The European Journal of Finance
Vol. 13, No. 7, 645–655, October 2007

Factor-based, Non-parametric Risk Measurement Framework for Hedge Funds and Fund-of-Funds

T. R. J. GOODWORTH** & C. M. JONES*,**
*Centre for Financial Research, Judge Institute, University of Cambridge, Cambridge, UK, **Key Asset Management (UK) Ltd, London, UK*

ABSTRACT *A factor-decomposition based framework is presented that facilitates non-parametric risk analysis for complex hedge fund portfolios in the absence of portfolio level transparency. This approach has been designed specifically for use within the hedge fund-of-funds environment, but is equally relevant to those who seek to construct risk-managed portfolios of hedge funds under less than perfect underlying portfolio transparency. Using dynamic multivariate regression analysis coupled with a qualitative understanding of hedge fund return drivers, one is able to perform a robust factor decomposition to attribute risk within any hedge fund portfolio with an identifiable strategy. Furthermore, through use of Monte Carlo simulation techniques, these factors can be employed to generate implied risk profiles at either the constituent fund or aggregate fund-of-funds level. As well as being pertinent to risk forecasting and monitoring, such methods also have application to style analysis, profit attribution, portfolio stress testing and diversification studies. This paper outlines such a framework and presents sample results in each of these areas.*

KEY WORDS: Hedge fund, fund-of-funds, risk, non-parametric, value-at-risk, multi-factor, Monte Carlo

1. Introduction and Literature Review

Hedge funds are a class of investment vehicle that aim to generate market independent returns by utilizing a range of non-traditional investment techniques and investing across a range of markets. Institutional and individual investors generally access hedge funds through hedge fund-of-funds, since they offer diversified access to a range of external hedge funds, which are selected and then monitored by the hedge fund-of-fund manager. For such a hedge fund-of-funds manager, risk monitoring of hedge funds can be difficult, not least because of issues surrounding transparency, liquidity, and pricing risk. It is standard practice for most hedge funds to produce partially transparent reports with respect to the underlying portfolio, and even when full transparency is possible this is usually only a monthly snapshot. Furthermore, even with full transparency, the portfolio risk is often difficult to measure as a result of the instruments within. For example the portfolio may contain OTC instruments with only indicative prices or difficult-to-value derivatives. Model risk is also prevalent in hedge fund strategies. Indeed, traditional value-at-risk (VaR) style risk

Correspondence Address: C. M. Jones, Key Asset Management Ltd., 34 Brook Street, London, W1K, 5DN, UK. Email: 1.chrisjones@gmail.com

1351-847X Print/1466-4364 Online/07/070645–11 © 2007 Taylor & Francis
DOI: 10.1080/13518470701322284

measures will often mis-estimate the risk in a hedge fund portfolio. Many such portfolios will be constructed by their managers to have carefully matched long and short positions in pairs or baskets as a hedge and so will have low VaR. However, this may not reflect the risk in the future since often such positions are event-driven with a severe covariance structure altering event yet to come. Similarly, a manager may invest in such a position straight after the event which, under historic techniques, would yield a scenario which reported a higher VaR within a lower risk environment. On a more pragmatic note, the complexity in many hedge fund portfolios can be significant, meaning there is just not enough time for analysts to accurately measure the risk of many hedge fund portfolios over a month.

The above views echo those of Lo (2001), whose review paper considers a wide range of risk management techniques applied over the last two decades. Lo then proceeds to highlight the shortcomings of mean-variance, historic VaR, and beta analysis techniques with respect to hedge fund analysis as a direct result of the complexity of hedge fund portfolios. Fung and Hsieh (1999) have also shown that normal mean-variance risk measurement techniques do not work for hedge funds due to the dominance of non-normal returns. To avoid distribution-based pitfalls such as underestimation of tail risk, this paper presents a method that is both non-parametric and forward-looking wherever possible. Furthermore, to avoid the problem of portfolio transparency, we have assumed zero position-level transparency at the underlying fund level and have instead built a factor-based model. This has been enhanced by the qualitative choice of factors based on our knowledge of the drivers of hedge fund returns for each strategy. Such a factor-based approach is also useful for hedge fund-of-funds risk measurement and portfolio construction.

