Arik Ben Dor 212-526-7713
abendor@lehman.com
Lev Dynkin 212-526-6302
ldynkin@lehman.com
Tony Gould, CFA 212-526-2821
agould@lehman.com

# STYLE ANALYSIS OF HEDGE FUND RETURNS: ACTUAL VERSUS SELF-PROCLAIMED

## Introduction

Institutional interest in hedge funds as an alternative to investments in traditional asset classes has increased substantially in recent years. Reflecting this trend, Lehman Brothers, in cooperation with HedgeFund.net (HFN), has recently launched a Global Hedge Fund Index.

As with all benchmarks in the Lehman Brothers Global Family of Indices, the Lehman Brothers/HFN Hedge Fund Index is constructed using an objective rules-based set of criteria to determine index eligibility.[1] However, while other indices represent composite returns of individual securities, the Hedge Fund Index reflects the performance of multiple investment strategies (or a single strategy in the case of style sub-indices) employed by the underlying funds.

The minimal disclosure requirements hedge funds face, coupled with an investment mandate that typically allows them to use leverage, short selling, derivatives and highly illiquid securities, present serious challenges for investors. How should they assess the risk/return characteristics of a certain hedge fund strategy in the context of their overall asset allocation? What is the correct approach for comparing the performance (or alpha) of individual hedge funds within the same style?

This article suggests one possible solution through the use of return-based style analysis, introduced in the early 1990s by William Sharpe and used primarily for analyzing mutual funds. This technique provides a way of identifying the asset mix style of a manager and comparing it with the asset mix style of a specified performance benchmark.

We provide a short overview of style analysis and discuss how it may be extended to hedge funds with some modifications. We demonstrate how investors can use its results to better understand the nature of risks and exposures of various strategies, and the extent to which investments in various individual funds are correlated.

Another application of style analysis is the construction of hedge funds portfolios. Lehman Brothers is planning to launch an investable index designed to closely track the broader index. Quantifying the sensitivities of individual funds to common market factors can be used to construct such investable indices in an efficient manner. Alternatively, it may be used to find the composition of portfolios with a minimum volatility for a pre-specified level of expected return.

We also examine the common practice of classifying hedge funds into styles based on their self-reported investment strategy. The actual style of a fund may differ from its self-proclaimed because a style may not be uniquely defined or because the information reported by the fund may be inaccurate. We present a simple technique that can help

[1] To be included in the index, funds must have at least $25 million in assets, a minimum one-year track record, an uninterrupted monthly return time series and an annual audit of fund returns. For more information see *Lehman Brothers/HFN Global Hedge Fund Index Rules*, Lehman Brothers, September 19, 2005.

identify inconsistencies between the actual and self-proclaimed style of a fund based on a measure of distance between the return time-series of the two.

## Return-Based Style Analysis

### Methodology

Factor models are commonly used to characterize how industry and economy-wide factors affect the return on individual securities and portfolios of securities. Sharpe's (1988, 1992) return-based style analysis can be considered a special case of a generic factor model.[2] In this model, we replicate the performance of a managed portfolio over a specified period by the return on a passively managed portfolio of style benchmark index portfolios:

$$(1) \qquad \tilde{R}_{p,t} = [d_{1,p}x_{1,t} + d_{2,p}x_{2,t} + .... + d_{n,p}x_{n,t}] + \tilde{e}_{t,p} \qquad t = 1,2,3,...T$$

where $\tilde{R}_{p,t}$ represents the managed portfolio return at time $t$ and $x_1, x_2...x_n$ are the returns on style benchmark index portfolios. The coefficients $d_1, d_2...d_n$ represent the managed portfolio average allocation among the different style benchmark index portfolios, or asset classes, during the relevant period. The sum of the terms in the square brackets is that part of the managed portfolio return that can be explained by its exposure to the different style benchmarks and is termed the style of the manager. The residual component of the portfolio return -$\tilde{e}_{t,p}$ is the part of the return attributable to the manager security selection ability. It reflects the manager's decision to deviate from the benchmark composition within each style benchmark class.

Given a set of monthly returns for a managed fund, along with comparable returns for a selected set of style benchmark index portfolios, the portfolio weights, $d_1, d_2...d_n$, in Equation (1) can be estimated. However, in order to get coefficients' estimates that closely reflect the fund's actual investment policy, it is important to incorporate restrictions on the style benchmark weights. The following two restrictions are often imposed:

$$(2) \qquad d_{j,p} \geq 0 \quad \forall\ j \in \{1,2,...n\}$$

$$(3) \qquad d_{1,p} + d_{2,p} + .... + d_{n,p} = 1$$

The first restriction corresponds to the constraint that the fund manager is not allowed to take short positions in securities. The second restriction imposes the requirement that we are interested in approximating the managed fund return as closely as possible by the return on a portfolio of passive style benchmarks. For funds known to employ some leverage or short selling, such as hedge funds, other bounds may be imposed.

The goal of return-based style analysis is to find the set of non-negative style-asset class exposures $d_1, d_2...d_n$ that sum to 1 and minimize the "unexplained" variation in returns

[2] Sharpe, William, 1992, "Asset Allocation: Management Style and Performance Measurement," *Journal of Portfolio Management*, 18, 7-19.
— "Determining a Fund's Effective Asset Mix." *Investment Management Review*, December 1988, pp. 59-69.

(i.e., the variance of $\tilde{\varepsilon}_{t,p}$ ) referred to as the fund's tracking error over the style benchmark. The exposures are estimated through the use of quadratic programming since the presence of inequality constraints in (2) does not allow the use of regression analysis. The objective is not to choose style benchmarks that make the fund "look good" or "bad," Rather, the goal is to infer as much as possible about a fund's exposures to variations in the returns of the given style benchmarks during the period of interest.

In the context of style analysis, $R^2$ provides a natural way to distinguish active from passive managers. An active manager is looking for ways to improve performance by investing in asset classes, as well as individual securities within each asset classes that she considers underpriced. Hence, the manager will typically have different exposure to the style benchmark asset classes compared with the performance benchmark. The manager will also be holding a different portfolio of securities within each style benchmark asset class. As a result, the selection component will be lower for passively managed funds than for actively managed funds.

(4) $$R^2 = 1 - \frac{Var(\tilde{\varepsilon}_p)}{Var(\tilde{R}_p)}$$

The right side of Equation (4) equals 1 minus the proportion of variance "unexplained". The resulting $R^2$ value thus indicates the proportion of the portfolio return variance "explained" by the *n* asset classes.[3]

### Application to Hedge Funds

Traditional return-based style analysis is unsuitable for determining the effective style of hedge funds for several reasons. First, unlike mutual funds that follow a defined investment strategy and are limited to investing in specific asset classes, hedge funds have substantial amount of freedom to choose from among a variety of investment strategies. Second, hedge funds can employ leverage and take short positions in securities whereas most mutual funds cannot do so. Third, many hedge funds exhibit an option-like return pattern that is hard to capture with a linear factor model. This arises from the use of derivatives, either explicitly or implicitly through the use of dynamic trading.[4]

To address these challenges, we modify return-based style analysis in several ways. To capture the investment universe available to hedge funds, we use an extensive set of over a 100 factors reflecting the returns to various asset classes, sectors, geographical regions, and currencies (for a compete description of the list of factors, see the appendix). In addition, we alleviate the no short-selling constraints and the requirement that the sum of the estimated coefficients be equal to one.

[3] Since the vector of residuals is not necessarily orthogonal to the matrix of benchmark returns, as is the case in multivariate regression, the alternative definition $R^2 = Var(\delta_{1,p}x_{1,t} + ... + \delta_{n,p}x_{n,t})/Var(R_p)$ is not in general equivalent to the definition given in (4).

[4] The 15%-20% performance-based fee charged by fund managers also contributes to an option-like return profile.

Although lifting the constraints enables the use of multivariate regressions, simply using ordinary least squares is inadequate, given the extended list of factors. Since a priori we have no exact knowledge of which factors should be included in the regressions, estimation with the full list of factors would require a long time-series of returns that is not available for many strategies and in particular for individual funds. The need to estimate the regressions over a relatively short timeframe is also necessitated by the flexibility hedge funds enjoy in their investment mandate. As a result, their style coefficients may be unstable and exhibit large variations across time.

We employ an estimation technique known as stepwise regression in which the variables are entered or removed from the model depending on the significance of the *F*-value (a 5% significance level is used for both inclusion and exclusion). The single best variable is chosen first; the initial variable is then paired with each of the other independent variables, one at a time, a second variable is chosen, and so on until no further variables are included or excluded from the estimation. Stepwise regression thus allows us to examine the importance of a large set of variables even when we have a relatively small number of observations.

