MN	Equity Non-Hedge	Event-Driven	Fixed Income	Funds of Funds	Market Timing	Managed Futures	Merger Arbitrage	Sectors
Alpha	-0.12	0.18	-0.34	0.34	-0.24	-0.12	-0.58	-0.56	0.02	-0.01	0.02	0.56	-0.07	-0.31
	48%	60%	71%	42%	36%	38%	47%	70%	47%	46%	40%	50%	100%	53%
Convertible Arbitrage	0.27	**0.89**	-0.25	1.84	0.3	0.12	-0.32	0.39	**0.42**	0.11	0.07	-0.27	**-0.06**	-0.74
	25%	**50%**	**6%**	**33%**	**25%**	**12%**	**11%**	**25%**	**47%**	**29%**	**33%**	**28%**	**75%**	**11%**
Regulation D	-0.03	**-0.01**	0	**-0.42**	0.18	-0.08	-0.32	-0.04	**-0.06**	0.1	**0.02**	-0.69	0.06	0.05
	27%	**50%**	29%	**42%**	25%	12%	26%	15%	**53%**	27%	**40%**	11%	25%	26%
Relative Value	-0.34	NA	-1.48	**-1.64**	-0.6	-0.24	-0.39	-0.04	**0.09**	-0.42	0.5	-0.76	-0.86	0.59
	28%	0%	6%	**50%**	23%	35%	21%	35%	**41%**	27%	33%	33%	25%	37%
Distressed Securities	0.39	0.23	**1.3**	3.39	-0.4	-0.12	0.12	**1.01**	-0.65	0.17	0.83	-0.61	0.11	-0.95
	30%	30%	**100%**	17%	26%	15%	26%	**60%**	18%	33%	7%	28%	25%	11%
Emerging Market	0.33	-0.03	-0.2	**1.44**	-0.02	0.19	0.15	0.29	0.04	0.21	-0.1	-0.26	-0.13	-0.1
	22%	30%	18%	**100%**	8%	15%	21%	35%	29%	17%	33%	11%	25%	16%
Equity Hedge	0.68	-0.25	-0.12	NA	**1.19**	0.98	1.1	0.46	NA	0.47	0.39	-0.58	-0.13	1.03
	21%	20%	12%	0%	**40%**	12%	11%	10%	0%	31%	13%	22%	25%	21%
Equity Market Neutral	0.29	-0.36	-0.43	1.56	0.05	**1.08**	0.81	0.06	-0.27	0.18	-0.25	NA	NA	-0.06
	24%	20%	18%	8%	30%	**58%**	21%	10%	18%	21%	27%	0%	0%	37%
Equity Non-Hedge	0.75	-0.2	0	-0.25	1.13	0.51	**1.14**	0.42	0.37	0.19	0.34	0.71	0.08	0.78
	16%	10%	12%	8%	19%	15%	**68%**	15%	12%	4%	13%	6%	25%	11%
Event-Driven	-0.13	NA	-0.65	**-1.65**	0.34	-0.04	0.18	0.56	-0.02	0.17	-0.69	-1.31	**0**	0.23
	23%	0%	18%	**42%**	11%	15%	16%	50%	12%	27%	20%	33%	**75%**	26%
Fixed Income (Total)	-0.18	**0.61**	**-0.54**	NA	0.49	-1.68	-0.15	-0.08	**-0.45**	-0.63	0.08	-2.44	**0.15**	**0.95**
	34%	**40%**	**47%**	0%	34%	19%	32%	35%	**53%**	38%	33%	17%	**50%**	**47%**
Fixed Income (Convertible Arbitrage)	-0.01	-0.19	0.2	0.86	-0.2	0.59	-0.2	0.13	0.19	0.13	0.08	-0.43	NA	-0.3
	20%	20%	35%	8%	17%	4%	26%	20%	35%	23%	7%	33%	0%	11%
Fixed Income (High Yield)	0.03	NA	0.07	**-0.26**	-0.13	0.24	0.96	-0.75	**0.78**	-0.05	-0.27	0.05	**0.02**	0.09
	26%	0%	35%	**42%**	19%	15%	16%	30%	**47%**	23%	40%	28%	**50%**	26%
Fixed Income (Arbitrage)	-0.09	0.26	**-0.22**	-0.57	0.13	-0.3	-0.49	-0.41	0.35	0.16	0.41	-1.02	0.53	-0.37
	25%	10%	**41%**	25%	25%	19%	26%	35%	35%	19%	33%	11%	25%	32%
Fixed Income (Diversified)	0.27	-0.4	0.66	0.76	-0.59	0.37	0.78	0.3	**0.46**	**0.17**	-0.8	**1.55**	**-0.19**	-0.48
	34%	30%	35%	25%	23%	19%	16%	15%	**41%**	**48%**	33%	**83%**	**50%**	37%
Fixed Income (Mortgage Backed)	-0.09	-0.51	0.02	-0.42	-0.25	0.37	0.14	0.08	0.19	0.19	-0.36	-0.22	**-0.2**	-0.67
	28%	20%	29%	33%	34%	19%	11%	20%	47%	29%	33%	28%	**50%**	26%