Using factors to understand hedge fund returns and risk is not a new approach. Based on Sharpe's (1992) style analysis, Fung and Hsieh (1997), Agarwal and Naik (2000a, 2000b), and more recently Assness *et al.* (2004a, 2004b), and Ibbotson and Chen (2005) have used a simple factor based approach to help understand how hedge funds generate returns, with Brealey and Kaplanis (2001) applying this approach specifically to understand hedge fund risk. Fung and Hsieh (2002) extended the factors used in such analysis to asset-based style (ABS) factors, which the authors considered to be a better descriptor of what hedge fund managers do. However, given the active nature of such hypothetical factors, an ABS factor is inappropriate for risk measurement in this context. As a result, recent papers by Fung and Hsieh (2004a, 2004b) looking at hedge fund risk using asset-based style factors are of interest but not extendable for this style of risk measurement. Finally, the most pragmatic and active use to date of style factor analysis in hedge funds is by Amenc *et al.* (2003), where this analysis is applied to tactical asset allocation for hedge funds.

As noted above, there are a number of challenges in the hedge fund-of-funds environment that prevent standard risk management frameworks being used accurately. Mostly, hedge fund investors are not privileged to total portfolio level transparency from their investments, but existing risk frameworks that are appropriate for hedge funds rely on such transparency. Furthermore, such risk frameworks usually assume a standard mean-variance environment or at least the existence of a standard return distribution for the underlying funds, even though this has been shown to be not the case by Fung and Hsieh (1999).

The primary objective of this paper is to present a risk measurement framework that is not reliant on underlying portfolio transparency and that is not dependent on the classic mean-variance framework for returns. We have chosen to develop the above risk framework through an approach based on factor decomposition of returns. This will give both depth and understanding of portfolio construction beyond basic time series analysis of underlying fund returns and can be applied in the absence of portfolio transparency. Such methods are well established for investigating financial time series, as discussed above. Consequently and to the best of our knowledge, this paper is the first to extend factor-based analysis of hedge fund returns in order to form a risk evaluation framework that should better estimate tail risk. The paper continues with a discussion of methods employed (Section 2), followed by the presentation of sample results in Section 3, where the paper is also summarized and further work considered.

2. Method

In this section we outline a prescriptive method necessary to analyse a hedge fund or, by aggregation of funds, a hedge fund-of-funds portfolio within a non-parametric factor-based framework. We begin by considering the criteria that a factor set should satisfy. This is followed by a presentation of the regression and simulation techniques employed and the results they generate.

2.1 Selecting a Factor Set

In line with the transparency requirements discussed above, the time series of a fund's returns forms the basis for analysis. Hedge fund reporting schedules and concerns regarding the reliability of intra-month pricing of less liquid portfolios dictate that we employ monthly data for portfolio returns and consequently also for associated factor data.

At a hedge fund-of-funds level, the issue of transparency is often mitigated as the institutionalization of this space has demanded greater look-through transparency from investors, which inherently includes fund composition and constituent manager weightings. As such, in this study, we consider a fund-of-funds portfolio to comprise the weighted set of constituent fund time series as oppose to a single aggregate time series.

Within a non-transparent reporting framework, a set of fundamental standardized factors can be adopted as a proxy for asset-specific risk analysis at the hedge fund portfolio level, see for example Fung and Hsieh (1997), Ibbotson and Chen (2005), and Brealey and Kaplanis (2001). In the present study, factor time series returns have been standardized to a unit standard deviation without the removal of outliers. The use of a comprehensive factor set not only permits conventional style analysis, but also provides an essential mechanism for cross-strategy fund analysis at the hedge fund-of-funds level, thus simplifying a very complex task when viewed from an instrument level perspective.

In constructing the factor universe, there should be enough breadth in factors to robustly describe all potential underlying portfolios within the hedge fund-of-funds, while minimizing factor redundancy at a conceptual level and co-linearity in the quantitative sense. For example, MSCI Europe value and growth indices would have significant correlation. Instead the use of a single factor describing the relative difference between the pair (i.e. a delta factor) would avoid such a problem. For example, a market-independent fund-of-funds universe, constructed according to instrument, region and duration, would typically contain 50 factors in order to reach a suitable level of descriptivity. To a certain extent, the factor universe is dependent on the range of underlying strategies employed within the hedge fund-of-funds, but it is also qualitatively dependent on a risk manager's interpretation of the drivers of return, which may vary over time as hedge fund strategies evolve. As such, inclusion of a prescriptive list of factors in this paper would be inappropriate. A satisfactory level of description for a broad-based multistrategy hedge fund-of-funds may be achieved, at the time of writing, with a working factor universe of approximately 100 factors, covering major themes including; equity, debt, credit, volatility, commodities, currency, as well as corporate event data. Within each of these broad themes, both factor and geographical subsets provide additional layers of granularity. For example; convertibles, high yield, sovereign debt, asset-backed and CBO/CDO factors along with yield curve, interest rate and issuance data form a core component of the debt theme.

