3rd IASTED Conference on Financial Engineering and Applications
October 2006, Cambridge, USA

DYNAMIC ANALYSIS OF HEDGE FUNDS

Michael Markov
Markov Processes International
25 Maple Street, Summit,
NJ 07901, USA
michael.markov@markovprocesses.com

Ilya Muchnik
Rutgers University
P.O. Box 8018, Piscataway,
NJ 08854, USA
muchnik@dimacs.rutgers.edu

Vadim Mottl, Olga Krasotkina
Tula State University
Lenin Ave. 92, Tula, 300600, Russia
vmottl@yandex.ru

ABSTRACT

In this paper, we review one of the most effective financial multi-factor models, called Returns Based Style Analysis (RBSA), from the standpoint of its performance in detecting dynamic factor exposures of hedge funds using only fund performance data. We analyze the shortcomings of earlier models in capturing changes in a dynamic portfolio structure and lay the groundwork for a new approach, which we call Dynamic Style Analysis (DSA). The problem is treated as that of estimating a time-varying regression model of the observed time series with the inevitable necessity to choose the appropriate level of model volatility, ranging from the full stationarity of instant models to their absolute independence of each other. We further propose an efficient method of model estimation and introduce a novel measure of the validity Predicted R^2 that is used to select the model parameters. Using both model and real hedge fund returns we illustrate the advantages of the proposed technique in analysis of hedge funds.

KEYWORDS

Style analysis of investment portfolios, hedge funds, time-varying regression, leave-one-out principle, Kalman filter.

1 Introduction

The Hedge fund industry has grown rapidly over the past decade to almost $1 trillion in assets and over 8,000 funds. At the same time, the amount of information on hedge funds available to investors is negligible as compared to traditional investment products such as mutual funds. In many cases, the only available information on a hedge fund is a time series of monthly returns and a vague description of the strategy. Returns are then analyzed using a variety of multi-factor models in order to detect the sensitivity of the hedge fund strategy to various risk *factors* (*factor exposures*) as well as to explain the fund's past performance.

One of the most effective and practical multi-factor models for analysis of investment portfolios, called Returns-Based Style Analysis (RBSA), was put forth by Sharpe [1,2]. In the RBSA model, the periodic return of a portfolio is approximated by the *constrained linear regression* of a relatively small number of single factors represented by the periodic returns of generic market indices, each of which represents a certain investment style or sector (market capitalization, quality, duration, region, etc.).

In order to account for allocation changes in active portfolios, Sharpe used a moving window of some preset length [2], assuming that the structure of the portfolio is constant inside the window.

Fung and Hsieh [3] applied RBSA to hedge funds where the method was reduced to unconstrained linear regression to account for shorting and leverage typical of hedge fund strategies. They note a significant loss of explanatory power applying RBSA to hedge funds as compared to traditional investment products such as mutual funds (0.25 and 0.75 median R^2 respectively). They conclude that such a low R^2 is due to the dynamic nature of hedge fund strategies and introduce generic indices designed to capture hedge fund dynamics, thus increasing median R^2 to 0.4 using the same static regression approach.

In [4], Fung and Hsieh further explore the issue of non-stationarity of RBSA and introduce a method to detect structural breakpoints in factor exposures to improve the R^2, but otherwise the allocations remain constant within the estimation window.

As a generalization of the static RBSA model, we propose a dynamic model in which a portfolio weights are considered as changing with time. The proposed approach, which we call Dynamic Style Analysis (DSA), consists of estimating a time-varying regression model of the observed time series of a portfolio's periodic returns and those of generic market indices. Time-varying regression has been subject of intensive study in statistical and econometric literature over the last fifteen years [5,6]. In this paper, we consider the problem of estimating a time-varying regression model of a portfolio in its inevitable connection with the necessity of choosing the appropriate level of volatility of results, ranging from the full stationarity of instant regression models to their absolute independence of each other. In selecting the volatility level, we use the "leave-one-out" principle widely adopted in Machine Learning [7,8].

One of the main results of this work is application of dynamic programming algorithm in implementing the "leave-one-out" procedure and providing, thereby, linear computational complexity of the algorithm relative to the length of the time series.

We illustrate the proposed approach to portfolio analysis by applying it to both model and real hedge fund strategies.

2 Returns Based Style Analysis

In RBSA model [1], the periodic return of a portfolio $r^{(p)}$ is approximated by the return on a portfolio of assets indices $r^{(i)}$ with weights $(\beta^{(1)},...,\beta^{(n)})$ equal to fractions invested in each asset at the beginning of the period under the assumption that the entire budget is fully spent on the investment:

$$r^{(p)} \cong \alpha + \sum_{i=1}^{n} \beta^{(i)} r^{(i)} . \qquad (1)$$

In [2], this model was used to analyze the performance of a group of US mutual funds and determined that a significant portion of a fund's return could be explained by a small number of assets. In order to estimate the parameters of the model (1), monthly returns of both the portfolio $\{r_t^{(p)}\}$ and asset indexes $\{r_t^{(i)}\}$ for consecutive months $t = 1,2,3,...$ were used to solve the following constrained quadratic optimization problem:

$$\begin{cases} (\hat{\alpha},\hat{\beta}^{(1)},...,\hat{\beta}^{(n)}) : \sum_{t=1}^{N} \left(r_t^{(p)} - \alpha - \sum_{i=1}^{n} \beta^{(i)} r_t^{(i)} \right)^2 \to \min, \\ \beta^{(i)} \geq 0, \ \sum_{i=1}^{n} \beta^{(i)} = 1. \end{cases} \qquad (2)$$

The resulting coefficients $(\hat{\beta}^{(1)},...,\hat{\beta}^{(n)})$ help to identify the major factors determining portfolio performance.

Further, recognizing that portfolio structure changes over time, Sharpe used a series of optimizations in moving windows of a smaller length K to determine the dynamics of portfolio factor exposures:

$$(\hat{\alpha}_t,\hat{\beta}_t^{(1)},...,\hat{\beta}_t^{(n)}) = \arg\min_{\alpha,\beta^{(i)}} \sum_{k=0}^{K-1} \left(r_{t-k}^{(p)} - \alpha - \sum_{i=1}^{n} \beta^{(i)} r_{t-k}^{(i)} \right)^2 . \qquad (3)$$

Model (2) has become commonly adopted in financial practice under the name of Returns Based Style Analysis (RBSA). The main appeal of this method for practitioners is that it is based solely on analysis of portfolio returns and does not require any other, very often proprietary, information about the portfolio composition.

The two major factors contributing to such wide acceptance of RBSA are its ease of interpretation and stability of results. It is worth noting that both of these factors are the direct result of the presence of non-negativity constraints in (2). These constraints, being the major innovation in RBSA, provide important prior information about the analyzed portfolio, i.e., the fact that most of investment portfolios such as mutual funds don't take short (negative) positions.

Since its introduction in 1992, RBSA model (1) has been criticized for its inability to capture an active portfolio's dynamics. Thus, because portfolio structure is assumed constant within the estimation window, the moving window technique (3) appears to be inadequate to capture rapid changes in portfolio structure.