E X H I B I T 3 (continued)
Multi-Factor HFRI Model

Funds of Funds	0.22	0.07	0.04	-1.34	-0.21	-0.55	-0.66	-0.49	0.33	**0.63**	0.08	**1**	NA	**-0.36**
	30%	*20%*	*18%*	*8%*	*23%*	*8%*	*26%*	*20%*	*18%*	***65%***	*13%*	***50%***	*0%*	***42%***
Market Timing	0.08	**-0.05**	**-0.14**	0.24	0.14	0.05	**-0.08**	0.12	0.07	0.05	**0.41**	-0.87	NA	0.46
	34%	***40%***	***41%***	*17%*	*38%*	*31%*	***42%***	*25%*	*29%*	*37%*	***67%***	*17%*	*0%*	*26%*
Macro	0.31	NA	0.12	-0.6	0.19	0.23	0.34	0.32	0	0.12	0.47	**1**	0.04	0.23
	28%	*0%*	*18%*	*25%*	*26%*	*27%*	*21%*	*25%*	*24%*	*23%*	*33%*	***72%***	*25%*	*37%*
Short Selling	0.04	0.22	-0.04	-0.69	-0.03	0.09	0.08	0.03	-0.12	0.11	0.24	0.23	NA	-0.23
	31%	*10%*	*35%*	*8%*	*36%*	*15%*	*37%*	*35%*	*18%*	*37%*	*47%*	*39%*	*0%*	*32%*
Merger Arbitrage	0.6	0.25	0.51	**2.24**	0.89	0.57	1.29	**0.24**	-0.33	-0.06	0.89	0.52	0.8	0.66
	30%	*10%*	*18%*	***42%***	*32%*	*23%*	*32%*	***50%***	*29%*	*31%*	*20%*	*6%*	*100%*	*37%*
Sector (Total)	-0.02	0.1	-0.05	-0.55	0.3	-0.13	-0.03	0.24	-0.21	-0.05	NA	0.76	NA	-0.54
	13%	*30%*	*35%*	*25%*	*13%*	*8%*	*11%*	*15%*	*12%*	*8%*	*0%*	*6%*	*0%*	*16%*
Sector Energy	0.05	**0**	0.06	0.12	0.04	-0.07	0.07	**0.02**	-0.02	0.01	0.02	**0.27**	NA	**0.11**
	30%	***40%***	*6%*	*8%*	*32%*	*35%*	*32%*	***45%***	*35%*	*23%*	*13%*	***44%***	*0%*	***58%***
Sector Financial	-0.03	-0.04	-0.14	-0.13	0.02	0.1	**-0.16**	0.06	NA	-0.03	-0.07	**-0.2**	-0.06	0.2
	31%	*10%*	*24%*	*25%*	*36%*	*15%*	***47%***	*35%*	*0%*	*37%*	*33%*	***50%***	*25%*	*32%*
Sector HC/Bio	-0.02	0.01	-0.11	-0.04	0.02	**-0.1**	-0.1	0.01	0.07	0.03	-0.19	-0.32	**0.02**	0.39
	24%	*30%*	*12%*	*25%*	*13%*	***54%***	*21%*	*30%*	*12%*	*23%*	*13%*	*28%*	***50%***	*32%*
Sector Real Estate	0.1	**-0.01**	-0.51	0.44	-0.16	0.41	0.33	0.13	0.11	0.07	0.56	-0.61	**-0.11**	0.15
	20%	***40%***	*12%*	*17%*	*17%*	*12%*	*32%*	*5%*	*29%*	*21%*	*27%*	*6%*	***50%***	*32%*
Sector Technical	0.37	NA	0.17	NA	0.41	-0.05	0.53	NA	-0.03	0.07	0.38	NA	NA	0.67
	7%	*0%*	*12%*	*0%*	*9%*	*8%*	*5%*	*0%*	*6%*	*2%*	*13%*	*0%*	*0%*	*26%*
Sector Miscellaneous	0.06	0.01	-0.1	0.81	0.16	-0.11	0	-0.06	NA	-0.03	-0.16	0.32	**0.09**	**0.01**
	22%	*30%*	*12%*	*8%*	*23%*	*15%*	*16%*	*30%*	*0%*	*23%*	*13%*	*39%*	***50%***	***42%***
Dummy (1st yr)	0.46	**0.9**	0.42	NA	0.52	0.81	-1.15	0.3	1.71	-0.02	1.14	1.99	NA	-0.69
	28%	***40%***	*35%*	*0%*	*34%*	*27%*	*21%*	*30%*	*12%*	*27%*	*27%*	*39%*	*0%*	*32%*
R^2	0.58	0.42	0.56	0.58	0.6	0.37	0.62	0.66	0.59	0.67	0.51	0.47	0.54	0.68