It is important to emphasise that factor descriptivity can be enhanced greatly by knowledge of the nature of the hedge fund portfolio. This understanding can be used to differentiate between factors that should fundamentally drive returns and those which are merely highly correlated. Importantly, the authors believe that this qualitative contextualization does not conflict with earlier assumptions regarding portfolio transparency. Data contemporaneity is also critical for any ex-ante risk management framework, to this extent data is sourced according to both quality and frequency. In certain circumstances (e.g. CPI data), calculation methods only permit use of monthly or quarterly lagged data, however contemporaneity is preserved as this is the most recent data available to the market.

Having identified a factor universe, a primary quantitative screen is employed to assess the universe for the impact of highly correlated factors. Unlike principal component analysis, multivariate linear regression (used to determine factor loadings to the underlying fund) does not dictate that factors form a mutually orthogonal basis, in which case these factors cannot be treated independently. Furthermore, factors possessing a high degree of co-linearity can significantly affect regression factor loadings even if the co-linearity is transient. Structural co-linearity can be removed from the factor universe through qualitative considerations. For example MSCI Europe and DAX indices are co-linear by composition, as the German equity index constituents are a significant component of MSCI Europe. Non-structural co-linearity is handled via correlation matrix screening, excluding factors with a correlation coefficient above a given ceiling level $|\rho_{\max}|$. In our empirical studies of a wide range of hedge fund strategies it has been found that $|\rho_{\max}| = 0.8$ provides a suitable balance between screening co-linear factors and creating an underspecified factor universe.

The screened factor universe can subsequently be used to select a factor sub-set to describe the fund. This process is described in the following sections, however the final outcome is a fund basis set of typically 5–8 factors. These describe the core drivers of risk and hence potential return in the portfolio. Over-specification of this sub-set will of course inherently improve the initial regression analysis, but will subsequently reduce the information content of the resulting factor decomposition as factors become less significant through a reduction in the number of degrees of freedom associated with the system. Furthermore a large factor subset is more likely to encounter the orthogonality constraints discussed above during construction.

2.2 *Establishing Quantitative Factor Significance*

In order to determine quantitatively significant factors, the factor universe is screened individually via an F-ratio test. Rather than relying on a conventional R^2 statistic, use of an F-ratio generated by a correlation coefficient transform (Koepf, 1998) permits exclusion of statistically insignificant correlated factors as well as uncorrelated factors (Barlow, 1996). This is of particular importance when applied to hedge fund time series as often data sets are limited, thus reducing the system's degrees of freedom. The factor universe is then screened according to standard critical values of the F distribution with 1 and T–2 degrees of freedom, where T is the length of the time series return.

At this point, the quantitatively screened factor universe can be combined with qualitative factor selection, as determined by the risk manager's understanding of the strategy, to form the final set of factors for the fund under consideration. Typically there is considerable overlap between quantitative and qualitative factor subsets. However, prior to excluding drivers that are not a member of the joint set, consideration should be given to the possibility of a quantitatively significant factor having a legitimate presence due to style-drift or mis-classification, as well as the possibility that a qualitatively significant factor possesses a dormant significance. As discussed earlier, a final factor set of 5–8 factors, depending on strategy, provides a suitable trade-off between descriptivity and factor statistical significance.[1] If the selected factor set accurately represents the drivers of the fund and results in a correctly specified model, then the systematic component of fund returns has been accounted for. In such circumstances, it is then assumed that the estimated residuals will have independently normally distributed with mean zero and constant variance. The normality of the estimated residuals is evaluated using the Kolmogorov–Smirnov test for goodness of fit to a standardized normal distribution (Conover, 1980). Given the typically small size of the data sets employed in such an analysis, coupled with the transient nature in which the factors represent the trading strategies employed by hedge fund managers vary, confidence intervals for the Kolmogorov–Smirnov test cannot be set too tightly. Our empirical studies identify 90% as a suitable confidence interval given that the authors seek to identify expected relationships in addition to proven relationships between factor sets and manager return series. Having screened individual factors prior to regression, the rejection frequency of a well-constructed set of factors is minimal. When factor basis sets are rejected it is invariably due to the lack of a significant period of manager returns, in which case recourse to factors determined qualitatively can be made. It should also be noted that basis sets can become non-stationary through style drift in the underlying fund.

2.3 Calculation of Factor Loadings Through Multivariate Linear Regression

The vector of betas corresponding to the factor subset are determined through linear regression according to equation (1). The regression model is

$$r_t = \alpha_t + \sum_{d=1}^{D} \beta_{dt} r_{dt} + \varepsilon_t \tag{1}$$

where r_t denotes the return on the portfolio for the period ending at time t and r_{dt}, $d = 1, \ldots, D$ denotes the return on factor d for the same period. By specification, the coefficients α_t and β_{dt} are time varying. The model is estimated using an optimum-period rolling window, thus generating a matrix of estimates of the betas and a vector of estimates of alpha via a weighted least-squares regression.