In addition, model (1) loses much of its advantage when it is applied to the analysis of portfolios which are allowed to take short (negative) positions. In such cases, the non-negativity constraints $\beta^{(i)} \geq 0$ have to be dropped from (2), and the problem is reduced to a simple linear regression. In many such cases, due to multicolinearity, the moving window method (3) produces highly unstable, meaningless results.

The two limitations above often make RBSA inadequate for analysis of hedge funds because, unlike traditional investment vehicles such as mutual funds, hedge funds can be extremely dynamic and take significant short positions.

Usually attempts to overcome these shortcomings of RBSA consist of the introduction of additional indices into the static model (1) to capture the specifics of a generic hedge fund strategy [3]. None of the methods available to date represent a true dynamic model and, therefore, their explanatory power remains low.

3 Limitations of RBSA: Dynamic model of a hedge fund

We will illustrate the shortcomings of RBSA using a simple model of an equity long-short hedge fund. The long position of the sample fund is created using Russell 1000 Value and Growth indices with weights following a sine-wave pattern as shown in Figure 1.

The fund is invested 100% in the Russell 1000 Growth as of Jan-96 and then shifts assets into Russell 1000 Value with a relatively low 50% annual turnover. At any point in time the sum of both index weights is equal to 100%. We then create a long-short model portfolio by 100% hedging the long portfolio with the S&P 500 Index, i.e., effectively subtracting the index returns from the long portfolio return.

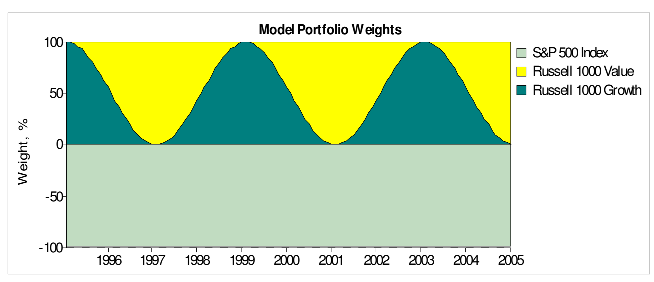

Figure 1. The long-short model portfolio.

Next, we apply a rolling 12-month window RBSA (3) on the composite monthly return time series of the model portfolio using, as regressors, the same three monthly indices that were used in its construction. The results are presented in Figure 2 where estimated allocations are stacked along the Y-axis.

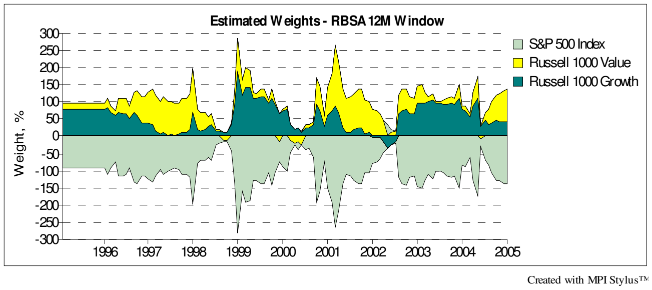

Figure 2. Estimated model portfolio, 12-month trailing window.

The results don't materially change when we vary the window size. In Figure 3 we show the result of rolling a 24-month window.

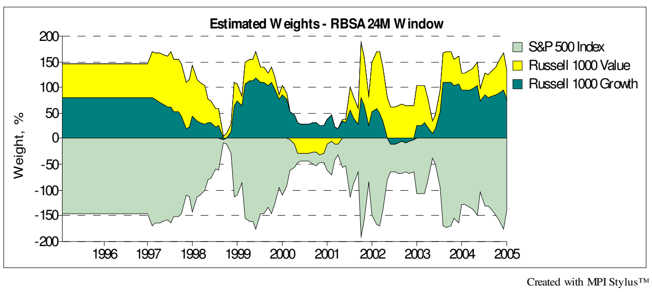

Figure 3. Estimated model portfolio, 24 month trailing window.

Although the turnover of the model portfolio is low and the number of assets very small, RBSA fails to adequately identify the model. Since no noise was added to the model portfolio returns, this provides clear indication that such poor model identification is the result of the window-based approach and multicolinearity as noted in Section 2, rather than noise in data as it is usually assumed. In Table 1 below we present the correlation matrix of assets used in construction of the model portfolio computed over the same 10-year period using monthly returns. The numbers in brackets represent the range of correlations computed over rolling 24-month windows.

The dynamic model introduced in this paper eliminates the shortcomings of traditional RBSA and makes it applicable to long-short strategies and hedge funds.

Table 1. The correlation matrix of assets constituting the model portfolio.

	R1000G	*R1000V*	*S&P500*
R1000G	1.00		
R1000V	0.71 (0.27;0.95)	1.00	
S&P500	0.94 (0.88;0.99)	0.90 (0.67;0.99)	1.00

4 Dynamic Style Analysis (DSA)

In contrast to static model (1), we propose a model in which factor exposures of the portfolio change in time. Let $t=1,2,...,N$ be a sequence of holding periods, for instance, days, weeks, months, quarters or years, and $B=(\boldsymbol{\beta}_t,\, t=1,...,N)$, $\boldsymbol{\beta}_t=(\beta_t^{(0)},\beta_t^{(1)},...,\beta_t^{(n)})$, $\sum_{i=0}^{n}\beta_t^{(i)}=1$, be the respective sequence of the portfolio's exposures at the beginning of each period. The notation $\beta_t^{(0)}$ is reserved in this case for a short-term instrument, such as bank deposit in an interest bearing account, often referred to as a risk-free asset.

For simplicity, we express the model in terms of excess returns on the portfolio $(r_t^{(p)}-r_t^{(0)})$ and assets $(r_t^{(i)}-r_t^{(0)})$ with respect to the return on the risk-free asset $r_t^{(0)}$. This equivalent notation effectively eliminates the need for the budget constraint $\beta_t^{(0)}+\sum_{i=1}^{n}\beta_t^{(i)}=1$ in (1). The new dynamic model of periodic portfolio returns can be written as follows:

$$y_t=(r_t^{(p)}-r_t^{(0)})=\sum_{i=1}^{n}\beta_t^{(i)}(r_t^{(i)}-r_t^{(0)})=\sum_{i=1}^{n}\beta_t^{(i)}x_t^{(i)}+e_t=\boldsymbol{\beta}_t^T\mathbf{x}_t+e_t \,. \tag{4}$$

Here $y_t=(r_t^{(p)}-r_t^{(0)})$ are known excess returns of the portfolio for each period t, and $\mathbf{x}_t=\left[(r_t^{(i)}-r_t^{(0)}),i=1,...,n\right]\in\mathbb{R}^n$ are known vectors of observed excess returns of assets for these periods, whereas $\boldsymbol{\beta}_t=(\beta_t^{(1)},...,\beta_t^{(n)})\in\mathbb{R}^n$ are vectors of time-varying fractional asset weights to be estimated.

The key element of the proposed *Dynamic Style Analysis* (DSA) is the treatment of fractional asset weights as a hidden process assumed *a priori* to possess the Markov property:

$$\boldsymbol{\beta}_t=\mathbf{V}_t\boldsymbol{\beta}_{t-1}+\boldsymbol{\xi}_t\,, \tag{5}$$

where matrices $\mathbf{V}_t$ determine the assumed hidden dynamics of the portfolio structure, and $\boldsymbol{\xi}_t$ is the vector white noise, non-stationary in the general case.