Exhibit 4
PCA Model

	All Funds	Convertible Arbitrage	Distressed Securities	Emerging Markets	Equity Hedge	Equity MN	Equity Non-Hedge	Event-Driven	Fixed Income	Funds of Funds	Market Timing	Managed Futures	Merger Arbitrage	Sectors
Alpha	-0.21	0.56	0.36	0.34	-0.57	0.01	-1.09	-0.15	0.3	-0.07	0.27	-0.92	0.31	-0.65
	50%	*80%*	*47%*	*33%*	*36%*	*31%*	*58%*	*50%*	*24%*	*56%*	*73%*	*78%*	*75%*	*58%*
PC1	0.6	0.08	0.51	1.8	0.93	0.16	1.17	0.8	0.3	0.39	0.06	-0.18	0.13	1.12
	79%	*50%*	*88%*	*100%*	*85%*	*46%*	*95%*	*100%*	*65%*	*92%*	*87%*	*28%*	*75%*	*89%*
PC2	0.19	0.03	0.06	0.27	-0.07	0.15	0.05	0.15	0.02	0.17	0.09	1.86	0	0.04
	39%	*20%*	*18%*	*33%*	*34%*	*46%*	*42%*	*30%*	*35%*	*44%*	*33%*	*94%*	*0%*	*37%*
PC3	-0.02	-0.07	-0.34	-1.37	0.28	0.03	0.06	-0.24	-0.19	0.06	-0.06	0.18	-0.04	0.16
	44%	*20%*	*47%*	*58%*	*45%*	*42%*	*37%*	*65%*	*35%*	*50%*	*33%*	*33%*	*25%*	*37%*
PC4	0.16	0.02	-0.17	0.21	0.24	0.16	0.59	0.14	-0.07	0.02	0.37	0.24	0.04	0.23
	29%	*10%*	*18%*	*17%*	*36%*	*38%*	*53%*	*35%*	*18%*	*25%*	*20%*	*22%*	*25%*	*37%*
Dummy (1st yr)	0.21	0.85	0.49	0.39	0.33	0.29	0.27	-0.16	0.24	0	0.51	0.51	-0.15	-0.45
	23%	*60%*	*12%*	*8%*	*30%*	*31%*	*16%*	*25%*	*35%*	*21%*	*13%*	*11%*	*0%*	*16%*
R^2	0.39	0.17	0.23	0.33	0.47	0.22	0.49	0.46	0.32	0.46	0.33	0.52	0.12	0.46

standard errors were computed using the Newey-West [1987] heteroscedasticity and autocorrelation consistent covariance matrix.

On average, the base case model explains only 27% of the total variance of fund excess returns (*Exhibit 1*). This is mainly due to the diverse dynamic strategies employed in the alternative investment industry which induce non-linear exposures to traditional asset classes. Still, 80% of funds are significantly correlated with the Wilshire 5000 excess returns (average beta = 0.3). Illiquidity is also important, since for 38% of funds the lagged Wilshire excess returns are significant determinants. However, bond index returns are significant for only 20% of funds. The average alpha is insignificant for approximately half the funds, being positive and significant for 48% of funds, and negative and significant for only three out of 282 funds.