As by specification the alpha and betas are time varying, the choice of rolling regression window size is critical for parameter estimation; if it is too long short-term reverting trends are washed out; if it is too short statistical significance is lost. By forming the coefficient of determination R_t^2 as a function of rolling window length t, the optimum period will manifest itself as a point of inflection, identifiable as a local minimum in the first derivative. Such a test is known as a Scree test, and was developed by Cattell (1966) to investigate eigenvalue problems.

In the present study, an absolute rolling regression period of 15–18 months, dependent on strategy, was identified as optimal, although this can extend out to over 30 months for longer-term investment strategies. Regression windows which cover the period since fund inception are also studied, but often prove to be of limited use given that hedge fund managers dynamically adjusted exposures on a frequent basis. Following Mina and Xiao (2001), the exponential smoothing decay factor is set to 0.94. Thus for 18 month regression window, the oldest observation will contribute with a weight of approximately 0.57 relative to the most recent.

2.4 Non-parametric Factor-correlated Monte Carlo Simulation

The estimated matrix of time-dependent factor exposures β_{dt}, and the estimated vector α_t are employed in conjunction with conventional multivariate Monte Carlo techniques to determine an associated time-dependent risk profile. Given the factor set is non-orthogonal, simulated return distributions for the factors are generated using a Cholesky factorization of their covariance matrix, Higham and Cheng (1998) to form a lower triangular matrix, which is then used as the basis for Monte Carlo simulation to generate uncorrelated, shocked factor return series that still possess the original covariance structure, while exploring the full scope of the factor distributions. Typically each factor is sampled of the order 10^5 times to assure convergence of the distribution.

As the initial factor set was standardized prior to regression, scale information is returned to the simulated system through beta-adjustment of the simulated returns and summed in conjunction with the alpha vector to form a sampled vector of aggregate portfolio returns. The histogram of this aggregate return vector is the implied risk profile. This technique generates simulated factor return vectors that replicate the historic factor covariance matrix, while allowing for the existence of portfolio returns which lie outside the scope of the time series of past returns. This factor-generated risk distribution is referred to as the implied risk profile, in the same sense that implied volatility is seen as a forward-looking indicator of realized volatility.

The avoidance of any explicit assumption of normality of the simulated fund returns mitigates a significant source of potential tail-risk estimation as identified by Sortino and Satchell (2001). It also has distinct advantages over single-series bootstrap methods, given that use of factor basis to describe a track record provides an extended-data set permitting out-of-sample simulation (Sharma, 1995).

Figure 1 illustrates the benefits of the present method by comparing typical cumulative distribution functions (cdfs) for the factor-driven approach (solid curve), the equivalent historic

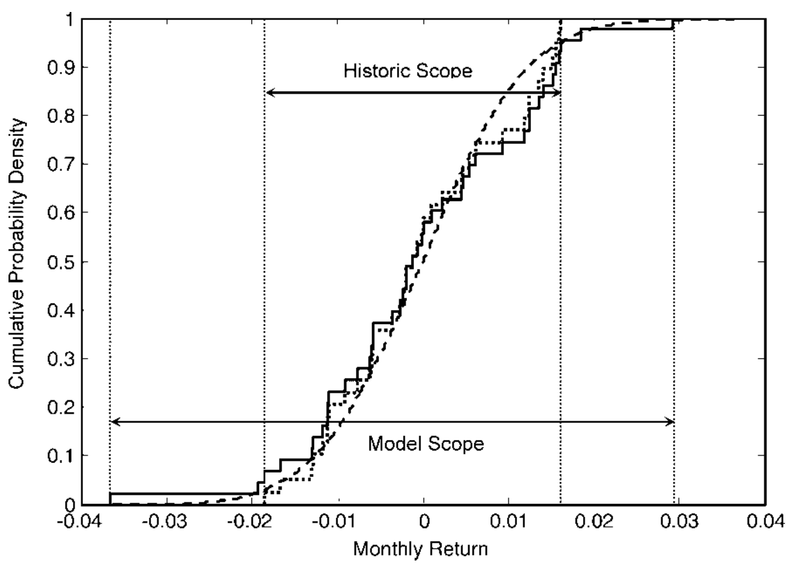

Figure 1. Typical monthly cumulative distribution functions constructed using: factor-driven non-parametric approach (solid); historic distribution (dotted); and a best-fit normal distribution (dashed), highlighting the impact of using historic data or placing an assumption of normality on the system

distribution (dotted curve), and a best-fit normal distribution (dashed curve). Figure 2 shows the corresponding histogram for the factor-driven approach. As shown in Figure 2, it is apparent that the implied risk distribution has a significant contribution from higher moments.