Equation (5) determines the state-space model of a dynamic system, while (4) plays the role of its observation model. In these terms, the DSA problem can be described as estimating the time-varying state of the respective dynamic system $B(Y,X)=\left(\boldsymbol{\beta}_t(Y,X),t=1,...,N\right)$ from observations $(Y,X)=\left[(y_t,\mathbf{x}_t)=\left((r_t^{(p)}-r_t^{(0)}),(r_t^{(i)}-r_t^{(0)}),i=1,...,n\right),t=1,...,N\right]$.

For estimating time-varying models of this kind (4)-(5), we use the *Flexible Least Squares* approach (FLS) first introduced in [5]. As applied to the DSA problem, the FLS criterion has the form

$$\begin{cases}\hat{B}(Y,X,\lambda)=\arg\min J(\boldsymbol{\beta}_t,t=1,...,N\,|\,Y,X),\\ J(\boldsymbol{\beta}_t,t=1,...,N\,|\,Y,X)=\\ \sum_{t=1}^{N}(y_t-\boldsymbol{\beta}_t^T\mathbf{x}_t)^2+\lambda\sum_{t=2}^{N}(\boldsymbol{\beta}_t-\mathbf{V}_t\boldsymbol{\beta}_{t-1})^T\mathbf{U}_t(\boldsymbol{\beta}_t-\mathbf{V}_t\boldsymbol{\beta}_{t-1}).\end{cases} \tag{6}$$

The assumed covariance matrices $\mathbf{Q}_t$ of white noise $\boldsymbol{\xi}_t$ in (5) occur here in the inversed form $\mathbf{U}_t=\mathbf{Q}_t^{-1}$. We shall additionally assume matrices $\mathbf{V}_t$ to be non-degenerate, in this case, they also determine the reversed dynamics of the time-varying regression coefficients.

The positive parameter λ in (6) is responsible for the noise ratio in (4)-(5), i.e. for the level of smoothness of regression coefficients. Thus, the smoothness parameter λ

balances the two conflicting requirements: to provide a close approximation of portfolio returns and, at the same time, to control the smoothness of asset weights $\boldsymbol{\beta}_t$ over time.

Note that proposed DSA approach to analysis of a portfolio returns makes it possible to identify structural shifts (breakpoints) in asset weights. Such shifts can be attributed to rapid changes in the portfolio positions, use of derivative instruments, etc. and in most cases are hidden from public. The DSA approach to breakpoint identification is based on local relaxation of smoothness requirement in (6) and is discussed at length in [9].

5 The cross validation principle of estimating the smoothness parameter

The FLS criterion (6) depends on a number of parameters. Matrices $\mathbf{V}_t$ and $\mathbf{U}_t$ of the transition model (5) can be defined a priori, depending on the model of hidden dynamics of time-varying regression coefficients and the "style" of smoothing of their estimates. For example, matrix $\mathbf{V}_t$ can be defined as the unity matrix thus requiring simple smoothness of estimates. Alternatively, we could allow for non-smoothness of asset weights by incorporating market-driven changes of weights into a transition model (5) as follows[1]:

$$\beta_t^{(i)} = \frac{1+r_{t-1}^{(i)}}{\sum_{k=0}^{n}\beta_{t-1}^{(k)}(1+r_{t-1}^{(k)})}\beta_{t-1}^{(i)} + \xi_t^{(i)}, \quad i=0,\ldots,n, \quad t=2,\ldots,N.$$

As to the coefficient λ, it is extremely problematic to preset its value a priori.

If the volatility parameter is given a certain value λ, the FLS estimate of time-varying regression coefficients (6) will be a function of λ, $\hat{B}(Y,X,\lambda) = \left(\hat{\boldsymbol{\beta}}_t(Y,X,\lambda), t=1,\ldots,N\right)$. It is impossible to find an "appropriate" value of λ by attempting to additionally minimize the residual sum of squares in (6) $\sum_{t=1}^{N}\left(y_t - \hat{\boldsymbol{\beta}}_{t-1}^T(Y,X,\lambda)\mathbf{x}_t\right)^2 \to \min_\lambda$.

Indeed, as $\lambda \to \infty$, the second sum in (6) totally prevails over the first sum, the values of the hidden process become functionally related to each other $\hat{\boldsymbol{\beta}}_t(Y,X,\lambda) = \mathbf{V}_t\hat{\boldsymbol{\beta}}_{t-1}(Y,X,\lambda)$, and the model is reduced to a static regression. Alternately, as $\lambda \to 0$, the instantaneous values become a priori independent, each estimate $\hat{\boldsymbol{\beta}}_{t-1}(Y,X,\lambda)$ is determined practically by only one current element of the time series $(y_t, \mathbf{x}_t)$, and the model will be "extremely" time-varying.

It is now easy to see that one can achieve 100% fit of y_t in (6) using arbitrary explanatory variables by adjusting the smoothness parameter. These variables can be totally unrelated to the analyzed portfolio return. Moreover, even when explanatory variables are selected properly, changing the smoothness parameter can lead to very different results. It is therefore crucial that this parameter is estimated from data, because in most cases analysts don't have enough information about underlying hedge fund positions and their dynamics.

Actually, the sought-for sequence of time-varying regression coefficients $B(Y,X,\lambda) = \left(\boldsymbol{\beta}_t(Y,X,\lambda), t=1,\ldots,N\right)$ is a model of the observed time series $(Y,X) = \left[(y_t,\mathbf{x}_t) = \left((r_t^{(p)} - r_t^{(0)}), (r_t^{(i)} - r_t^{(0)}), i=1,\ldots,n\right), t=1,\ldots,N\right]$, and the choice of λ is the choice of a class of models which would be most adequate to the data [7,8]. A commonly used measure of regression model fit is its coefficient of determination R^2. In [2], the R^2 was defined as the proportion of the portfolio volatility explained by systematic exposures using the moving window technique (3). In terms of the FLS criterion (6), the coefficient of determination is expressed by the ratio

$$R^2 = \frac{\sum_{t=1}^{N}(y_t)^2 - \sum_{t=1}^{N}\left(y_t - \hat{\boldsymbol{\beta}}_t^T(\lambda)\mathbf{x}_t\right)^2}{\sum_{t=1}^{N}(y_t)^2} = 1 - \frac{\sum_{t=1}^{N}\left(y_t - \hat{\boldsymbol{\beta}}_t^T(\lambda)\mathbf{x}_t\right)^2}{\sum_{t=1}^{N}(y_t)^2}. \quad (7)$$

By decreasing λ, it is easy to drive R^2 up to 100% but at the same time, obtain highly volatile, meaningless estimates of fractional asset weights $\hat{\boldsymbol{\beta}}_t(Y,X,\lambda)$.