When a manager's return is related to the benchmark returns in a nonlinear manner, it would be difficult to identify his performance using linear factor models, of which return-based style analysis is a special case. For example, if investors were to evaluate the performance of a manager selling call options on a standard benchmark by measures such as Jensen's alpha or the Treynor-Black appraisal ratio, such a manager would be falsely classified as a superior performer.[5]

A suggested remedy to this problem that we use in our analysis is to augment the returns on style benchmark indexes with returns on selected options on the style benchmark. We include the returns to six strategies that involve buying put and call options on the S&P500 and holding them to expiration (the exact nature of the strategies is described in the appendix).

### *Empirical Results*

To illustrate the use of return-based style analysis for hedge funds in practice, we use data collected by HedgeFund.net. The dataset includes a total of 2,712 distinct funds and a total of 147,261 monthly return observations for January 1991-June 2004. There are a total of 30 self-reported investment styles. Each fund reports the strategies it employs, and based on this information it is classified into the appropriate style.

Figure 1 presents the style analysis results for each strategy with available return history of at least three years. Monthly returns to a strategy are the equal weighted returns of all funds in that strategy, if at least 30 individual funds' returns are available.[6] Panel A reports the degrees of freedom, number of significant factors, explanatory power ($R^2$), and the variance of the selection components (regression residuals).

[5] For a discussion of this issue, see Ben Dor, Jagannathan and Meier, 2003, "Understanding Mutual Fund and Hedge Fund Styles Using Return-based Style Analysis," Journal of Investment Management, 1, 94-134.

[6] We use equal weighting since data on assets under management is missing for many of the funds. In addition, we do not want the results to be affected by the highly skewed size distribution of hedge funds.

The number of significant factors and the explanatory power of the regression for each strategy vary substantially. For example, roughly 45% of the variation in returns to convertible arbitrage can be explained using five factors, whereas we are able to explain 88% of the return variation in long-short equity using ten factors. Not surprisingly, the return profile of equity market-neutral turns out to be the most difficult to explain because in principle it should have no exposure to systematic factors (the $R^2$ is only 20%). In general, directional strategies are more sensitive to market movements and, therefore, are better explained than arbitrage/market neutral strategies.

Two more points are worth mentioning. The first five factors in order of significance (e.g., order of entry into the estimation) account for almost all the explanatory power of the regression. Hence, the systematic portion of the returns to most strategies can be captured by only five factors (and sometime fewer). Second, strategies with high explanatory power can still have a higher volatility of the regression residuals (the selection volatility shown in the last column) relative to other strategies with lower explanatory power. For example, the selection volatility of equity market neutral and small/micro cap is 0.77% and 2.07%, respectively, although the $R^2$ of the latter is more than four times larger (87% versus 20%). This is simply due to the fact that the overall risk (or return variation) of the two is very different. The market neutral strategy exhibits relatively stable returns and low market exposure, whereas the opposite is true for the small/micro cap strategy.

Panel B of Figure 1 displays the first five factors in order of significance for each strategy and the exposure to them. Convertible Arbitrage, for example, has a beta of -0.07 and -0.09 with respect to the U.S. stocks in the Industrial and Non-cyclical sectors, respectively. In contrast, the strategy has a positive loading on U.S. corporate bonds, emerging markets bonds, and the out-of-the-money call strategy. Indeed, convertible arbitrage often involves owning a convertible security and shorting the underlying stock to hedge the equity component. The exposure to the Industrials and Non-cyclical sectors probably reflects the fact that the companies issuing the convertibles belong primarily to these sectors. The long position in the convertible security is mimicked by the positive exposure to the returns of corporate bonds and call-option strategy.[7]

Figure 2 presents the results of style analysis for individual hedge funds assigned to one of eight investment styles (arbitrage, event-driven, sector, directional, fixed income, global, Commodity Trading Advisors (CTA), and other).[8] The estimation is based on the

[7] The weight of -0.22% on the call option appears to be small. However, it can still have a significant amount of sensitivity to tail events. For example, consider investing $100 in cash and writing 1.2 index put options with an exercise price of $90 and 3 months to maturity when the current index value is $100. If the interest rate is 5% and index volatility is 20% per year, the put option value will be $0.55 based on Black-Scholes. The portfolio will have $100 in T-bills and -$0.66 in index put options, i.e. 100.7% of the funds invested in T-bills and -0.7% invested in out-of-the-money index put options. Suppose the index value drops steeply by 20% to $80 right after forming the position then the position will lose $12, or a 12%. Hence, the position can lose a significant amount in severely depreciating markets even though most of the money is in T-bills.

[8] The investment styles are defined as follows: Arbitrage: Statistical arbitrage, Capital structure arbitrage, Convertible arbitrage, Market neutral equity and other arbitrage.Event driven: Special situations, Reg D, Merger/risk arbitrage, Event driven, Distressed.Sector funds: Value, Technology, Energy, Healthcare, Finance, Small and Micro cap. Directional: Short bias, Long/short equity, Long-only, Market timing. Fixed Income: Fixed income non arbitrage, Fixed income arbitrage, MBS.Global: Macro, Emerging markets, Country specific.CTA: Managed futures. Other: Short term trading, Options strategies, Multi strategy.

period June 2001-June 2004. Only funds that have at least 36 consecutive monthly returns were analyzed. With the exception of CTA, the median number of significant factors for individual funds is 4-5. As before, the explanatory power is higher for funds that employ directional strategies than for funds that use non-directional strategies.

Panel B presents the five factors that are most frequently found to be significant in each style. The exposures to each factor represent the average exposures across all funds within a style. For example, funds in the event-driven style have on average a beta of 0.3 against the returns of the Lehman HY index. This reflects the fact that issuers comprising

Figure 1. **Style Analysis for Hedge Funds' Strategies**

The Number of Observations Used in the Estimation Varies by Strategy; Monthly Returns to a Strategy Are Calculated If at Least 30 Individual Funds' Returns Are Available and Are Equal Weighted; A Description of the Factors Is Provided in the Appendix

Panel A: **Main Statistics**

| Strategy | DF | # of Significant Loadings | Adj $R^2$ (%) All Factors | Adj $R^2$ (%) First 5 Factors | Selection Volatility (Monthly, %) |
|---|---|---|---|---|---|
| Convertible Arbitrage | 78 | 5 | 45.6 | 43.8 | 0.77 |
| Market-Neutral Equity | 63 | 3 | 20.6 | 20.6 | 0.72 |
| Merger Arbitrage | 67 | 3 | 57.2 | 57.2 | 0.82 |
| Event Driven | 76 | 7 | 79.6 | 77.4 | 0.96 |
| Distressed | 59 | 6 | 79.5 | 78.4 | 0.82 |
| Value | 65 | 5 | 86.6 | 86.6 | 1.41 |
| Finance | 62 | 9 | 84.6 | 78.2 | 1.70 |
| Small/Micro Cap | 40 | 7 | 87.3 | 84.1 | 2.07 |
| Long/Short Equity | 73 | 10 | 87.6 | 82.4 | 1.16 |
| Long-Only | 31 | 8 | 98.3 | 97.3 | 0.67 |
| FI Arbitrage | 70 | 7 | 66.4 | 61.9 | 0.79 |
| FI Non-Arbitrage | 65 | 4 | 55.1 | 55.1 | 0.89 |
| Macro | 77 | 6 | 51.7 | 49.8 | 1.42 |
| Emerging Markets | 68 | 12 | 86.2 | 75.3 | 2.77 |
| CTA | 79 | 4 | 30.4 | 30.4 | 2.35 |
| Multi Strategy | 74 | 9 | 61.1 | 46.4 | 0.98 |