The presence of extra-historic, unrealized portfolio risk can be identified in greater detail, as illustrated in Figure 3, by overlaying a typical monthly implied risk distribution with the ex-post

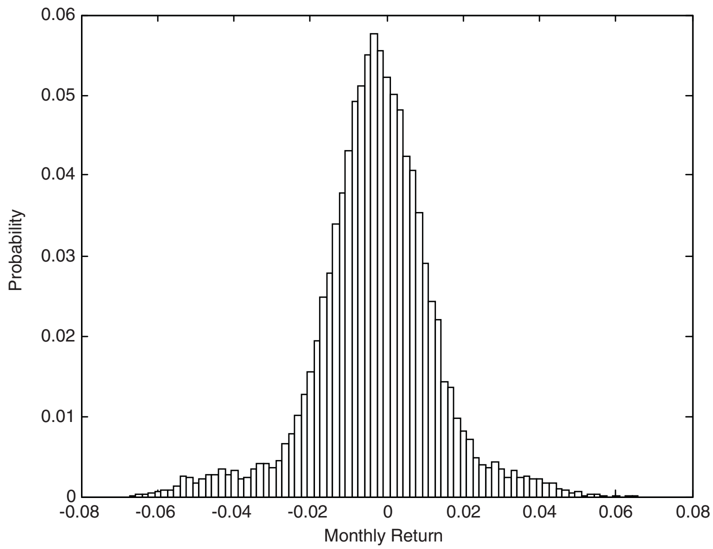

Figure 2. Typical implied risk profile generated by the factor-driven non-parametric model using a six- factor basis set and a Monte Carlo sample size of 5×10^5 events

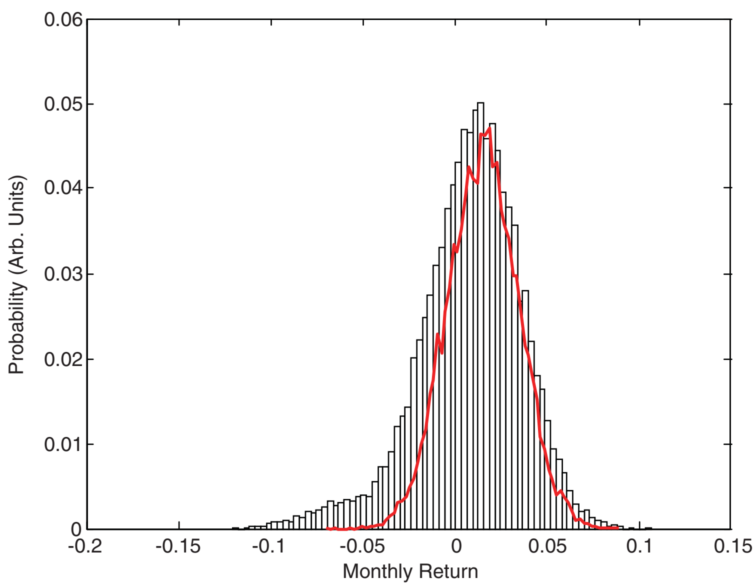

Figure 3. Implied monthly risk distribution for hedge fund H4 (bars) superimposed with the ex-post historic return distribution (line), highlighting the contribution of ex-ante risk through factor modelling

historic risk distribution. In the example shown in Figure 3, the left-hand or downside tail shows that there is a greater probability of large losses under the factor model than is the case using the historic return distribution.

3. Results and Discussion

In this section, we exemplify the method described above using two funds. The first is a convertible arbitrage hedge fund, referred to as H4. The second is a hedge fund-of-funds, referred to as F1.

Figure 4 depicts the actual returns for the hedge fund H4. Located either side of this returns series are the 97.5% non-parametric value-at-risk (VaR) bands along with our own expected tail loss (ETL) calculation at the 95% level, see for example Artzner *et al.* (1999). Conventionally, value-at-risk is calculated using a multiple of the standard deviation, σ, away from the distribution mean. For example a 97.5% VaR would equate to approximately $\langle r \rangle - 1.96\sigma$ under the assumption of a normal distribution. In the present study, it is this implicit assumption of the form of the distribution that we seek to avoid. Consequently, the VaR bands shown in Figure 4 have been calculated under a percentile counting regime, working in from a negative infinity limit until a cumulative probability of 2.5% is reached. Similarly, the expected tail loss is formed from the probability weighted integral of all returns up to a 5% cumulative probability.