The major reason for this shortfall of the R^2 statistic is that it uses the same data set for both estimation and verification of the model. The *Cross Validation* method suggested by Allen [10] under the name of *Prediction Error Sum of Squares* (PRESS) is aimed at overcoming this obstacle. According to this method, an observation is removed from the sample, the model is evaluated on the remaining observations, and the prediction error is calculated on the removed observation. This procedure is then repeated for each observation in the sample, and the sum of squared errors is computed. The Cross Validation (CV) principle is widely used in data analysis [11,12], including pattern recognition, where the procedure is known under the name of "leave-one-out" [8,13].

The essence of the Cross Validation principle can be explained as the assessment of the adequacy of the given model by estimating the variance of the residual noise $D(e)$ in (4) and comparing it with the full variance of the goal variable $D(y) = (1/N)\sum_{t=1}^{N}(y_t)^2$. When computing the error at a time t, it is incorrect to use the estimate $\hat{\boldsymbol{\beta}}_t$ obtained by minimizing the criterion (6) including the observation at that time $(y_t, \mathbf{x}_t)$. The CV principle leads to the following procedure that provides a correct estimate of the observation noise variance.

In the full time series $\left((y_1,\mathbf{x}_1),\ldots,(y_N,\mathbf{x}_N)\right)$, single elements $t=1,\ldots,N$ are skipped one by one $\left((y_1,\mathbf{x}_1),\ldots,(y_{t-1},\mathbf{x}_{t-1}),\ (y_{t+1},\mathbf{x}_{t+1}),\ldots,(y_N,\mathbf{x}_N)\right)$, each time replacing the sum $\sum_{t=1}^{N}\left[y_t - \left(\boldsymbol{\beta}_t(\lambda)\right)^T\mathbf{x}_t\right]^2$ in (6) with the truncated sum

[1] This clearly leads to nonlinearity in criterion (6) which can be addressed, for example, by using iterations of linear models [9].

$\sum_{s=1,s\neq t}^{N}\left[y_s-(\boldsymbol{\beta}_s(\lambda))^T\mathbf{x}_s\right]^2$. The optimal vector sequences $(\hat{\boldsymbol{\beta}}_1^{(t)},...,\hat{\boldsymbol{\beta}}_N^{(t)})$ are found where the upper index (t) means that the observation $(y_t,\mathbf{x}_t)$ was omitted when computing the respective estimate. For each t, the instantaneous squared prediction error is calculated using the respective single estimate $\left[y_t-(\boldsymbol{\beta}_t^{(t)}(\lambda))^T\mathbf{x}_t\right]^2$. The cross-validation estimate of the noise variance is found as the average over all the local squared prediction errors

$$\hat{D}_{CV}(e\,|\,\lambda)=\frac{1}{N}\sum_{t=1}^{N}\left[y_t-\left(\hat{\boldsymbol{\beta}}_t^{(t)}(\lambda)\right)^T\mathbf{x}_t\right]^2 \quad (8)$$

The smaller $\hat{D}_{CV}(e\,|\,\lambda)$, the more adequate the model with the given value of the smoothness parameter λ for the observed time series $\left((y_1,\mathbf{x}_1),...,(y_N,\mathbf{x}_N)\right)$.

The cross-validation estimate of the residual noise variance $\hat{D}_{CV}(e\,|\,\lambda)$ can be further scaled to make it comparable across different analyzed portfolios. We suggest the cross-validation statistic

$$PR^2(\lambda)=\frac{D(y)-\hat{D}_{CV}(e\,|\,\lambda)}{D(y)}=1-\frac{\hat{D}_{CV}(e\,|\,\lambda)}{D(y)}. \quad (9)$$

which we call *Predicted R-squared*. Note that it is computed similarly to the regression R-squared statistic (7).

We suggest a method of determining optimal model parameters that consists of processing the given time series $\left((y_1,\mathbf{x}_1),...,(y_N,\mathbf{x}_N)\right)$ several times with different tentative values of λ. Each time, the model adequacy is assessed by the averaged squared prediction error (8) estimated by the cross validation procedure. The value λ^* that yields the maximum value of the cross-validation statistic (9) is to be taken as the smoothing parameter recommended for the given time series:

$$\lambda^*=\arg\max_{\lambda} PR^2(\lambda). \quad (10)$$

It should be noted that the selection of model parameters through minimizing the prediction error makes this method a version of the James-Stein estimator [11].

6 Kalman filter and smoother for minimization of flexible least squares and cross validation

The FLS criterion (6) is a quadratic function, and its minimization leads to a system of linear equations. At the same time, it belongs to the class of pair-wise separable optimization problems [14]. In these, the objective function is the sum of functions each dependant on not more than two vector variables, in this case $\boldsymbol{\beta}_{t-1}$ and $\boldsymbol{\beta}_t$ associated with immediately successive time moments. As a result, the matrix of the system of linear equation in respect to variables $\boldsymbol{\beta}_1,...,\boldsymbol{\beta}_N$ has a block-threediagonal structure, which is efficiently solved by the double-sweep method, a quadratic version of the much more general dynamic programming method [14]. These algorithms are, in turn, equivalent to the Kalman filter and smoother [15].

First, the Kalman filter runs along the time series (signal)

$$\boldsymbol{\beta}_{1|1}=\left(y_1/\mathbf{x}_1^T\mathbf{x}_1\right)\mathbf{x}_1,\ \mathbf{Q}_{1|1}=\mathbf{x}_1\mathbf{x}_1^T \text{ at } t=1,$$

$$\boldsymbol{\beta}_{t|t}=\mathbf{V}_t\boldsymbol{\beta}_{t-1|t-1}+\mathbf{Q}_{t|t}^{-1}\mathbf{x}_t(y_t-\mathbf{x}_t^T\mathbf{V}_t\boldsymbol{\beta}_{t-1|t-1}),\ t=2,...,N, \quad (11)$$

$$\begin{aligned}\mathbf{Q}_{t|t}=&\mathbf{x}_t\mathbf{x}_t^T+\mathbf{U}_t\mathbf{V}_t\left(\mathbf{V}_t^T\mathbf{U}_t\mathbf{V}_t+(1/\lambda)\mathbf{Q}_{t-1|t-1}\right)^{-1}\mathbf{Q}_{t-1|t-1}\mathbf{V}_t^{-1}=\\&\mathbf{x}_t\mathbf{x}_t^T+\left(\mathbf{V}_t\mathbf{Q}_{t-1|t-1}\mathbf{V}_t^{-1}+(1/\lambda)\mathbf{U}_t^{-1}\right).\end{aligned} \quad (12)$$

The intermediate vectors $\boldsymbol{\beta}_{t|t}$ and matrices $\mathbf{Q}_{t|t}$ are parameters of the Bellman functions $J_{t|t}(\boldsymbol{\beta}_t)$ which are quadratic in this case [15]:

$$J_{t|t}(\boldsymbol{\beta}_t)=\min_{\boldsymbol{\beta}_1,...,\boldsymbol{\beta}_{t-1}}J_t(\boldsymbol{\beta}_1,...,\boldsymbol{\beta}_t)=(\boldsymbol{\beta}_t-\boldsymbol{\beta}_{t|t})^T\mathbf{Q}_{t|t}(\boldsymbol{\beta}_t-\boldsymbol{\beta}_{t|t})+const,$$

$$J_t(\boldsymbol{\beta}_1,...,\boldsymbol{\beta}_t)=\sum_{s=1}^{t}(y_s-\boldsymbol{\beta}_s^T\mathbf{x}_s)^2+\lambda\sum_{s=2}^{t}(\boldsymbol{\beta}_s-\mathbf{V}_s\boldsymbol{\beta}_{s-1})^T\mathbf{U}_s(\boldsymbol{\beta}_s-\mathbf{V}_s\boldsymbol{\beta}_{s-1}).$$

Here $J_t(\boldsymbol{\beta}_1,...,\boldsymbol{\beta}_t)$ are partial criteria of the same structure as (6). The minimum points $\boldsymbol{\beta}_{t|t}$ of the Bellman functions yield the filtration estimates of the unknown regression coefficients at current t under the assumption that the time series is only observed up to point t.