Panel B: **First Five Factors in Order of Significance**

| Strategy | 1 | | 2 | | 3 | | 4 | | 5 | |
|---|---|---|---|---|---|---|---|---|---|---|
| Convertible Arbitrage | US Industrial | -0.07 | US NonCyc | -0.09 | US Credit | 0.70 | EM FI | 0.08 | Call OTM | 2.2E-03 |
| Market Neutral Eq | Slope 10-30 | -0.12 | Slope 2-10 | 0.20 | HML | -0.05 | | | | |
| Merger Arbitrage | EM FI | 0.15 | Slope 2-10 | 0.17 | Put OTM Deep | -4.9E-04 | | | | |
| Event Driven | Wilshire 5000 | 0.15 | EM FI | 0.17 | SMB | 0.13 | VIX | -0.08 | US HY | 0.14 |
| Distressed | EM Healthcare | 0.03 | EM FI | 0.15 | SMB | 0.09 | VIX | -0.07 | US HY | 0.21 |
| Value | Wilshire 5000 | 0.65 | EM Tech | 0.05 | SMB | 0.16 | HML | 0.30 | US HY | 0.17 |
| Finance | Wilshire 5000 | 0.27 | EM FI | 0.32 | SMB | 0.39 | Ausy Dollar | -0.27 | Call OTM Deep | 5.0E-03 |
| Small/Micro Cap | Wilshire 5000 | 0.46 | US Healthcare | 0.20 | SMB | 0.56 | Developed NonCyc | -0.31 | EM Telecom | 0.19 |
| Long/Short Equity | Wilshire 5000 | 0.23 | VIX | -0.10 | SMB | 0.22 | HML | -0.14 | EM Telecom | 0.06 |
| Long-Only | Wilshire 5000 | 0.54 | Developed Cyc | 0.14 | SMB | 0.20 | US Treasury | 0.15 | Slope 10-30 | 0.21 |
| FI Arbitrage | Far East HML | 0.07 | Developed Cyc | 0.10 | Global Treas | 0.15 | EM FI | 0.12 | JPY | -0.18 |
| FI Non-Arbitrage | MBS | 0.99 | US Credit | 0.37 | JPY | -0.13 | VIX | -0.09 | | |
| Macro | Slope 10-30 | -0.33 | Developed Cyc | 0.18 | JPY | -0.20 | EM FI | 0.17 | GS Commodity | 0.06 |
| Emerging Markets | EM NonCyc | 0.38 | Far East HML | -0.20 | Global Treas | -0.43 | EM FI | 0.55 | VIX | -0.26 |
| CTA | US Reits | 0.16 | Latin America | 0.06 | US HY | -0.42 | MLM Commodities | 0.56 | | |
| Multi Strategy | Developed Tech | 0.05 | EM FI | 0.12 | VIX | -0.08 | Developed Fin | -0.06 | Slope 2-10 | 0.19 |

this index may be in distress or takeover candidates. Notice also that the positive coefficient may reflect an exposure not only to bonds but also to the issuers' stocks as the returns of the two are highly correlated.[9]

## Hedge Funds Portfolios

A byproduct of the tremendous growth in the hedge funds industry was the creation of so-called funds of funds (FOF), which are essentially portfolios of individual hedge funds. Many institutional investors (and individuals) are interested in getting exposure to hedge funds but do not posses the knowledge to identify funds that suit their investment objectives and evaluate them. FOF can be attractive to such investors since they can offer exposure to a certain style with increased liquidity and reduced risk. In addition, they may have expertise that allows them to identify the better-performing funds in each style.

Several providers of hedge funds indices offer similar products known as investable indices. These are designed to track a broader index closely and provide investors with a practical way of getting exposure to hedge funds.[10]

[9] For further evidence on the relation between stocks and bonds of issuers in the HY index see "Empirical Duration of High-Yield Credit," Global Relative Value, Lehman Brothers, November 8, 2004.

[10] A hedge funds index may be composed of many funds that are closed to new investments or have liquidity provisions that preclude many institutions from investing in them. Lehman Brothers is planning to launch in the near future an 'Investable Index' based on the new Global Hedge Funds Index.

Figure 2. **Style Analysis Results for Individual Hedge Funds**
All Figures in the Table Represent Medians; Only Funds with at Least 36 Consecutive Monthly Returns Are Analyzed; Estimation Is Based on January 1997-June 2004

Panel A: **Main Statistics**

| Style | # of Funds | # of Significant Loadings | DF | Adj. $R^2$ (%) |
|---|---|---|---|---|
| Arbitrage | 149 | 4 | 54 | 39.9 |
| Event driven | 164 | 4 | 53 | 44.7 |
| Sector funds | 137 | 5 | 52 | 65.9 |
| Directional | 402 | 5 | 51 | 61.0 |
| Fixed Income | 95 | 4 | 54 | 41.2 |
| Global | 132 | 4 | 53 | 54.2 |
| CTA | 168 | 3 | 54 | 36.3 |
| Other | 91 | 4 | 51 | 50.8 |

Panel B: **Five Most Frequent Significant Factors**

| Strategy | 1 | | 2 | | 3 | | 4 | | 5 | |
|---|---|---|---|---|---|---|---|---|---|---|
| Arbitrage | LEH Vol Index | -0.11 | US Credit | 0.18 | US HY | 0.42 | EUROPE HY | 0.14 | HML | 0.003 |
| Event Driven | US HY | 0.30 | LEH Vol Index | -0.05 | SMB | 0.34 | Slope 2-10 | 0.18 | EUROPE HY | 0.24 |
| Sector Funds | HML | 0.00 | SMB | 0.66 | US Financial | 0.33 | LEH Vol Index | -0.46 | US Tech | 0.41 |
| Directional | SMB | 0.41 | HML | -0.05 | LEH Vol Index | -0.25 | EM Telecom | 0.21 | Wilshire 5000 | -0.11 |
| FI | MBS | 1.52 | JPU | -0.50 | VIX | -0.08 | Far East HML | 0.20 | Global treas | 1.16 |
| Global | EM Telecom | 0.42 | EM FI | 0.46 | LEH Vol Index | -0.19 | MLM Commodity | 0.30 | US Healthcare | -0.45 |
| CTA | GS Commodity | 1.11 | EUROPE AGG | 0.83 | US Healthcare | -0.41 | US Cyc | -0.38 | GS Commodity | 0.35 |
| Other | EM Telecom | 0.22 | US Healthcare | -0.13 | EUROPE HY | 0.24 | SMB | 0.17 | Slope 2-10 | 0.30 |

In what follows, we demonstrate how the sensitivities of individual funds to common market factors can be used in constructing investable indices or portfolios with a minimum volatility for a pre-specified level of expected return. We compare the results with those of two other approaches that are commonly used: stratification and maximum correlation.

*Constructing Investable Indices*

We start by defining a reference benchmark (an index) that we are interested in replicating. For the purpose of our analysis, the index includes all funds in HedgeFunds.Net database, irrespective of style, size (AUM), available track record, or investment capacity (open or close to additional investments). The index returns are calculated monthly from January 1999-June 2004 as the equal-weighted returns of all funds.

The replicating portfolio is constructed of hedge funds that are part of the eligible universe. The composition of the eligible universe is updated quarterly. A fund is included in the eligible universe in a certain quarter only if it satisfies the following requirements (as of the end of the previous quarter):

1) It is open for additional investments.
2) Has a track record of at least three years (return data must be consecutive).
3) It is larger than the 75th percentile of the size distribution (in terms of AUM).

This last condition attempts to insure that the funds that ultimately comprise the replicating portfolio have adequate capacity in terms of new investments. As of June 2004, the eligible universe included 155 hedge funds with a minimum size threshold of $115 million.

To reflect real practices, the replicating portfolio is composed of at most 40 funds, with no fund having a weight larger than 5%. In order to reflect at least partially hedge funds' liquidity constraints, the replicating portfolio is re-balanced only quarterly.[11] In each re-balancing date, the composition of the replicating portfolio is determined based on one of the following approaches:

i) **Stratification**: The composition of the replicating portfolio mimics the style composition of the broad index. The weight of each style in the index is assigned to the two largest funds in that style that are part of the eligible universe. If a style has no representation in the eligible universe, its weight is divided proportionally among the other styles.
ii) **Maximum correlation**: The objective is to identify the set of funds that minimize the in-sample tracking error relative to the broader index. First, a variance-covariance matrix based on excess returns of all the funds in the eligible universe is constructed using the previous 36 months of data (excess return is defined as the fund return less the broader index return). The weights assigned to each fund are determined in a two-stage process. Initially the optimization is performed using all the funds in the eligible universe. In the second stage, the 40 funds with the largest weights are selected and their weights are re-optimized (which can lead to some of them ending with a zero weight). In essence, this approach is equivalent to treating each fund as a separate factor.

[11] For a discussion of the effects of liquidity in hedge funds portfolios see Lo, Andrew, W., 2005, The Dynamics of the Hedge Funds Industry, CFA Institute, 61-84.

iii) **Style-based optimization**: This approach identifies the optimal replicating portfolio based on a set of risk factors. The risk exposures of both the broad index and the individual funds comprising the eligible universe are estimated as described in the previous section.

The replicating portfolio is found by minimizing overall risk, systematic and idiosyncratic. Systematic risk is measured as the product of a variance-covariance matrix of the risk factors and the respective loadings' estimates. Idiosyncratic risk is represented by a variance-covariance matrix that is calculated using the entire set of residuals of each fund obtained from the regressions. Both matrices are based on the previous 36 months of data used in the estimation. Similar to the maximum correlation approach, weights are determined in a two-stage process: first identifying the set of 40 funds that minimize the overall risk and then recalculating their weights to minimize the tracking error volatility.