Figure 4 clearly shows that the VaR bands for fund H4 are dynamic over the analysis period, with the risk profile tightening throughout 2002, and maintaining a conservative phase throughout 2003 before reverting outwards to historically aggressive levels in 2004 onwards. The VaR bands also reveal higher frequency structure, most notable from 2005 onwards, where shorter-term opportunistic investments result in localized widening of the implied risk profile. Use of historic

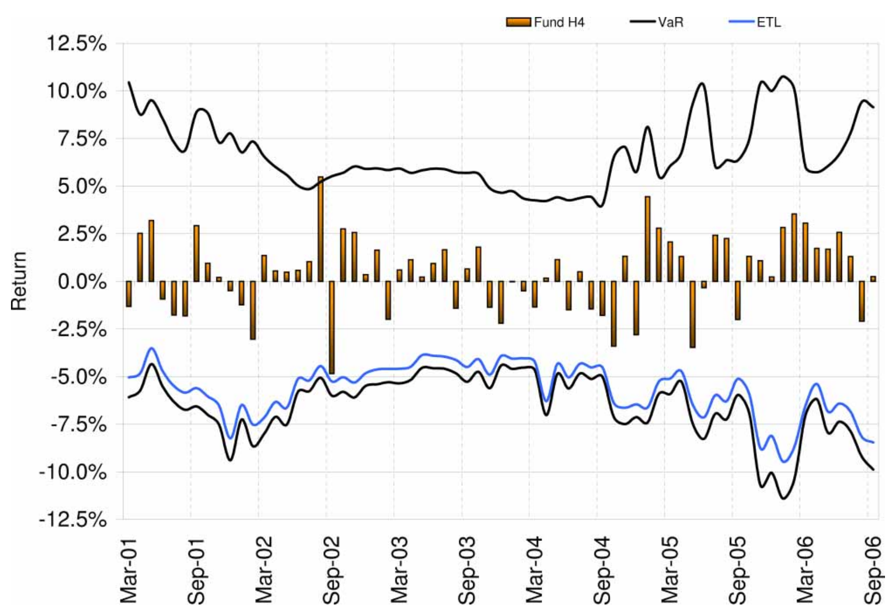

Figure 4. Realized return series for hedge fund H4 (bars), overlaid with 97.5% non-parametric value-at-risk boundaries (positive, and lower negative lines) and the expected tail loss at the >95% level (upper negative line)

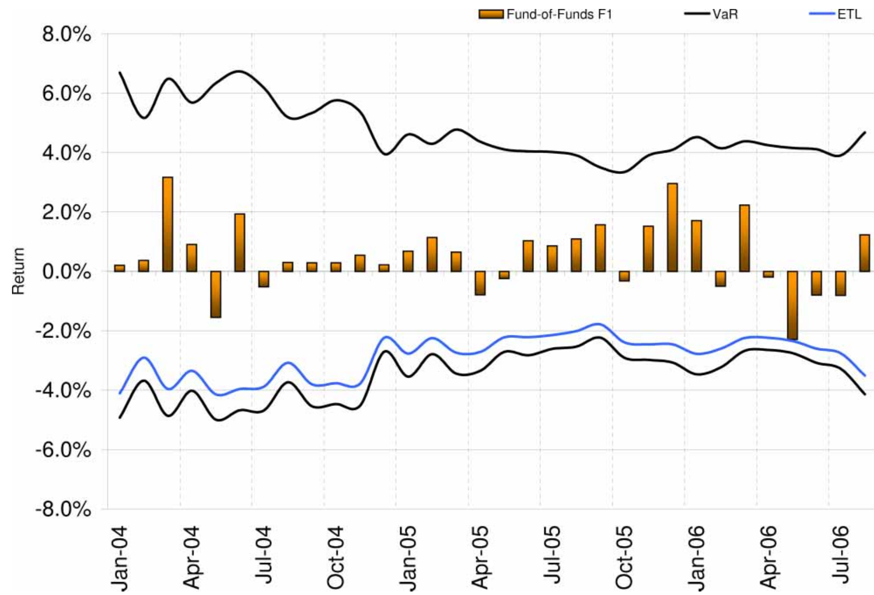

Figure 5. Realized return series for hedge fund-of-funds F1 (bars), overlaid with 97.5% non-parametric value-at-risk boundaries (positive, and lower negative lines) and the expected tail loss at the >95% level (upper negative line)

return analysis techniques in this instance would not reveal this structure. In addition to confirming such strategic trends or biases in the portfolio, it is also constructive to compare position level, ex-post, value-at-risk figures, thus providing a necessary, and useful out-of-sample accuracy check.