Then, the Kalman smoother runs backwards $t=N-1,...,1$:

$$\hat{\boldsymbol{\beta}}_t=\boldsymbol{\beta}_{t|t}+\mathbf{H}_t(\hat{\boldsymbol{\beta}}_{t+1}-\boldsymbol{\beta}_{t|t}), \quad (13)$$

$$\begin{aligned}\mathbf{H}_t=&\left(\mathbf{V}_{t+1}^T(\lambda\mathbf{U}_{t+1})\mathbf{V}_{t+1}+\mathbf{Q}_{t|t}\right)^{-1}\mathbf{V}_{t+1}^T(\lambda\mathbf{U}_{t+1})\mathbf{V}_{t+1}=\\&\left(\mathbf{I}+(1/\lambda)\mathbf{V}_{t+1}^{-1}\mathbf{U}_{t+1}^{-1}(\mathbf{V}_{t+1}^{-1})^T\mathbf{Q}_{t|t}\right)^{-1}.\end{aligned} \quad (14)$$

The resulting sequence is just the minimum point of the FLS criterion (6) $\hat{B}(Y,X,\lambda)=\left(\hat{\boldsymbol{\beta}}_t(Y,X,\lambda),t=1,...,N\right)$.

To compute the "leave-one-out" estimate of the noise variance (8), we have to find the estimate of each regression coefficient vector $\hat{\boldsymbol{\beta}}_t(Y^{(t)},X^{(t)},\lambda)$ from the time series $(Y^{(t)},X^{(t)})=\left((y_s,\mathbf{x}_s),s=1,...,t-1,t+1,...,N\right)$ where the element $(y_t,\mathbf{x}_t)$ is cut out. This means that, when running the Kalman filter, we have to use the matrix $\mathbf{Q}_{t|t}^{(t)}=\mathbf{Q}_{t|t}-\mathbf{x}_t\mathbf{x}_t^T=\left(\mathbf{V}_t\mathbf{Q}_{t-1|t-1}\mathbf{V}_t^{-1}+(1/\lambda)\mathbf{U}_t^{-1}\right)^{-1}$ at step t instead of $\mathbf{Q}_{t|t}$ (12).

7 The Computational Complexity of DSA

Straightforward application of the cross validation principle (8)-(10) in determining the value of the smoothness parameter implies running the Kalman filtration-smoothing procedure (11)-(14) N times for each removed observation corresponding to time period t, which prevents maintaining the linear computational complexity of the algorithm. To analyze N=120 monthly returns of a portfolio using n=10 economic sectors as variables, the quadratic problem (6) with Nn=1,200 variables can easily be done using the standard Kalman filter-smoother with linear computational complexity with respect to N. But in order to compute the cross-validation statistic (9) corresponding to a single value of the smoothness parameter, N=120 such optimizations are required, and computing the CV statistic on a grid of 20 values of this parameter requires solving $20N = 2{,}400$ problems (6), i.e. 2,400 runs of the optimization procedure.

To avoid repeated processing of the signal for each tentative value of λ, a technique incorporating the computation of the "leave-one-out" error (8) into the main dynamic programming procedure is proposed in [15]. It is shown that the estimate $\hat{\boldsymbol{\beta}}_t^{(t)}(\lambda)$ is determined by the expression

$$\hat{\boldsymbol{\beta}}_t^{(t)}(\lambda) = \hat{\boldsymbol{\beta}}_t(Y^{(t)}, X^{(t)}, \lambda) = (\mathbf{Q}_{t|N} - \mathbf{Q}_t^0)^{-1}(\mathbf{Q}_{t|N}\boldsymbol{\beta}_{t|N} - \mathbf{Q}_t^0\boldsymbol{\beta}_t^0),$$

where matrices $\mathbf{Q}_{t|N}$ are also computed on the backward run of the Kalman smoother for $t = N-1, \ldots, 1$

$$\mathbf{Q}_{t|N} = \left(\mathbf{H}_t \mathbf{V}_{t+1}^{-1}\mathbf{Q}_{t+1|N}^{-1}(\mathbf{V}_{t+1}^{-1})^T \mathbf{H}_t^T + (\mathbf{Q}_{t|t} + \lambda \mathbf{V}_{t+1}^T \mathbf{U}_{t+1}\mathbf{V}_{t+1})^{-1}\right)^{-1},$$

starting with matrix $\mathbf{Q}_{N|N}$ found at the last step of the Kalman filter (12).

8 Testing the DSA Approach: Dynamic model of a hedge fund

We applied the DSA approach (6) to the model portfolio developed in Section 3 with the smoothness parameter λ selected in accordance with (10). The result of this analysis is shown in Figure 4.

In order to test the sensitivity of the model to noise in the data, we added idiosyncratic white noise to the portfolio's monthly returns in the amount of 20% of the portfolio volatility[1]. The resulting portfolio returns were analyzed, and the output corresponding to the maximum value of the CV statistic is shown in Figure 5. This result corresponds to the optimal smoothness parameter $\lambda = 0.2$ selected to provide the maximum value of the PR^2 statistic (10).

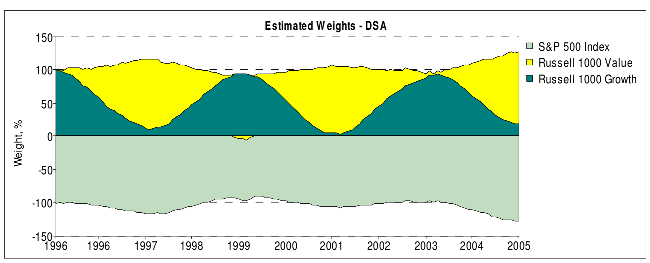

Figure 4. DSA approach – estimation of the model portfolio.

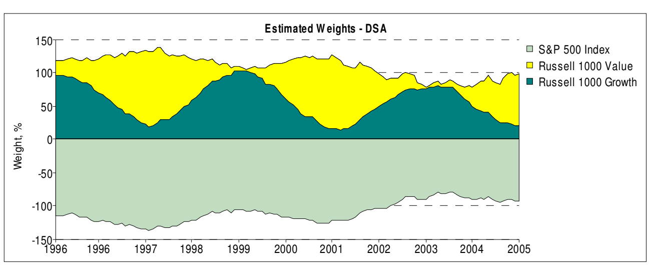

Figure 5. DSA approach – noisy model portfolio.

[1] Since the standard deviation of monthly returns over 120 months makes $\sigma = 1.1\%$, we applied white noise $N(0, 0.22)$.