Figures 3 and 4 illustrate the use of style analysis for the broad index. Each quarter beginning with December 2001, the index style exposures are estimated using the

Figure 3. **Estimated Style Exposures for Broad Index**
Estimated Quarterly Starting in December 2001 Using the Previous 36 Months

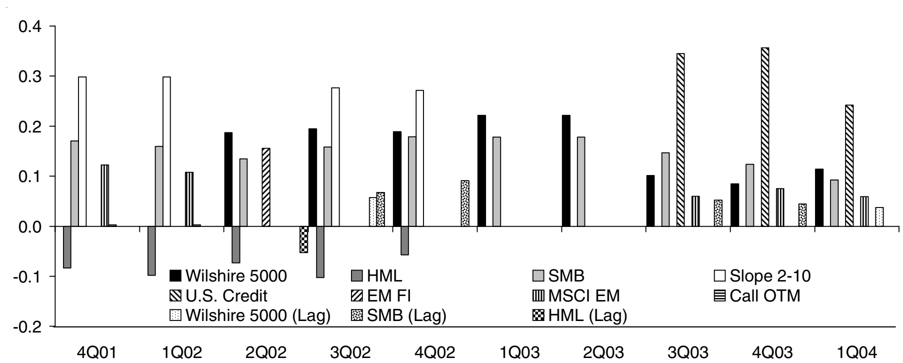
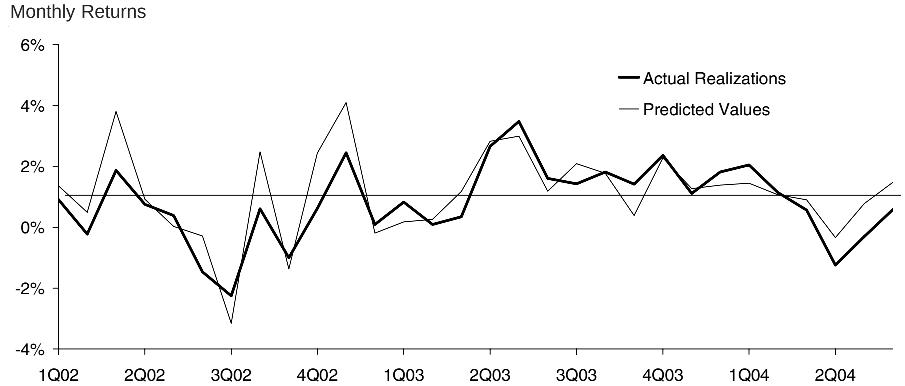



Figure 4. **Realized Returns versus Model Predictions**
Based on Broad Index Returns during January 2002-June 2004;
Results Are Computed Out-of-Sample



previous 36 months of data. Figure 3 plots the in-sample estimated style exposures quarter over quarter. In order to make the estimation robust and more tractable the list of factors used in the analysis is reduced, but it still spans the investment universe.[12] In fact, we find that the reduction in the number of factors improves the regression out-of-sample predictive power.

For example, at the end of 2001, the index had positive exposures to the medium term of the yield curve (10 years), emerging markets, and small-growth stocks and negative exposures to large-value stocks and the short end of the yield curve (two years). The in-sample explanatory power of the model was 80%-90% and was generally higher in the second half of the period.

Figure 4 plots the actual index realizations and respective predicted values using the estimated factor loadings in the three months following the estimation date. The out-of sample tracking error over the entire period (January 2002-June 2004) was 83 bp/month. The improvement in the in-sample explanatory power (e.g. $R^2$) since 1Q03 is also reflected in smaller tracking errors out-of-sample (the far right part of Figure 4).

*Out-of-Sample Performance*

Figure 5 presents the mean and standard deviation of the tracking errors when the replicating portfolio is constructed according to each of the methods described earlier. The figures in the table are based on the out-of-sample realized tracking errors calculated in the three months following the formation date of the replication portfolio.

Using the stratification approach results in a tracking error volatility of 69 bp/month, whereas the maximum correlation approach faired much better with a T.E.V of 48 bp/ month. The style-based optimization achieved the lowest tracking error (44 bp/month), which implies an annual tracking error of 152 bp.

[12] The reduced list includes 5 equity factors (Wilshire 5000, SMB, HML, MSCI EAFE+CANADA and MSCI EM), 6 Fixed Income factors (Aggregate Treasury, two yield curve slopes: 2-10 and 10-30, US credit, Global Treasury ex US and EM), two currencies (JPY, EUR), the GS Commodity Index and two options strategies (described in the appendix). In addition lagged realizations of the following four factors are included: Wilshire 5000, SMB, HML and MSCI EM. Lagged factors were introduced after an analysis of the residuals from regressing the broad index returns against the contemporaneous factors detected signs of autocorrelation. After the lagged variables were introduced, no autocorrelation was detected using the Durbin-Watson test.

Figure 5. **Replication Results by Approach**
January 2002-June 2004; in %
Based on Out-of-Sample Return Realizations
of the Broad Index and Replicating Portfolio

| Approach | Stratification | Maximum Correlation | Style Factors | Modified Style Factors* |
|---|---|---|---|---|
| T.E.V | 0.69% | 0.48% | 0.44% | 0.51% |
| Mean TE | -0.19% | -0.21% | -0.17% | -0.23% |

* Systematic risk is minimized using actual factors rather than individual hedge funds.

The latter result may seem impressive given the fact the broad index includes over 2,000 funds, while the replicating portfolio is composed of no more than 40 funds. However, the results in Figure 5 should in no way be taken to represent actual replication results. Hedge funds data are notoriously affected by survivorship bias: if a fund stops reporting for any reason, its entire return history may be eliminated.[13] Hence, we are ex-ante guaranteed that all funds in the replicating portfolio do not "blow up" in the quarter following their inclusion date. However, since the effect of the survivorship bias is independent of the replication approach, the analysis still provides a good comparison of the three approaches.

The last column in the table (Modified Style Factors) contains results for a variant of the style-based optimization approach. Systematic risk is completely hedged using the actual factors (i.e., the portfolio allocation to the factors corresponds to the estimates of the loadings; typically, this implies positions in 4-5 factors). In contrast, idiosyncratic risk is minimized using the subset of hedge funds that have no exposure to any of the factors (i.e., none of the loadings in the individual regressions is significant, and the $R^2$ is zero).

The approach generates a higher tracking error (51 bp/month) than the "regular" style-based optimization. It seems that the cost involved in reducing the set of funds in the eligible universe (e.g., only funds with no systematic exposures) outweighs the benefit of using fewer loadings' estimates in the optimization (i.e., only for the index and not for the individual funds). However, from a liquidity standpoint, the use of this approach may still prove beneficial since it decreases the number of hedge funds used (on average, only 20-25 funds comprise the replicating portfolio).

Interestingly, the average tracking error is negative (about 20 bp/month) regardless of the replication technique. One explanation is that the index includes many funds that were recently launched, whereas the replicating portfolio mandates a minimum three year's track record. Since many studies argue that hedge funds' performance declines as they mature, the difference in composition can account for this result.[14] Repeating the analysis with the same track record requirement applied to both the index and the replicating portfolio reduces the average negative tracking error by half (about 10 bp/month). Another way to reduce the negative tracking error may be to incorporate explicitly a no-underperformance constraint in the optimization since it decreases the number of hedge funds used (e.g., require the index alpha to equal that of the replicating portfolio in-sample).

### Actual versus Self-Proclaimed Investment Style

The classification of hedge funds by investment style has important implications. Because of the low level of disclosure required from hedge funds, investors use style classification in asset allocation to characterize investment strategies with different risk-return profiles. Similarly, when evaluating the performance of a specific fund, it is typically carried out against a peer group of funds in the same style category.

[13] For a comprehensive discussion of the survivorship bias and other biases in hedge funds' data see Malkiel Burton, G. and Saha, Atanu., 2005, "Hedge Funds: Risk and Return," working paper, Princeton University.

[14] It is not yet clear if this finding reflects capacity constraints (older funds tend to also be larger) or deterioration of skill among 'mature' managers or simply survivorship bias (e.g. "young" hedge funds that prove unsuccessful are closed and subsequently their return data is excluded).

Style classification is usually performed based on the self-reported investment strategy of the fund. However, a fund's actual style may differ from its self-proclaimed. This can happen because in some cases, there is no clear definition of what constitutes a certain style. A fund may also report it employs a different strategy than it does in practice to attract investors.[15]

Although style analysis can be used to compare the actual and self-proclaimed style of a given fund, it has several shortcomings. First, the time-series of returns to the "pure" style itself is typically unavailable. As a result, the appropriate set of factors and exposures that define a certain style is often unclear. Second, even if a style is clearly defined, determining if a fund belongs to it requires specifying acceptable ranges for the exposures to multiple factors. Third, style analysis cannot be applied to funds with relatively short histories that constitute a substantial fraction of the hedge funds universe.