As well as providing a risk measurement framework for the underlying hedge funds, this analysis can also be performed at an aggregate hedge fund-of-funds level. This is illustrated for manager F1 in Figure 5. Such an analysis allows the composite risk distribution to be calculated frequently and for risk drift at the hedge fund-of-funds level to be identified and corrected over time. As a hedge fund-of-funds is represented by the weighted set of constituent funds, a further use of this factor-based framework is to calculate marginal contributions to portfolio risk at the constituent fund, strategy or factor level. Such a decomposition can be seen in Figure 6, where the marginal risk contribution to the fund has been attributed at both the constituent manager and strategy levels, and depicts how a nominal unit of risk would be allocated if deployed in the given portfolio. As a result of such information, quantitative portfolio construction by strategy, manager and factor can be facilitated, for example providing convergent solutions for optimal initial weightings to ensure maximum portfolio diversification, or to illustrate the potential effects of removing a fund from a given portfolio.

4. Summary

In summary, we have shown how combining conventional multivariate linear regression with non-parametric Monte Carlo simulation within a factor basis framework provides access to both time-dependent factor exposures and implied risk profiles, facilitating active style analysis in addition to a wealth of risk analysis beyond the conventional, and often misleading, standard deviation multiple based value-at-risk measures. Furthermore, this framework has proved suitable