9 Case Studies

In this section, we present examples of the application of the dynamic multi-factor methodology developed in this paper to real-life hedge funds.

9.1 Laudus Rosenberg Value Long/Short Fund

According to the fund prospectus, the Laudus Rosenberg Value Long/Short mutual fund[2] used computer models to buy underpriced US stocks and sell short other stocks in order to maintain both market and sector neutrality. Such neutrality is very important for investors because it protects their investment in market downturns.

Fund monthly returns are shown in Figure 6. We will compare performance of the traditional RBSA and Dynamic Style Analysis (DSA) in determining sensitivity of the fund returns to economic sectors.

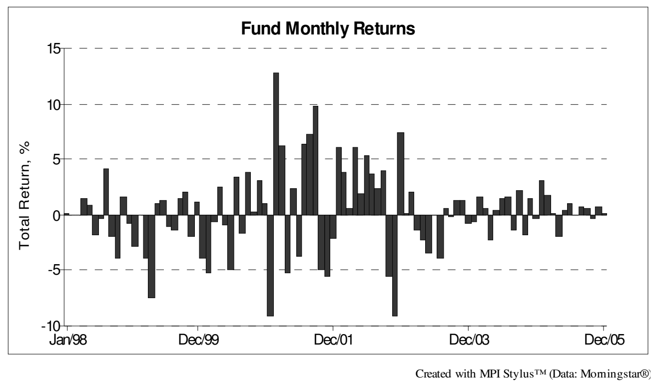

Figure 6. Fund returns.

For our analysis we used 10 indexes provided by Dow Jones Indexes[3]. The result of the analysis using a 36-month window RBSA (3) is presented in Figure 7.

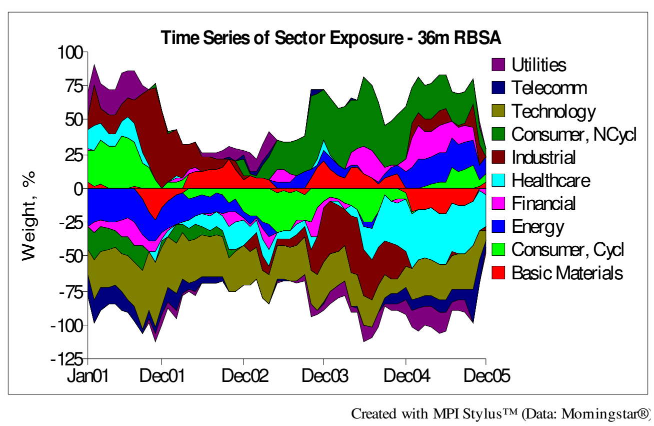

Figure 7. Results using RBSA in 36 month window.

[2] Laudus Rosenberg Value Long/Short (Ticker: BRMIX) is a mutual fund employing a strategy similar to a long-short hedge fund. Information on this fund is available on finance.yahoo.com and www.morningstar.com

[3] Source: indexes.dowjones.com

The result is very volatile and unrealistic with R^2=0.60 (7). Shortening the estimation window produces higher R^2 value but much more volatile results.

We then applied the DSA approach using the same return data and selected the optimal parameter λ in accordance with (10). The resulting sector exposures are presented in Figure 8.

In Figure 9 we present the values of criterion $PR^2(\lambda)$ (10) corresponding to various λ plotted along the logarithmic axis and the optimal value of the smoothness parameter λ^* at the intersection of the two dashed lines. Even though the weights are much less volatile in Figure 8 than those in Figure 7, the resulting $R^2 = 0.86$ (7) is much greater than 0.6, which was determined by the RBSA trailing window. The corresponding optimal value of the *Predicted* R^2 (10) is $PR^2(\lambda^*) = 0.53$.

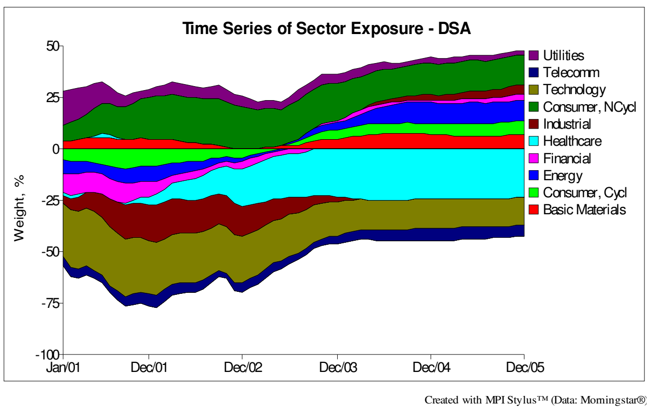

Figure 8. DSA-estimated asset weights.

Therefore, the proposed technique allowed us to achieve a much closer approximation of the fund return pattern with a significantly more realistic pattern of sector exposures.

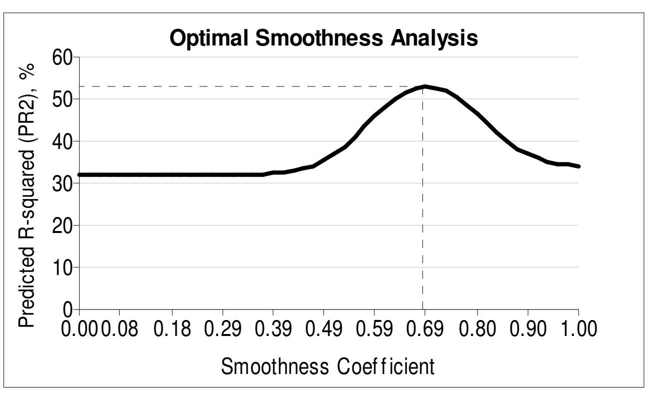

Figure 9. Selection of smoothness parameter.

9.2 Long Term Capital Management (LTCM)

The collapse of this highly-leveraged fund in 1998 is by far the most dramatic hedge fund story to date. At the beginning of 1998, the $5B fund maintained a leverage ratio of about 28:1 [16,17]. LTCM's troubles began in May-June 1998. By the end of August, the fund had lost 52%. By the end of September, 1998, a month after the Russian crisis, the fund had lost 92% of its December 1997 assets. Fearing the destabilizing impact of the highly leveraged fund on global financial markets, on September 23rd, the Federal Reserve Bank of New York orchestrated a bailout of the fund by a group of 14 major banks.

The investment strategy of the fund was based on spread bets between thousands of closely related securities in various global markets. Such bets are based on the assumption that the securities' prices will eventually converge and the arbitrage position will result in a profit. In fact, spreads continued to increase, which eventually led to the collapse of the fund. It took several major investigations, including one commissioned by President Clinton [17], to determine that the major losses sustained by LTCM came from bets on global credit spreads.

We use the DSA methodology to determine the major factors explaining LTCM's losses in 1998 using the fund's monthly returns. We also determine the leverage ratio, an important risk factor which, according to published figures [16,17], increased from 28:1 to 52:1 in 1998. The fund's 1998 monthly returns in January-August shown in Figure 10 were obtained from public sources [18]. Daily or weekly returns, if available, would provide far greater accuracy.