This section presents a different approach to evaluating the accuracy of a self-reported style. The main idea is that the proximity between any two funds (or styles) can be represented using a distance metric that measures the similarity in their performance. Similarly, the distance between a certain fund and a group of peer funds (identified based on the funds' self-declared strategy) can be measured using the return time-series of the two. The advantage of this one-dimensional metric lies in its simplicity and the fact it does not require specifying a set of factors (benchmarks).

### Method and Data

Let $D^T{}_{i,J}$ denote the distance of fund $i$ from strategy $J$ such that,

$$D_{i,J}^{T} = (1-\rho_{i,J}^{T})^2 \tag{5}$$

where $\rho_{i,J}^{T}$ is the correlation between the returns of fund $i$ and strategy $J$ over the last $T$ months. The returns to strategy $J$ are calculated as the time-series of average monthly returns of all funds that report themselves as employing strategy $J$. Depending on the correlation, the value of $D_{i,J}^{T}$ can be 0-4. When the correlation is negative, the distance will be larger than 1, whereas if the correlation is positive the distance is below 1.

Defining the distance based on correlation of returns has several appealing features. First, the correlation measure is intuitive and does not require a long time series to compute. Second, since the degree of leverage is unobservable (but exhibits large cross-sectional variation) the distance should be invariant to the amount of leverage. The division by the standard deviation controls for the effect of leverage. Third, the distance function is convex, which will help to identify funds that are very dissimilar to their self-reported style.

The distance measure can be used to examine the classification of hedge funds using a two-step approach. First, the distance between a fund and its self-reported strategy, which we term 'self-distance' (that is, $D_{i,J}^{T}$ where $i \in J$ ), is calculated. Funds with self-distances

15 For example, if a certain investment style performed well in the past and enjoys positive inflows, a fund employing a different style may claim to be employing that style.

exceeding a given threshold will be candidates for reclassification. The threshold can be specified in absolute terms (e.g., 0.25, which corresponds to a correlation higher or equal to 0.5) or in relative terms (e.g., the top 5% of the funds ranked by self-distance). Second, we calculate the distance between each fund and the rest of the strategies. A fund $i$ will be reclassified from its self-reported strategy $J$ to another strategy $K$ if exceeds the threshold and indicates it is relatively close to strategy $K$ (in a fashion which will be made precise in the next section).

Notice that if a fund's self-distance is below the specified threshold, it would not be re-classified even if its self-distance were much larger than its distance to a certain strategy $K$. Hence, a fund is required to be both dissimilar to its self-reported style and sufficiently similar ("close") to another strategy in order to be re-classified.

To illustrate how the distance measure is employed in practice, we use data from January 1999-June 2004. The first classification is performed in December 1999 and then in six-month intervals until June 2004 (a total of 10 classifications). The distance is computed based on the previous 12 monthly returns. At each classification, only funds that have a complete history over the past year are examined. Monthly returns to a strategy are simply the equal-weighted returns of all funds reporting to be in that strategy. In order to insure meaningful results, only strategies represented by at least 10 funds are included in the analysis.

## Classification Results

### *The Distance between Actual and Self-Reported Style*

Figure 6 presents the upper part of the distance distribution by strategy (median, 75 percentile, 90 percentile, 95 percentile, and maximum) based on 10 classification periods and illustrates that the distribution varies significantly from one strategy to

Figure 6. **Distribution of Distance by Strategy**
Based on Ten Consecutive Classifications with 6-Month Intervals Starting December 1999; Each Classification Is Based on the Past 12 Monthly Returns

| Strategy | # of Funds | $P_{50}$ | $P_{75}$ | $P_{90}$ | $P_{95}$ | Max |
|---|---|---|---|---|---|---|
| Stat Arb | 172 | 0.42 | 0.74 | 1.11 | 1.41 | 1.91 |
| Convert Arb | 712 | 0.08 | 0.23 | 0.50 | 0.75 | 2.34 |
| Market Neutral | 625 | 0.56 | 1.04 | 1.65 | 1.98 | 2.97 |
| Other Arbitrage | 103 | 0.56 | 1.08 | 1.42 | 1.58 | 2.32 |
| Reg D | 43 | 0.35 | 0.79 | 1.45 | 1.67 | 1.75 |
| Risk Arb. | 340 | 0.11 | 0.33 | 0.61 | 0.87 | 1.59 |
| Event Driven | 702 | 0.12 | 0.40 | 0.83 | 1.22 | 2.48 |
| Distressed | 429 | 0.11 | 0.34 | 0.83 | 1.02 | 2.79 |
| Value | 392 | 0.10 | 0.32 | 0.73 | 1.13 | 2.12 |
| Tech | 179 | 0.05 | 0.27 | 0.61 | 1.44 | 3.40 |
| Finance | 410 | 0.14 | 0.46 | 1.31 | 1.83 | 3.07 |
| Small Cap | 282 | 0.08 | 0.28 | 0.62 | 1.04 | 2.07 |
| Short Bias | 142 | 0.03 | 0.11 | 0.34 | 0.66 | 2.02 |
| Long/Short Eq | 3610 | 0.18 | 0.55 | 1.26 | 1.86 | 3.67 |
| Long Only | 221 | 0.04 | 0.15 | 0.30 | 0.51 | 1.31 |
| FI Non-Arb | 384 | 0.21 | 0.48 | 1.00 | 1.45 | 3.38 |
| FI Arb | 542 | 0.42 | 0.79 | 1.16 | 1.79 | 3.03 |
| Macro | 561 | 0.41 | 0.77 | 1.17 | 1.47 | 3.00 |
| EM | 599 | 0.07 | 0.22 | 0.68 | 0.95 | 2.85 |

another and that certain strategies tend to be more cohesive then others. For example, the median distance for convertible arbitrage is 0.08, whereas it is 0.56 for equity market neutral (these figures correspond to correlations of 0.71 and 0.2, respectively). The difference becomes even greater if we look at the 95 percentile (0.75 versus 1.98). In general, the table demonstrates that we should not use a uniform threshold but rather that the threshold should vary by strategy.

The distance distribution is also not stable across time for the same strategy. Figure 7 plots the 95 percentile of the distribution for each of the 10 classification periods for two popular strategies: long/short equity and event-driven. It is clear that the 95 percentile (as well as the rest of the distribution, which is not shown) varies significantly from period to period. In general, we may expect some time variation due to changes in the composition of strategies. In addition, for strategies with a relatively small population of funds, the 95 percentile may be heavily affected by only a few funds.

Clearly, there are many possible ways to determine the potential candidates for reclassification. We use the 95 percentile of the distance distribution determined separately for each strategy and period as the threshold. Any fund with a self-distance that exceeds the 95 percentile will be further examined relative to other strategies.

*Robustness of the Reclassification Criterion*

An effective criterion for selecting candidates for reclassification should identify funds that are very different than their peers (based on their self-reported style). Another desirable feature is that funds that are not selected (e.g., have self-distances below the threshold) are on average “closer” to their strategy than to any other strategy.

Figure 8 reports average distances between the subset of funds not selected for potential reclassification and each of the strategies. If the threshold criterion is effective, then for

Figure 7. **95 Percentile of Distance Distribution by Classification Period**
For Funds with Self-Reported Styles of Event-Driven and Long/Short Equity;
Based on Ten Consecutive Classifications with a 6-Month Interval
Starting on December 1999