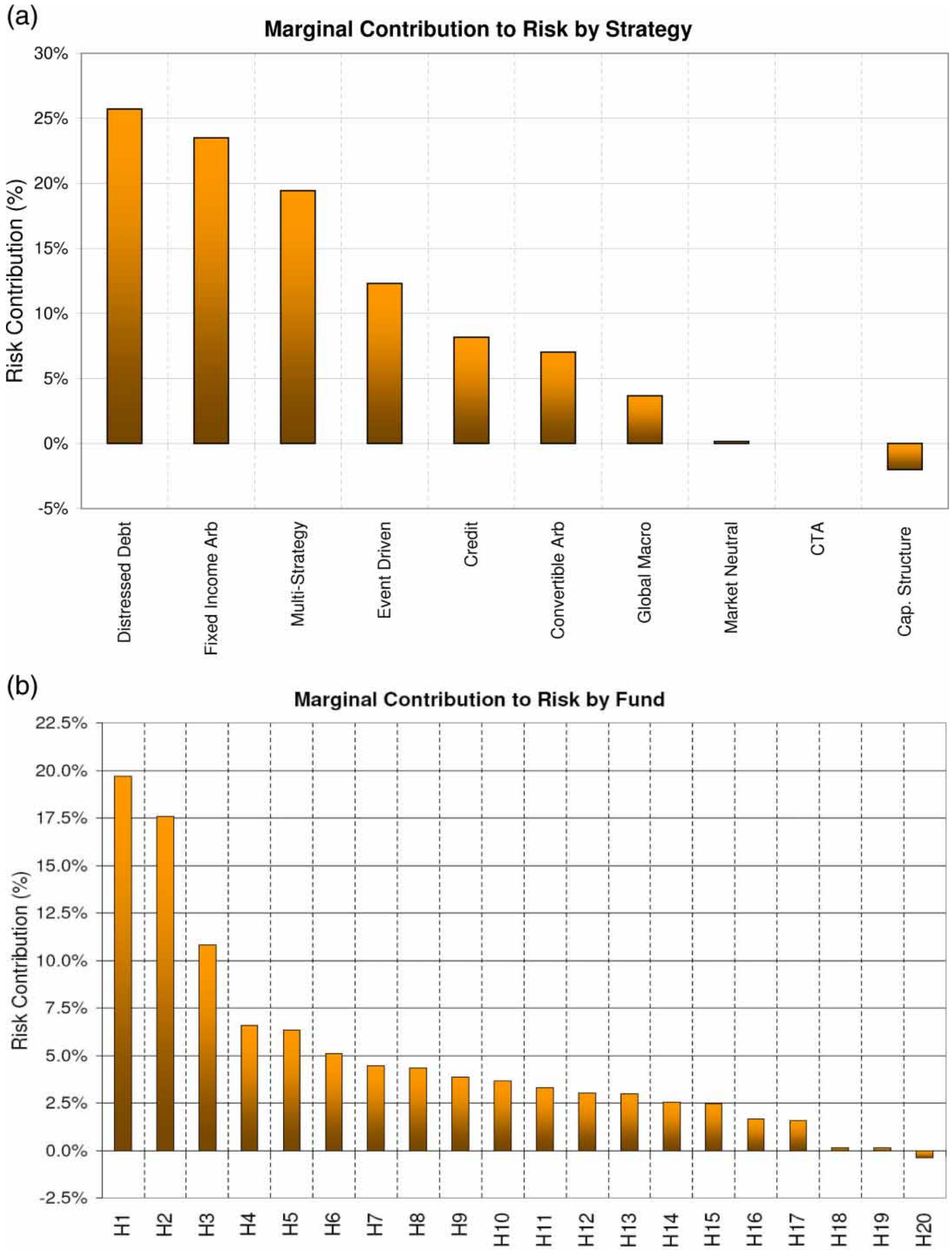

Figure 6. Hedge fund-of-funds marginal risk attribution at (a) the strategy level, and (b) the constituent fund level

for use across all strategies (as identified by the CSFB/Tremont classifications) with both the underlying hedge funds and collectively through hedge fund-of-funds without the need for full transparency. Further work is in progress and should yield prescriptive information on building diversified hedge fund portfolios and managing their risk.

Note

[1] If too many factors are selected, standard techniques like stepwise regression can be employed to determine a more parsimonious set.

References

Agarwal, V. and Narayan, Y. (2000a) On taking the alternative route: risks, rewards, and performance persistence of hedge funds, *Journal of Alternative Investments*, 2(4), pp. 6–23.

Agarwal, V. and Narayan, Y. (2000b) Generalized style analysis of hedge funds, *Journal of Asset Management*, 1(1), pp. 93–109.

Amenc, N., El Bied, S. and Martinelli, L. (2003) Predictability in hedge fund returns, *Financial Analysts Journal*, Sept/Oct, pp. 32–46.

Artzner, P., Delbaen, F., Eber, J. M. and Heath, D. (1999) Coherent measures of risk, *Mathematical Finance*, 9, pp. 203–228.

Asness, C. (2004a) An alternative future: Part I, *Journal of Portfolio Management* (30^{th} Anniversary Issue 2004), pp. 94–103.

Asness, C. (2004b) An alternative future: Part II, *Journal of Portfolio Management*, Fall, pp. 8–23.

Barlow, R. J. (1996) *Statistics* (New York: Wiley).

Brealey, R. A. and Kaplanis, E. (2001) Changes in the factor exposures of hedge funds. Working Paper HF-004, Hedge Fund Centre, London Business School.

Cattell, R. B. (1966) The scree test for the number of factors, *Multivariate Behavioral Research*, 1, pp. 245–276.

Conover, W. J. (1980) *Practical Nonparametric Statistics* (New York: Wiley).

Fung, W. K. H. and Hsieh, D. A. (1997) Empirical characteristics of dynamic trading strategies: the case of hedge funds, *Review of Financial Studies*, 10, pp. 275–302.

Fung, W. K. H. and Hsieh, D. A. (1999) Is mean-variance analysis applicable to hedge funds? *Economics Letters*, 62, pp. 53–58.

Fung, W. K. H. and Hsieh, D. A. (2002) Asset based style factors for hedge funds, *Financial Analysts Journal,* 57, pp. 16–33.

Fung, W. K. H. and Hsieh, D. A. (2004a) The risk in hedge fund strategies: alternative alphas and alternative betas. Working Paper, HF-013, Hedge Fund Centre, London Business School.

Fung, W. K. H. and Hsieh, D. A. (2004b) Hedge fund benchmarks: a risk based approach, *Financial Analyst Journal*, 60(5), pp. 65–80.

Higham, N. J. and Cheng, S. H. (1998) A modified Cholesky algorithm based on a symmetric indefinite factorization, *Journal of Matrix Analysis and Applications*, 19, pp. 1097–1110.

Ibbotson, R. G. and Chen, P. (2005) Sources of hedge fund returns: alphas, betas and costs. Yale ICF Working Paper 05-17.

Koepf, W. (1998) Hypergeometric summation: an algorithmic approach to summation and special function identities (Weisbaden: Vieweg Verlag), pp. 6–9.

Lo, A. (2001) Risk management for hedge funds: introduction and overview, *Financial Analysts Journal*, 57, pp. 16–33.

Mina, J. and Xiao, J. Y. (2001) *Return to RiskMetrics – The Evolution of a Standard* (New York: RiskMetrics Group).

Sharpe, W. F. (1992) Asset allocation, management style and performance measurement, *Journal of Portfolio Management*, 18, pp. 7–19.

Sharma, S. (1995) *Applied Multivariate Techniques* (New York: Wiley).

Sortino, F. and Satchell, S. (2001) *Managing Downside Risk in Financial Markets* (New York: Butterworth-Heinemann).