The broad fundamental model (*Exhibit 2*) includes a total of 17 factors but the average number of significant factors for an individual fund was only 2.5. Nevertheless, the average R^2 across all funds was 36%, a considerable increase from the base case model. The broad fundamental model better explained the returns of funds in the following classes: emerging markets, equity hedge and non-hedge, event driven, convertible bonds, financial and technology sectors. However, the returns for some fixed income, macro, and relative value funds were not well modeled by this approach. The most significant factor, determining the excess returns of 38% of the funds is the small-cap S&P index which influences funds trading on distressed securities, equity hedge and non-hedge, event driven, funds of funds, and technology. Additionally, the squared returns of the small cap S&P index are significant in 40% of models, indicating use of leverage and market timing abilities. The other equity style indices are only significant in 6%-10% of the funds. Another important factor is the Lehman High Yield index (for convertible arbitrage, distressed securities, event driven, managed futures, merger arbitrage, and funds of funds, as well as equity hedge and non-hedge funds, the latter two having a negative average beta) and its squared excess returns (highly significant and positive for funds in distressed securities, equity non-hedge, and managed futures, and significant but negative for convertible arbitrage, equity hedge, funds of funds, market timing, and technology funds). The MSCI emerging markets index was significant for 20% of the funds: in addition to the funds primarily trading in emerging markets and funds of funds investing in these, distressed securities, equity hedge and non-hedge, event driven, and technology funds all have significant exposure to emerging markets. The GS commodity index is a significant factor for emerging markets, event driven, funds of funds, equity hedge and non-hedge, and, as expected, managed futures. While the broad equity market indices and the Fed trade weighted forex index are generally less significant than other factors, the change in S&P 500 implied volatility and the Dow Jones price dispersion index are among the most significant factors. Each of these factors has positive coefficients and is significant for almost 30% of funds. The first-year reporting dummy confirmed a significant positive bias for most strategies. Alpha was most significant for emerging markets and financial sector funds and also for convertible arbitrage, relative value, and short-selling, but for these a high alpha could just result from the lower explanatory power of the model.

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

The estimation results for the PCA factor model are presented in Exhibit 4. The only strategies with positive and significant alphas are convertible arbitrage and merger arbitrage, but they also have a low average R^2, so the abnormal return could be due to omitted risk factors. The average R^2 across all strategies is 39%, greater than for the fundamental factor model but less than the HFR model. The strategies with the highest R^2 are the equity hedge and non-hedge, funds of funds, and technology funds, as expected since the database is dominated by these types of funds. The PC1 portfolio is a significant factor for 79% of funds, PC2 portfolio for 39% funds, PC3 for 44% funds, and PC4 is significant for 29% of funds. All strategies, except for managed futures have positive average betas on the PC1 portfolio. Also, most strategies have positive betas on the PC2 portfolio and negative betas on PC3 portfolio.

RANK ALPHA

Substantial differences in alphas estimated from the four factors models have been identified in Exhibits 1-4. For the "average fund,"[8] the two index model estimated an alpha of 0.69% per month (highly significant and is equivalent to 8.5% per annum). The "average fund" has an alpha of 5.1% in annual terms according to the broad fundamental model, but a negative alpha (though not statistically significant) according the HFR model, and significantly negative according the PCA model. Despite this, we find significant agreement on the sign of alpha from different models, and on the rank of a funds' alpha. Of the funds having at least one positive alpha estimate, in 30% of cases there is perfect agreement between all models in terms of alpha's sign. The largest difference in ranks of alpha estimates arose between the base case and the PCA models (rank correlation = 0.19), the most similar rank alphas were from the base case and the broad fundamental models (rank correlation = 0.68), and the other rank correlations were all between 0.31 and 0.45. Hence the factor models tend to agree on the sign of alpha and on the alpha ranking of funds, despite the fact that the range of alpha estimates produced by them is wide. Interestingly, this feature has also been found in the TASS hedge fund database by Amenc and Martellini [2003a].

In summary, there is a significant disagreement between the factor models on the "average fund" alpha, ranging from –2.5% to 8.5% per annum, and the dispersion of alpha estimates is even higher at the level of individual funds. This leads us to conclude that funds of hedge funds that rely on the accuracy of alpha estimates cannot be implemented without significant model risk. However, some agreement can be achieved on the sign and the ranking of alpha for individual funds and there is considerable scope for more flexible models based precisely on ranking alpha.

OPTIMIZING FUNDS OF HEDGE FUNDS

Optimizers are well known to be error enhancers (Michaud [1989]), so the quality of the data is essential. Since this is not likely to be achieved with hedge funds in the near future one can only aim for a better solution than naïve diversification. Therefore we begin by testing the benefits of naïve diversification through simulation, randomly drawing without replacement 5, 10, and up to 80 funds in multiples of 5 from our database and independently forming equally weighted portfolios having no