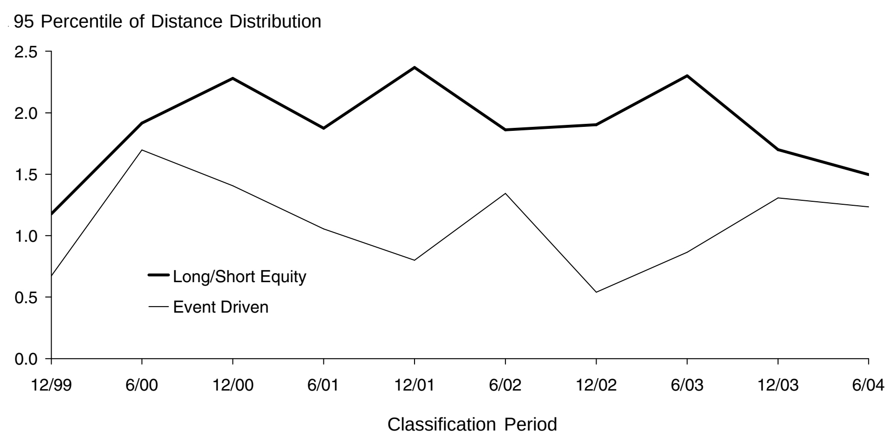


Figure 8. **Average Distance of "Non-Candidate" Funds to Each Strategy**
Using Only the Subset of Funds with Self-Distance Below the 95 Percentile of the Distribution;
Figures in the Table Are Based on Ten Consecutive Classifications with 6-Month Intervals during December 1999-June 2004

| Strategy | Stat Arb | Conv Arb | Mkt Neutral | Other Arb | Reg D | Risk Arb | Event Driven | Distr | Value | Tech | Finance | Small Cap | Short Bias | Long/ Short Eq | Long Only | FI Non-Arb | FI Arb | Macro | EM |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Stat Arb | **0.38** | 0.94 | 0.77 | 0.93 | 0.58 | 1.01 | 0.88 | 0.91 | 0.89 | 0.85 | 0.82 | 0.87 | 1.12 | 0.73 | 0.87 | 0.89 | 0.96 | 0.87 | 1.03 |
| Convert Arb | 0.80 | **0.07** | 0.83 | 0.35 | 0.43 | 0.56 | 0.52 | 0.37 | 0.65 | 0.71 | 0.74 | 0.62 | 1.32 | 0.70 | 0.78 | 0.39 | 0.40 | 0.61 | 0.66 |
| Market Neutral | 0.76 | 0.88 | **0.52** | 0.91 | 0.80 | 0.71 | 0.81 | 0.83 | 0.96 | 0.98 | 0.95 | 0.96 | 1.05 | 0.85 | 0.95 | 0.78 | 0.75 | 0.85 | 0.81 |
| Other Arbitrage | 0.82 | 0.76 | 0.75 | **0.51** | 0.83 | 0.93 | 0.92 | 0.79 | 1.01 | 0.88 | 0.82 | 0.92 | 1.00 | 1.01 | 1.04 | 0.77 | 0.88 | 0.78 | 1.06 |
| Reg D | 0.42 | 0.64 | 0.57 | 0.70 | **0.30** | 0.79 | 0.49 | 0.51 | 0.59 | 0.55 | 0.58 | 0.56 | 1.39 | 0.40 | 0.55 | 0.59 | 0.53 | 0.48 | 0.50 |
| Risk Arb | 0.97 | 0.58 | 0.62 | 0.96 | 0.83 | **0.09** | 0.28 | 0.40 | 0.37 | 0.50 | 0.44 | 0.42 | 1.85 | 0.34 | 0.41 | 0.36 | 0.47 | 0.46 | 0.33 |
| Event Driven | 0.85 | 0.38 | 0.78 | 0.80 | 0.45 | 0.21 | **0.11** | 0.14 | 0.24 | 0.28 | 0.26 | 0.21 | 2.25 | 0.19 | 0.25 | 0.24 | 0.41 | 0.37 | 0.22 |
| Distressed | 0.92 | 0.32 | 0.71 | 0.54 | 0.38 | 0.35 | 0.15 | **0.09** | 0.34 | 0.36 | 0.35 | 0.31 | 1.85 | 0.28 | 0.40 | 0.24 | 0.38 | 0.41 | 0.31 |
| Value | 1.02 | 0.60 | 0.91 | 1.14 | 0.55 | 0.32 | 0.15 | 0.25 | **0.09** | 0.21 | 0.22 | 0.16 | 2.67 | 0.14 | 0.12 | 0.38 | 0.70 | 0.51 | 0.22 |
| Tech | 0.68 | 0.72 | 0.75 | 1.34 | 0.56 | 0.39 | 0.20 | 0.32 | 0.18 | **0.03** | 0.15 | 0.10 | 2.85 | 0.08 | 0.08 | 0.47 | 0.65 | 0.51 | 0.26 |
| Finance | 0.86 | 0.73 | 0.95 | 1.01 | 0.52 | 0.34 | 0.27 | 0.35 | 0.22 | 0.24 | **0.13** | 0.22 | 2.31 | 0.16 | 0.18 | 0.46 | 0.81 | 0.56 | 0.30 |
| Small Cap | 0.84 | 0.56 | 0.74 | 1.18 | 0.41 | 0.34 | 0.12 | 0.19 | 0.15 | 0.13 | 0.17 | **0.07** | 2.54 | 0.09 | 0.10 | 0.32 | 0.53 | 0.37 | 0.23 |
| Short Bias | 1.08 | 1.59 | 1.10 | 0.84 | 1.42 | 2.11 | 2.60 | 2.24 | 3.05 | 2.94 | 2.64 | 2.82 | **0.03** | 3.02 | 3.16 | 1.99 | 1.43 | 1.79 | 2.62 |
| Long/Short Eq | 0.83 | 0.68 | 0.76 | 1.15 | 0.55 | 0.45 | 0.24 | 0.34 | 0.27 | 0.23 | 0.27 | 0.19 | 2.31 | **0.16** | 0.18 | 0.45 | 0.65 | 0.45 | 0.30 |
| Long Only | 0.84 | 0.66 | 0.77 | 1.27 | 0.56 | 0.31 | 0.14 | 0.26 | 0.09 | 0.06 | 0.11 | 0.07 | 3.10 | 0.05 | **0.04** | 0.37 | 0.71 | 0.40 | 0.14 |
| FI Non-Arb | 0.81 | 0.44 | 0.69 | 0.55 | 0.41 | 0.43 | 0.29 | 0.36 | 0.44 | 0.55 | 0.55 | 0.48 | 1.75 | 0.44 | 0.55 | **0.19** | 0.44 | 0.40 | 0.37 |
| FI Arb | 0.92 | 0.70 | 0.73 | 0.78 | 0.72 | 0.68 | 0.68 | 0.73 | 0.86 | 0.87 | 0.91 | 0.87 | 1.10 | 0.87 | 0.92 | 0.57 | **0.39** | 0.70 | 0.70 |
| Macro | 0.83 | 0.75 | 0.66 | 0.70 | 0.61 | 0.61 | 0.51 | 0.64 | 0.73 | 0.76 | 0.67 | 0.62 | 1.38 | 0.58 | 0.67 | 0.47 | 0.62 | **0.36** | 0.53 |
| EM | 0.84 | 0.51 | 0.65 | 0.96 | 0.55 | 0.28 | 0.15 | 0.22 | 0.18 | 0.27 | 0.26 | 0.19 | 2.48 | 0.15 | 0.17 | 0.25 | 0.45 | 0.33 | **0.06** |

Figure 9. **Ratio of Each Distance to Self-Distance by Strategy**
The 25 Percentile of the Distribution of the Ratio is Reported for Each Combination of Two Strategies; Using Only the Subset of Funds with Self-Distance Below the 95 Percentile of the Distribution; Figures Are Based on Ten Consecutive Classifications with 6-Month Intervals during December 1999-June 2004