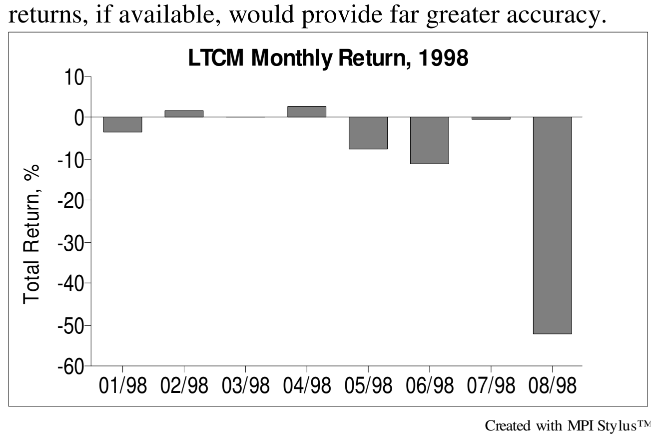

Figure 10. LTCM monthly returns in 1998.

For our analysis we used the following indexes provided by Lehman Brothers and Merrill Lynch: US Corporate and Government Long Bond Indices, European (EMU) Corporate and Government Bond Indices and the US Mortgage-Backed Securities Index. The result of the analysis is presented in the Figure 11. Asset exposures β_t of the fund for each time period are "stacked" along the Y-axis, with the sum equal to 100%. The negative weights shown below the X-axis correspond to borrowed assets and represent leverage. Evidently, the leverage comes from credits spreads – both US (Corp Index vs. Govt Index) and EMU (Corp Index vs. Govt Index). There is also significant exposure to Mort gages. Our results show that the leverage increased from 35:1 (or 3,500%) to 45:1 (4,500%) during 1998, which is close to the figures published in [16,17].

The result corresponds to the optimal smoothness coefficient λ which was selected to provide the maximum value of the Predicted R^2 statistic. The R^2 of this result is 0.99, while Predicted R^2 is 0.98.

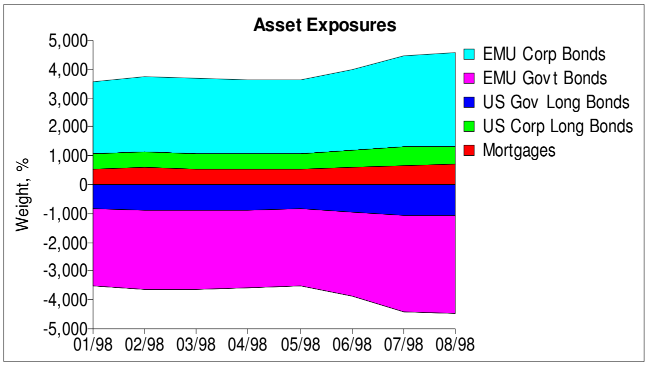

Figure 11. LTCM Estimated Asset Weights $\boldsymbol{\beta}_t$.

The "growth of $100" chart in Figure 12 showing the performance of LTCM in 1998 presents an excellent model fit, where the performance of the fund is very closely approximated by the model. In chart, the "Total" line represents the fund and the "Style" represents the model.

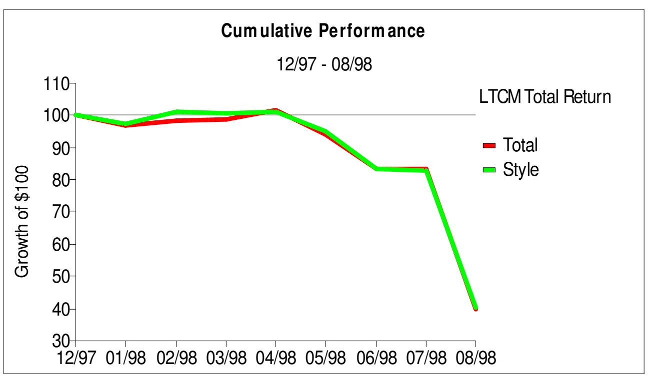

Figure 12. LTCM performance tracking.

Following Allen's application of leave-one-out PRESS statistic in model selection [10], in Table 2 we use cross-validation static Predicted R^2 (10) to illustrate "optimality" of the result obtained in our analysis of the LTCM. In the table, we show the impact on Predicted R^2 from either removal of each of the model assets or adding new assets. It shows that the full model with 5 assets is preferable, having the highest Predicted R^2 equal to 0.98.

Table 2. Model selection in the LTCM analysis.

Asset Name	Max *Predicted* R^2	
	Asset removed	Asset added
EMU Corp Bonds	0.861	
EMU Govt Bonds	0.857	
US Gov Long Bonds	-0.538[1]	
US Corp Long Bonds	0.839	
Mortgages	0.965	
European Stocks		0.948
US Small Stocks		0.977

We then computed monthly Value-at-Risk (VaR) corresponding to asset exposures in Figure 11 using two years of monthly returns for the asset indices employed in the analysis. Depending on the parameters of calculation (such as decay factor, distribution assumptions, etc.), the 99% systematic VaR for June-August 1998 is in the 30%-55% range. Therefore, 10-50% of the monthly losses sustained by the fund during this period should have been expected if proper VaR methodology was used. As mentioned above, applying DSA to higher frequency data (daily or weekly) could have produced much more accurate estimates of potential losses.

9.3 Replicating a Hedge Fund Index

The purpose of this section is to demonstrate how the DSA methodology developed in previous sections can be employed to replicate the performance of a hedge fund strategy index using generic asset indices.

Most hedge fund database vendors publish performance of hedge fund strategy indices – weighted aggregates of funds within groups representing a similar investment strategy. These indices are readily available from a number of hedge fund database vendors[2]. For our analysis we used monthly returns of the HFR Equity Hedge Index representing the Long/Short category, which is one of the most representative. Below, we provide the category definition from the HFR website:

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

[1] Models with negative cross-validation R^2 are typically rejected as inadequate [19].

[2] Among the most widely used: HFR (Hedge Fund Research) www.hedgefundresearch.com, CSFB/Tremont www.hedgeindex.com, Eurekahedge www.eurekahedge.com, and others.

The HFR Equity Hedge index represents an equal-weighted composite of the 615 funds in the category. Even though individual hedge funds engage in frequent trading and utilize derivatives, our contention is that in the index, specific risk is diversified and its returns can be explained by a handful of systematic factors. Since most hedge funds in the category invest in equities, we used the following indices for our analysis: 6 Russell Equity Indices (Top 200 Value/Growth, Midcap Value/Growth, Russell 2000 – Small Cap Value/Growth) as proxies for US Equities, and MSCI EAFE Index as the proxy for international equities and ADRs. We used the Merrill Lynch 3-Month TBill index as a proxy for cash. Monthly returns for 7 years from July 1999 to June 2006 of both the hedge fund index and generic asset indices were used.

The results of DSA analysis (6) corresponding to the optimal value of parameter λ is shown in Figure 13. The quality of regression fit is very high: $R^2 = 0.98$ and *Predicted* $R^2 = 0.90$.