EXHIBIT 5

Out-of-Sample Performance of Randomly Selected Portfolios

	Unbounded		Bounded	
	Sample matrix	Cleaned matrix	Sample matrix	Cleaned matrix
Annual volatility	2.32	2.3	1.97	2.54
Annual returns	7.49	7.2	7.92	7.69
Skewness	0.03	-0.26	-0.15	-0.4
Excess kurtosis	5.47	5.73	4.14	5.28
Information ratio	3.22	3.12	4.01	3.02

EXHIBIT 6

Out-of-Sample Performance of Rank Alpha Minimum Variance (RAMV) Portfolios, Maximum Information Ratio (MIR) Portfolios, and the Equally Weighted (EW) Portfolio of All Funds

	Base Case		Broad Fundamental		HFR		PCA		Overall	EW
	RAMV	MIR	RAMV	MIR	RAMV	MIR	RAMV	MIR	RAMV	
Annual returns	8.28	9.24	8.15	8.55	9.44	10.39	8.94	8.98	9.06	10.44
Annual volatility	1.35	1.86	1.34	1.81	1.7	1.97	1.29	1.44	1.51	6.89
Skewness	0.3	0.71	0.06	0.01	0.1	0.57	0.22	0.5	0.49	-0.07
Excess kurtosis	-0.08	0.57	-0.03	1.92	0.24	0.46	-0.34	0.1	0.35	1.91
Information ratio	6.15	4.97	6.06	4.73	5.56	5.28	6.91	6.22	5.99	1.51
Turnover	6.3	8.99	7.13	9.77	7.66	7.05	4.94	7.59	7.02	4.92

style consideration. For each size of portfolio we repeat the experiment 1,000 times, and estimate the first four moments of the fund of funds portfolio returns distributions for each portfolio size. Our results indicate that diversification across all strategies works well, with most of the diversification benefits obtained at around 30 funds in the portfolio. Practitioner standards also appear to favor portfolios of at least 20 to 30 funds so in the following our portfolios will include a number of funds in this range.

Assuming investors have quadratic preferences, the classical portfolio theory of Markowitz [1952] requires knowledge of the first two moments of the returns distribution. Estimating expected returns has been shown to be a difficult task even for traditional asset classes (Merton [1980]; Jorion [1985]), and this can be a main reason why mean-variance efficient portfolios perform poorly out-of-sample (Frost and Savarino [1986, 1988]; Jorion [1986]; Michaud [1989]; Best and Grauer [1991]). Accurate estimates of expected returns are even more difficult to obtain for alternative investments, as shown in the previous section. On the mean-variance efficient frontier the only portfolio that does not require estimates of expected returns is the minimum variance portfolio, so in the hedge funds world this is a natural choice.

CLEANING THE COVARIANCE MATRIX

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

$$\mathbf{W}_{MV} = \frac{\Sigma^{-1}\underline{1}}{\underline{1}'\Sigma^{-1}\underline{1}}$$

where Σ is the covariance matrix of the fund returns and $\underline{1}$ is a vector of ones. If restrictions are imposed on weights, the solution can be obtained numerically. Since the covariance matrix of fund returns is the only input required in the model, the results will strongly depend on its accuracy. Given that the number of funds in our sample is much larger than the number of data points, we are concerned with the estimation risk in the sample covariance matrix. Indeed, a large part of the information content of the empirical correlation matrix is driven by randomness. We found that the "true" correlation structure is captured by the first four eigenvalues, while the rest can be ascribed to noise and measurement errors. Since the presence of such noise is likely to perturb the minimum variance optimization, its reduction is essential. Several other solutions have been offered, including imposing a factor structure (e.g., Sharpe [1963]; Chan, Karceski, and Lakonishok [1999]), the use of an optimal shrinkage towards the mean or the single factor model (Jorion [1985, 1986]; Ledoit and Wolf [2003]), and the introduction of portfolio constraints (Jagannathan and Ma [2003]). Following Plerou et al. [2002] for traditional assets, and Amenc and Martellini [2002, 2003b] for alternative investments, we reconstruct the empirical correlation matrix using only the eigenvalues that deviate from those of a random correlation matrix. We therefore construct the "cleaned" correlation matrix as:

$$\mathbf{C} = \mathbf{W}\Lambda_{\mathbf{c}}\mathbf{W}'$$

where Λ_c is the diagonal matrix of the ordered eigenvalues of the correlation matrix of fund returns with all but the first four eigenvalues replaced by zeros, and $\mathbf{W}$ is the matrix of eigenvectors. The diagonal elements of $\mathbf{C}$ are set equal to one, and then the "cleaned" covariance matrix is $\mathbf{V} = \mathbf{DCD}$, where $\mathbf{D}$ is the diagonal matrix of standard deviations of each fund returns.