| Strategy | Stat Arb | Conv Arb | Mkt Neutral | Other Arb | Reg D | Risk Arb | Event Driven | Distr | Value | Tech | Finance | Small Cap | Short Bias | Long/ Short Eq | Long Only | FI Non-Arb | FI Arb | Macro | EM |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Stat Arb | 1.00 | 1.15 | 1.10 | 1.39 | 1.02 | 1.02 | 1.08 | 1.04 | **0.95** | **0.73** | **0.71** | **0.85** | 1.37 | **0.90** | **0.94** | 1.20 | 1.08 | 1.08 | 1.05 |
| Convert Arb | 3.63 | 1.00 | 4.41 | 3.06 | 4.09 | 3.01 | 2.49 | 2.26 | 3.10 | 3.74 | 3.55 | 3.05 | 6.35 | 3.51 | 3.88 | 2.84 | 2.80 | 3.81 | 4.06 |
| Market Neutral | **0.75** | **0.76** | 1.00 | **0.95** | **0.83** | **0.67** | **0.69** | **0.74** | **0.68** | **0.76** | **0.77** | **0.77** | **0.82** | **0.77** | **0.77** | **0.74** | **0.76** | **0.84** | **0.70** |
| Other Arb | **0.75** | **0.76** | **0.67** | 1.00 | **0.72** | 1.05 | **0.72** | **0.88** | **0.75** | **0.65** | **0.69** | **0.78** | 1.01 | **0.67** | **0.95** | **0.75** | **0.81** | **0.76** | 1.02 |
| Reg D | 1.11 | 1.33 | 1.08 | 1.02 | 1.00 | **0.73** | **0.67** | **0.85** | 1.19 | 1.28 | **0.64** | 1.02 | 1.66 | **0.99** | 1.10 | 1.04 | 1.00 | **0.77** | **0.83** |
| Risk Arb | 3.15 | 2.25 | 2.12 | 2.69 | 1.90 | 1.00 | 1.38 | 1.93 | 1.85 | 2.11 | 1.98 | 1.92 | 4.95 | 1.92 | 2.19 | 1.76 | 2.12 | 2.00 | 1.44 |
| Event Driven | 2.13 | 1.40 | 1.57 | 2.26 | 1.73 | **0.79** | 1.00 | **0.68** | **1.00** | 1.06 | 1.09 | **0.98** | 4.93 | **0.94** | 1.07 | 1.07 | 1.40 | 1.34 | 1.11 |
| Distressed | 2.87 | 1.30 | 2.13 | 2.69 | 2.34 | 1.52 | 1.06 | 1.00 | 1.48 | 1.70 | 1.64 | 1.47 | 4.76 | 1.36 | 1.61 | 1.18 | 1.70 | 1.84 | 1.47 |
| Value | 3.25 | 2.15 | 2.62 | 4.37 | 1.93 | 1.31 | **0.92** | 1.19 | 1.00 | 1.12 | 1.04 | **0.84** | 8.26 | **0.78** | **0.86** | 1.64 | 2.53 | 1.87 | 1.07 |
| Tech | 3.57 | 3.09 | 3.72 | 5.33 | 4.17 | 2.58 | 1.97 | 2.11 | 1.61 | 1.00 | 1.50 | 1.23 | 12.49 | **0.95** | **0.99** | 4.10 | 3.86 | 2.78 | 2.23 |
| Finance | 1.92 | 1.88 | 1.96 | 2.80 | 1.69 | **0.99** | **0.91** | 1.13 | **0.80** | 1.18 | 1.00 | **0.99** | 5.97 | **0.88** | **0.93** | 1.38 | 2.21 | 1.67 | **0.91** |
| Small Cap | 3.28 | 2.45 | 2.88 | 4.20 | 2.51 | 1.54 | **0.78** | 1.19 | **0.95** | **0.98** | 1.00 | 1.00 | 8.88 | **0.78** | **0.87** | 1.78 | 2.23 | 1.53 | 1.31 |
| Short Bias | 14.77 | 19.11 | 14.22 | 8.71 | 9.24 | 17.14 | 27.22 | 22.30 | 27.07 | 27.92 | 22.93 | 26.92 | 1.00 | 27.45 | 30.38 | 23.34 | 15.10 | 20.85 | 25.06 |
| Long/Short Eq | 1.66 | 1.49 | 1.41 | 2.07 | 1.31 | 1.02 | **0.83** | 1.04 | **0.85** | **0.87** | **0.95** | **0.88** | 3.44 | 1.00 | **0.83** | 1.15 | 1.39 | 1.09 | **0.95** |
| Long Only | 7.30 | 6.47 | 5.96 | 10.52 | 6.83 | 3.58 | 1.62 | 2.45 | 1.14 | 1.10 | 1.51 | 1.12 | 27.40 | **0.82** | 1.00 | 4.04 | 7.71 | 2.86 | 1.86 |
| FI Non-Arb | 1.55 | 1.17 | 1.40 | 1.15 | 1.25 | 1.21 | **0.61** | **0.74** | **0.85** | 1.04 | 1.01 | **0.76** | 2.86 | **0.74** | **0.80** | 1.00 | 1.17 | **0.97** | **0.88** |
| FI Arb | 1.03 | **0.97** | 1.01 | **0.96** | **0.83** | **0.92** | **0.95** | **0.96** | 1.18 | 1.11 | 1.02 | 1.06 | 1.14 | 1.13 | 1.19 | **0.88** | 1.00 | **1.00** | 1.00 |
| Macro | 1.12 | 1.11 | **1.00** | 1.12 | 1.16 | 1.00 | **0.92** | **0.97** | 1.06 | 1.13 | **0.98** | **0.81** | 1.23 | **0.76** | **0.82** | **0.83** | **0.99** | 1.00 | **0.95** |
| EM | 3.94 | 2.67 | 2.62 | 3.40 | 2.47 | 1.34 | **0.98** | 1.47 | 1.35 | 1.61 | 1.48 | 1.35 | 11.58 | 1.13 | 1.13 | 1.20 | 2.56 | 1.60 | 1.00 |

each row of the matrix the figures on the diagonal should be the smallest. Indeed, this is what we observe in practice. For example, the average self-distance of funds in the statistical arbitrage strategy is 0.38, whereas the distance of these funds to the second nearest strategy is 0.73 (long/short equity). This result can be expected because statistical arbitrage is based on taking offsetting positions in closely related securities (often stocks) that trade at very different prices.

Despite the results in Figure 8, some funds are closer to a different strategy than the one they report. To see this, we calculate for each fund its distance to a certain strategy divided by its self-distance. Figure 9 reports the 25 percentile of the distribution of this ratio. If the self-distance is always smaller than any other distance, the 25 percentile of the distribution would be above 1.

However, as Figure 9 illustrates, in many cases more than 25% of the funds in a certain strategy are actually closer to another strategy(such cases are marked by bold figures). This gives us a good sense of the cohesiveness of each strategy or level of diversification it offers. For example, looking at risk-arbitrage or short-bias funds, we see that over 75% of the funds are closer to their strategy than to any other strategy. In contrast for the case of market neutral, there are always at least 25% of the funds that are closer to any of the other strategies.

### *Re-classifying Funds*

For a fund to be reclassified, we require not only that its self-distance is greater than the specified threshold. It should also be relatively “close” to at least one other strategy to which it will be reclassified. For the set of funds identified as candidates for reclassification, Figure 10 reports (by strategy) the distance to the closest strategy, the self-distance,

Figure 10. **Summary Statistics for Re-Classified Funds**
All Figures Represent Medians; Based on Ten Consecutive Classifications with 6-Month Intervals during December 1999-June 2004

| Strategy | # of Reclassified Funds | Distance to Closest Strategy | Self-Distance | Ratio of Distance to Closest Strategy over Self-Distance |
|---|---|---|---|---|
| Stat Arb | 12 | 0.25 | 1.44 | 0.18 |
| Convert Arb | 41 | 0.24 | 0.84 | 0.30 |
| Market Neutral | 36 | 0.37 | 2.25 | 0.19 |
| Other Arbitrage | 9 | 0.25 | 1.61 | 0.18 |
| Reg D | 3 | 0.42 | 1.67 | 0.25 |
| Risk Arb | 24 | 0.16 | 0.79 | 0.23 |
| Event Driven | 43 | 0.27 | 1.34 | 0.25 |
| Distressed | 29 | 0.29 | 1.09 | 0.25 |
| Value | 24 | 0.23 | 1.15 | 0.24 |
| Tech | 13 | 0.15 | 1.65 | 0.08 |
| Finance | 26 | 0.23 | 1.83 | 0.15 |
| Small Cap | 18 | 0.15 | 1.01 | 0.18 |
| Short Bias | 11 | 0.32 | 0.38 | 1.00 |
| Long/Short Eq | 184 | 0.16 | 2.36 | 0.07 |
| Long only | 16 | 0.16 | 0.56 | 0.30 |
| FI Non-Arb | 24 | 0.18 | 1.57 | 0.16 |
| FI Arb | 30 | 0.46 | 1.94 | 0.20 |
| Macro | 32 | 0.40 | 1.83 | 0.20 |
| EM | 34 | 0.32 | 1.43 | 0.30 |

and the ratio of the two (calculated individually for each fund; the figures in the table represent medians).

Except for short bias, the figures in the last column are all substantially smaller than 1. This indicates that for funds identified as candidates for reclassification, we can find strategies that their return histories resemble much more than their self-reported strategy. Based on Figure 10, we decided to reclassify a fund to a different strategy if it is the closest to the fund (relative to all other strategies) and the distance to it is smaller than the self-distance.

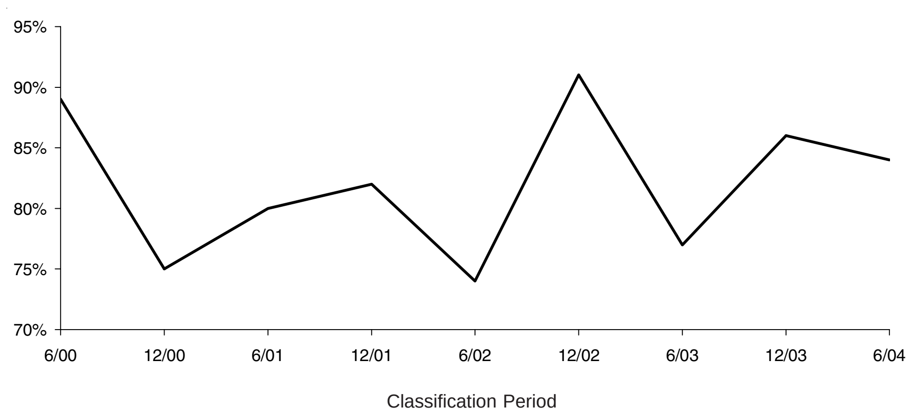
A desirable feature of any classification scheme is stability. In our context, we think of stability in the following sense: if a fund is re-classified at time *t* to a certain strategy, then the results of the next classification (at *t+6*) should point to the same strategy. Figure 11, plots by period the percentage of funds that were reclassified at time *t* but were not reclassified again at *t+6* (i.e., these funds remained in their new assigned strategy). The results suggest that the classification scheme is stable because, on average, 80% of the funds that are reclassified at time *t* remain in their newly assigned strategy following the next classification.

## Conclusion

Over the last decade, hedge funds have probably been the fastest growing sector in the financial services industry, with over $1 trillion in assets currently under management. Their growth was largely fueled by the view that they are consistently able to generate positive alphas and provide diversification benefits due to the nature of the strategies they employ.

Hedge funds' investment mandates typically allow them to use leverage, short selling, derivatives, and highly illiquid securities. These characteristics can generate return profiles that exhibit auto-correlation, fat tails, options-like payoffs, and unstable

Figure 11. **Stability of Classification**
% of Reclassified Funds at Time *t* Remaining in Their New Assigned Strategy at Time *t*+6; Based on Ten Consecutive Classifications with 6-Month Intervals during January 1999-June 2004



correlations with other asset classes. The lack of almost any investment constrains, coupled with the minimum disclosure hedge funds offer (due to the lax of regulatory oversight), presents serious challenges for investors. Are the returns to a certain strategy indeed uncorrelated with the returns to other assets in their portfolios? How should the left tail-tail risk of a strategy with a relatively short return history be measured? What is the correct approach for comparing the performance (or alpha) of individual hedge funds within the same style?

This article discusses several quantitative tools that can help investors address some of these issues. We start by showing how return-based style analysis, originally applied to mutual funds, may be extended to hedge funds with some modifications. We demonstrate how investors can use the factor exposures estimates to better understand the nature of risks and exposures of various strategies and the extent to which investments in various individual funds are correlated. In addition, style analysis can be used to measure the alpha or added value of a certain investment style or individual fund.

Style analysis can also be used in constructing investable indices or portfolios of hedge funds. We construct an investible index using three distinct approaches: stratification, maximizing in-sample correlation, and based on factor exposures estimated using style analysis. We find that under realistic conditions and using actual hedge funds data, the style-based optimization achieves the lowest out-of-sample tracking error volatility.

Another important issue we examine is the investment style classification of hedge funds. The style assigned to a fund carries major implications for asset allocation and performance evaluation against peers. Yet in practice, style classification commonly relies on hedge funds' self-reported investment strategies. We present a simple technique that is able to identify inconsistencies between a fund's actual and self-reported strategies and reclassify it to another strategy that it more closely resembles (based on its return history). Despite the short data history we use in the analysis, which does not allow to us to perform traditional statistical tests, we find that the technique generates stable and meaningful results.

## Appendix

Figure 12 presents the list of factors used in the style analysis of hedge funds, separately by asset class. The equity component is broken down by geography (U.S., developed countries excluding U.S. and emerging markets) and sector (ten basic industries) using the Dow Jones indices. The Wilshire 5000 and MSCI EAFE+Canada serve as proxies for the aggregate U.S. market and developed countries (ex-U.S.), respectively. Emerging markets are broken down further to three regions: Latin America, Europe and Middle East, and the Far East. We also construct value-growth and size factors for the U.S. using the six Wilshire sub-indices (Small-Value, Large-Growth, etc.). For the other developed countries and emerging markets, we use the respective MSCI-value index less the MSCI-growth index.

The fixed income component is similarly broken into three regions but has less detailed coverage. All factors, except the 1-month LIBOR rate, which serves as the risk-free rate, are based on Lehman indices. The U.S. market is represented by the Lehman Treasury, Credit, MBS, and HY indices (the return to all spread asset classes are in excess of the duration-matched treasury returns). In addition, variations in the slope of the yield curve are modeled using two factors that are actually returns to long-short strategies: 10-year less 2-year and 30-year less 10-year (both strategies are duration neutral). For the non-U.S. markets, we use as factors the Global Treasury (ex-U.S.), Euro Aggregate, Euro HY, and EM aggregate indices.

Additional factors are the returns to four currencies (JPY, EUR, Swiss Franc, and Australian Dollar), a commodity index, and two measures of implied volatility (VIX and LEH volatility index). We also include the returns to six options strategies that involve buying a one-month put or a call on the S&P500 and holding it until expiration.[16] Finally, we include lagged factor realizations to control for possible serial correlation.[17]

[16] We compute the returns to holding at-the-money (ATM), out of the money (OTM) and deep OTM Calls and Puts. ATM, OTM and deep OTM are defined based on the strike price being equal to the index price at the time of purchase, index price + 0.5 std and index price + 1 std, respectively.

[17] See, for example, Asness, C., Krail, R. and J. Liew, 2001, "Do Hedge Funds Hedge?," *The Journal of Portfolio Management*, 28, 6-19.

Figure 12. **List of Factors for Return-Based Style Analysis**

**Equities**

| | U.S | Developed Countries Excl. U.S. | EM | |
|---|---|---|---|---|
| Agg Market | Wilshire 5000 | MSCI EAFE+Canada | MSCI EM | Separately for: Latin America, Europe and Middle East, Far East |
| Value-Growth (HML) | Wilshire | MSCI EAFE+Canada | MSCI EM | 1/3[(large value-large growth)+ (mid value-mid growth)+ (small value-small growth)] |
| Small-Big (SMB) | Wilshire | | | 1/2[(large value - small value) + (large growth -small growth)] |
| <u>Sector</u> | | | | |
| Basic Materials | DJ | DJ | DJ | |
| Industrials | DJ | DJ | DJ | |
| Cyclical | DJ | DJ | DJ | |
| Non Cyclical | DJ | DJ | DJ | |
| Financials | DJ | DJ | | |
| Energy | DJ | DJ | DJ | |
| Healthcare | DJ | DJ | DJ | |
| Technology | DJ | DJ | DJ | |
| Telecom | DJ | DJ | DJ | |
| Utilities | DJ | DJ | DJ | |
| REITs | DJ | | | |

**Fixed Income (All Lehman Brothers Indices)**

| | U.S | Developed Countries Excl. U.S. | EM |
|---|---|---|---|
| Risk-Free Rate | 1-Month Libor | | |
| Treasury | US Treasury | Global Tresury ex. U.S. | |
| Credit (Excess Returns) | US Credit | Euro. Agg. | EM Agg. |
| MBS (Excess Returns) | MBS | | |
| HY (Excess Returns) | US HY | Euro. HY | |
| Yield-Curve Slope | 2 - 10, 10 - 30 | | |

**Other**

| | |
|---|---|
| Currencies | JPY, EUR, Swiss Franc, Australian Dollar |
| Commodities | Gold, Oil, GS commodity index |
| Volatility | VIX (equity), Lehman vol index (swaptions) |
| Options | S&P500 Call and Put (ATM, OTM, deep OTM) |

The views expressed in this report accurately reflect the personal views of Lev Dynkin, Arik Ben Dor, and Tony Gould, the primary analyst responsible for this report, about the subject securities or issuers referred to herein, and no part of such analyst's compensation was, is or will be directly or indirectly related to the specific recommendations or views expressed herein.

Any reports referenced herein published after 14 April 2003 have been certified in accordance with Regulation AC. To obtain copies of these reports and their certifications, please contact Larry Pindyck (lpindyck@lehman.com; 212-526-6268) or Valerie Monchi (vmonchi@lehman.com; 44-(0)207-102-8035).

Lehman Brothers Inc. and any affiliate may have a position in the instruments or the companies discussed in this report. The firm's interests may conflict with the interests of an investor in those instruments.

The research analysts responsible for preparing this report receive compensation based upon various factors, including, among other things, the quality of their work, firm revenues, including trading, competitive factors and client feedback.

Lehman Brothers usually makes a market in the securities mentioned in this report. These companies are current investment banking clients of Lehman Brothers or companies for which Lehman Brothers would like to perform investment banking services.

*Publications-L. Pindyck, B. Davenport, W. Lee, D. Kramer, R. Madison, A. Acevedo, M. Graham, V. Monchi, K. Banham, G. Garnham, Z. Talbot*