In order to test the efficiency of the "noise cleaning"process, we compared, out-of-sample, the variances of minimum variance portfolios constructed from 1) the empirical covariance matrix and 2) the "cleaned"covariance matrix. We constructed 1,000 portfolios, each with 25 randomly selected funds, and optimized them based on both the "cleaned" and the sample covariance matrices to achieve minimum variance. For estimating the covariance matrices we used a rolling sample of 60 months. The first portfolios were constructed in January 1998 and rebalanced every six months until January 2003. Between two rebalancing moments, the portfolios are left unmanaged and their out-of-sample performance monitored in order to determine the out-of-sample variance. Exhibit 5 summarizes these results: when no constraints are imposed on the portfolio weights the out-of-sample variance of the portfolio constructed on the "cleaned" covariance matrix is smaller than that based on the sample covariance matrix, indicating the effectiveness of the noise cleaning process. However, no short sales are allowed in a hedge funds portfolio and upper bounds are also normally imposed to reduce concentration risk. We therefore introduce a non-negativity constraint and an upper bound on individual fund weights at 20%, following standard practice in this respect. When we repeated the previous analysis with the constraints in place, the out-of-sample results were in favor of the sample covariance matrix. Interestingly, this feature has also been observed by Jagannathan and Ma [2003]. Hence in the following we use the ordinary sample covariance matrix instead of the "cleaned" one for constrained optimization purposes.

PERFORMANCE ANALYSIS OF RANK ALPHA PORTFOLIOS

The simulation results can only give a rough idea of the average performance expected from minimum variance portfolios of hedge funds. Next we enhance the minimum variance portfolios by introducing a fund selection criterion based on the funds' alpha from each of the four factor models presented in the previous section. For a given factor model, we first select all the funds having positive alphas that are significant at 10%. We also examine a more restrictive selection criterion, which requires that all models rank the fund with positive alpha, and in at least three models the fund's alpha is significant at the 10% significance level. In each case, the portfolio weights are allocated so that the fund of funds has minimum variance.

In order to test the out-of-sample performance of these models we use the period January 1990 to December 1997 to calibrate the factor models and select the first set of funds. We keep the non-negativity and 20% upper bound constraints in place. The first minimum variance portfolios are set up in January 1998 and left unmanaged for the next six months. The portfolios are then rebalanced every six months, reselecting funds based on the sign and significance of the alphas estimated over the entire data sample available at the portfolio construction moment and the covariance matrices estimated over the previous 60 months. For reference, we compare these results with an equally weighted (EW) portfolio of all funds in our database.

The portfolios of hedge funds that employ the alpha selection criterion have a significantly improved out-of-sample performance.[9] The portfolio statistics are reported in the columns labeled RAMV (Rank Alpha Minimum Variance) in Exhibit 6. All portfolios have average annual returns in the range of 8% to 9.5%, with an annual volatility of only 1.3% to 1.7%. Their evolution is very constant, with no more than three months (out of 66) having negative returns in any of the models. The lowest volatility is displayed by the PCA portfolio and this also has the highest average annual information ratio (6.91). The highest returns are produced by the HFR factor model portfolio but given the higher volatility, the information ratio of this portfolio was the lowest (5.56). The null hypothesis of normally distributed out-of-sample returns cannot be rejected for any of these portfolios. We have also examined the out-of-sample performance of portfolios excluding funds of funds, as some institutional investors would prefer to invest solely in hedge funds in order to avoid paying two layers of fees. The information ratios of these portfolios were 5.11 for the base case model, 5.26 for the fundamental model, 5.58 for the HFR index model, and 5.69 for the PCA model. The skewness and excess kurtosis stay in the same range as for the portfolios including funds of funds. By contrast, the equally weighted portfolio has comparable returns to the rank alpha portfolios but much higher volatility (resulting in an information ratio of only 1.5, if funds of funds are included, and even lower than this if funds of funds are excluded) and its returns display significant excess kurtosis.

Given the significant trading limitations for hedge funds portfolios we investigate the structural stability of these portfolios. The portfolio turnover is computed as the absolute difference in fund weights from one rebalancing period to the next. With 10 rebalancing points in our out-of-sample test, the maximum turnover is 20 (10 $\times$ 200%). The estimated turnover ranges from 4.9 (equivalent to 24% restructuring p.a.) for the PCA portfolio and the EW portfolio to 7.6 (equivalent to 38% restructuring) for the HFR portfolio. We note that even the base case portfolio produces good results and the more conservative "overall" approach, based on ranking alpha over all four models, generates average results. In summary, compared with the randomly selected funds portfolio results and the EW portfolio of all funds in the database, the factor models have provided valuable information for hedge fund selection.

For comparison, we also implement a maximum information ratio (MIR) optimization based on the alpha estimates from all four models and the sample covariance matrix, the fund selection being as before. The results are presented in the columns labeled MIR in Exhibit 6. Interestingly, the lack of accuracy of individual alpha estimates results in *lower* information ratios out-of-sample for these portfolios, compared with minimum variance portfolios, even though the objective of the optimization was to achieve the maximum in-sample information ratio. Additionally, the MIR portfolios are less stable so turnover costs are usually higher. Therefore, rather than optimizing on alpha estimates that we know have a high degree of model risk, we are better off just selecting funds based on alpha estimates and then optimizing solely on the covariance matrix.

CONCLUSIONS

Despite the modeling complexity caused by data biases, noisy correlation, inaccurate alphas, and institutional limitations to trading, there is no doubt that alternative investments continue to present attractive opportunities. Their popularity among institutional investors is increasing and the industry requires more academic studies on funds of hedge fund optimization. This article is one such study. Aiming to develop a fund selection and optimal allocation process for funds of hedge funds, we have analyzed the out-of-sample performance of portfolio construction models that target alpha through fund selection based on ranking alpha from factor models. To deal with data biases our database contains both dead and alive funds, we controlled for the instant history bias, and, to neutralize the effect of any selection, survivorship, and multi-period biases on our results, we reported performance on a relative basis, benchmarked against portfolios affected by the same biases.

Since most traditional portfolio construction models require an estimate of expected returns, we have used a number of factor models to estimate the funds' alphas. Despite significant disagreement on individual alpha estimates, the factor models largely agreed on the ranking of funds' alphas. Thus, while the uncertainty in alpha estimates impairs the efficiency of mean-variance optimal portfolios, we have found that the best model for portfolio optimization is based solely on the covariance matrix. The sample covariance matrix had a high element of randomness, but we found no benefit from imposing a factor structure on the correlation matrix, as the sampling errors are already significantly reduced through the weights constraining that is operationally necessary for funds of hedge funds.

When funds are selected by the rank of their alphas from any factor model and the portfolios are optimized to have minimum variance, the performance is superior to that of an equally weighted portfolio of all funds and to that of randomly selected minimum variance optimal portfolios. Of the four factor models considered, all out-of-sample performance measures favor the fund of hedge funds selected using alphas from the PCA factor model. Between January 1998 and May 2003 this had the highest average annual information ratio (6.91), the lowest turnover (24% p.a.), and its returns were also close to normally distributed. Also, for each of the four factor models considered here, the rank alpha minimum variance optimization produced better results than maximum information ratio optimization. We attribute this to the high degree of model risk arising from the factor model dependency of alpha estimates. To conclude, we have shown that, with some refinements, the traditional tools of portfolio management can still be applied to portfolios of hedge funds to achieve excellent results.

ENDNOTES

The authors would like to thank Michiel Timmerman, Lionel Martellini, and Drago Indjic for very helpful comments on the full version of this article, "The Art of Investing in Hedge Funds: Fund Selection and Optimal Allocations," available at www.ismacentre.rdg.ac.uk/dp. Also the financial support from Foundation of Managed Derivatives Research is gratefully acknowledged. All errors remain our responsibility.

[1]Indic [2003] finds 4,589 funds reporting to at least one of the three most important datebase providers, TASS, HFR, and CISDM.

[2]See for instance Black and Litterman [1992], Pastor [2000], and Baks et al. [2001].

[3]The strategy representation in our database is the following: Convertible Arbitrage (10 funds), Distressed Securities (17 funds), Emerging Markets (12 funds), Equity Hedge (53 funds), Equity Market Neutral (26 funds), Equity Non-Hedge (19 funds), Event-Driven (20 funds), Fixed Income (17 funds), Funds of Funds (52 funds), Market Timing (15 funds), Managed Futures (18 funds), Merger Arbitrage (4 funds), Sectors (19 funds).

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

[5]Principal component analysis relies on the quality of the correlation matrix, so we must separate "true" correlation from noise or measurement errors. Especially for hedge fund returns, where the correlation matrix is typically computed on relatively small samples the measurement risk is large and separating real information from noise becomes essential. To this end, we use random matrix theory to compare the properties of the correlation matrix of all the funds in our sample—the empirical correlation matrix—with the properties of a correlation matrix of an identical number of mutually uncorrelated returns series, following a method proposed by Plerou et al. [2002]. Deviations of the empirical matrix from the properties of the random matrix reveal information about "true" correlation between hedge fund returns. We find that only the four largest empirical eigenvalues significantly exceed the range of the random ones and this result is robust to changes in sample periods. This suggests that the "true" correlation pattern can be captured with just the first four principal components and the variation in fund returns relating to the higher eigenvalues is uncorrelated noise.

[6]Results available from the authors on request.

[7]Potential multicollinearity problems were addressed by identifying any pairs of factors that were highly correlated over the sample and dropping the factor having lower correlation with the returns of the fund.

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

[9]We have also investigated the out-of-sample performance of portfolios based on the same alpha selection criterion but with allocations determined by the ordinal rank of the alpha. That is, the higher the alpha estimate, the greater the weight in the portfolio. Several *ad hoc* rank alpha allocation methods were considered. Such portfolios again outperformed the equally weighted portfolios, with information ratios typically between 2 and 4 depending on the factor model and the allocation rule. Also, given the stability in alpha ranks over time, turnover costs were often quite low. But in general their performance was inferior to that of portfolios where a simple "in-out" rank was used for fund selection and allocations were based on minimum variance. More detailed results are available from the authors on request.

REFERENCES

Agarwal, V., and N. Naik. "On Taking the 'Alternative' Route: Risks, Rewards, and Performance Persistence of Hedge Funds." *The Journal of Alternative Investments*, Vol. 2, No. 4 (2000), pp. 6-23.

Alexander, C., and A. Dimitriu. "Indexing, Cointegration and Equity Market Regimes." *International Journal of Finance and Economics*, 10 (2005), pp. 1-19.

Amenc, N., and L. Martellini. "Portfolio Optimization and Hedge Fund Style Allocation Decisions." *The Journal of Alternative Investments*, Vol. 5, No. 2 (2002), pp. 7-20.

——. "The Alpha and Omega of Hedge Fund Performance Measurement." Discussion paper, EDHEC Business School, www.edhec-risk.com, 2003a.

——. “Optimal Mixing of Hedge Funds with Traditional Investment Vehicles.” Discussion paper, EDHEC Business School, www.edhec-risk.com, 2003b.

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

Baks, K., A. Metrick, and J. Wachter. “Should Investors Avoid All Actively Managed Mutual Funds? A Study in Bayesian Performance Evaluation.” *Journal of Finance*, 56 (2001), pp. 45-85.

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

Black, F., and R. Litterman. “Global Portfolio Optimization.” *Financial Analysts Journal*, 48 (1992), pp. 28-43.

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

Dimson, E. “Risk Measurement When Shares Are Subject to Infrequent Trading.” *Journal of Financial Economics*, 72 (1979), pp. 197-226.

Edwards, F.R., and M.O. Caglayan. “Hedge Fund Performance and Manager Skill.” *Journal of Futures Markets*, Vol. 21, No. 11 (2001), pp. 1003-1028.

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

——. “For Better Performance: Constrain Portfolio Weights.” *The Journal of Portfolio Management*, 15 (1988), pp. 29-34.

Fung, W., and D. Hsieh. “Empirical Characteristics of Hedge Fund Strategies.” *Review of Financial Studies*, 10 (1997), pp. 425-461.

——. “Performance Characteristics of Hedge Funds and Commodity Funds: Natural vs. Spurious Biases.” *Journal of Financial and Quantitative Analysis*, 35 (2000), pp. 291-307.

Indjic, D. “Hedge Fund Strategy Classification: AIMA Survey and Analysis of Commercial Classifications.” AIMA Research Day October (2003), Paris, www.aima.org, 2003.

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

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

——. “Bayes-Stein Estimation for Portfolio Analysis.” *Journal of Financial and Quantitative Analysis*, Vol. 21, No. 3 (1986), pp. 279-292.

Lhabitant, F. “Hedge Funds Investing: A Quantitative Look inside the Black Box.” *Journal of Financial Transformation*, 11 (2001), pp. 82-90.

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

Liang, B. “Hedge Fund Performance: 1990-1999.” *Financial Analysts Journal*, 57 (2001), pp11-18.

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

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

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

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

Pastor, L. “Portfolio Selection and Asset Pricing Models.” *Journal of Finance*, 55 (2000), pp. 179-223.

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

Schneeweis, T., and R. Spurgin. “Multifactor Analysis of Hedge Fund, Managed Futures, and Mutual Fund Return and Risk Characteristics.” *The Journal of Alternative Investments,* Vol. 1, No. 2 (1998), pp. 1-24.

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

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