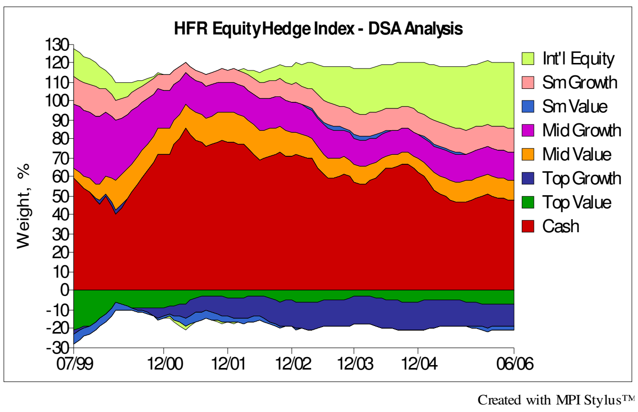

Figure 13. Equity Hedge Index: DSA analysis

In Figure 14 we present values of criterion $PR^2(\lambda)$ corresponding to various values of coefficient λ plotted along the logarithmic X-axis. Note that the results in Figure 13 were obtained using λ corresponding to the highest $PR^2(\lambda)$. The analysis of exposure levels in Figure 13 present several interesting observations. First, the average leverage level in this category (as measured by the magnitude of short exposures below the X-axis) is relatively small and stable. Note also that market exposure has increased dramatically over 2005-2006 (especially to international equity markets represented by EAFE index), almost to year 2000 levels.

After creating the "in-sample" replication described above of the Equity Hedge index, we used the same methodology to replicate the index "out-of-sample." We used the first 60 months of data from July 1999 through June 2004 to determine allocations to generic indices as of June 2004. We then computed the return for the replication portfolio of generic indices for July 2004 using index returns for July 2004 and allocations estimated for June 2004 via DSA. We then expanded the estimation interval to include 61 months through July 2004 and estimated the replication portfolio return for August 2004. We then repeated the process for each of the remaining months through June 2006.

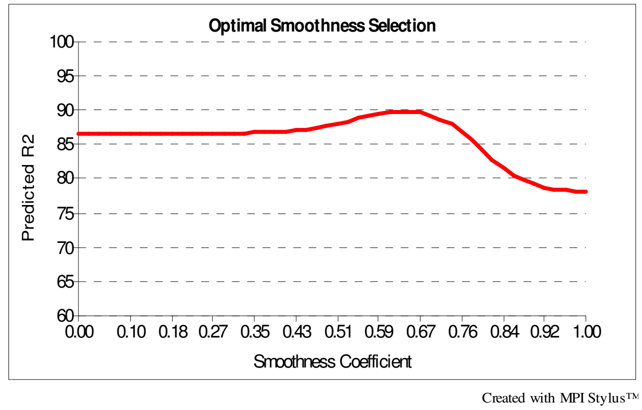

Figure 14. Equity Hedge Index: smoothness selection.

In Figure 15 we show cumulative performance for the HFR Equity Hedge Index and its "out-of-sample" replication ("Benchmark") for the two years July 2004 – June 2006. It is clear that index performance has been replicated very closely.

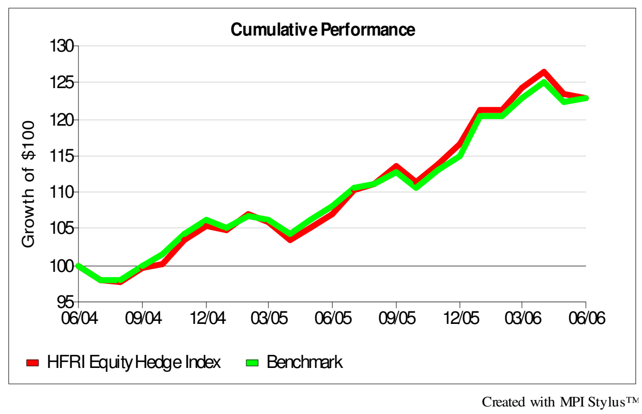

Figure 15. Equity Hedge Index: performance replication

In Figure 16 we compare allocations of both "in-sample" DSA analysis of the Equity Hedge Index (equivalent to the one in Figure 13) and allocations of its "out-of-sample" replication. Note that the latter starts 60 months after the start of the data sample and the asset weight estimates are more volatile than the former.

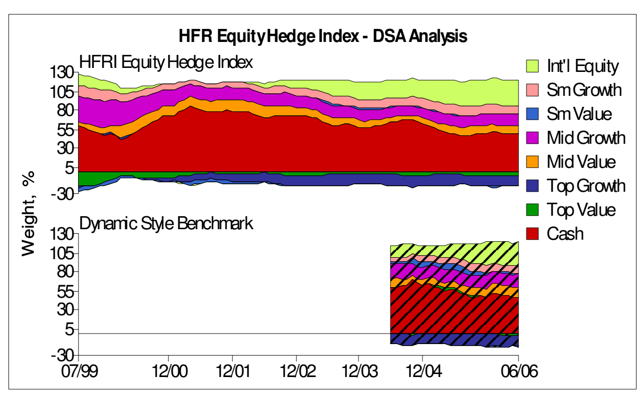

Figure 16. Equity Hedge Index: replication portfolio

10 Conclusions

In this paper we provide a framework for truly dynamic analysis of hedge funds. The proposed Dynamic Style Analysis approach is implemented as a parametric quadratic programming model with a single parameter controlling the smoothness of estimated asset weights on the one hand, and regression fit on the other.

Typically, there's no prior information available about dynamics of hedge fund asset weights or factor exposures, which makes it important to have a way to determine the optimal value of the smoothness parameter based on available return data. In addition, there is a need to measure the model adequacy, because by adjusting the smoothness parameter it is easy to achieve a perfect fit.

Instead of the traditionally used coefficient of determination R^2 (7), we use here the "leave-one-out" cross-validation and *Predicted* R^2 to solve both of the above-mentioned issues. We illustrate our approach using a model hedge fund portfolio as well as three real-life case studies as examples of detecting leverage and changing exposures of hedge funds. Aside from its use in the parameter selection above, the *Predicted* R^2 serves as a measure of the model validity. Similarly to the PRESS statistic that it is based on, it can be used to select the best set of factors as the one providing the highest value of the *Predicted* R^2.

A modification of the Kalman filter-smoother by incorporating the "leave-one-out" procedure has allowed us to escape the seemingly unavoidable loss of the linear computational complexity with respect to the length of the time series. The proposed technique for implementing quadratic optimization algorithms is very practical and can be executed on a personal computer.

The proposed DSA approach has made it possible to fundamentally improve the existing RBSA methodology currently employed in Finance. It results in increased transparency and better hedge fund due diligence which is of crucial importance for financial institutions today.

References

[1] W.F. Sharpe. Determining a fund's effective asset mix. *Investment Management Review*, Vol. 2, 1988, No. 6 (December), pp. 59-69.

[2] W.F. Sharpe. Asset allocation: Management style and performance measurement. *The Journal of Portfolio Management*, Winter 1992, pp. 7-19.

[3] W. Fung, D.A. Hsieh. Empirical characteristics of dynamic trading strategies: The case of hedge funds. *Review of Financial Studies*, 1997, V.10, pp. 275–302.

[4] W. Fung, D.A. Hsieh. Hedge fund benchmarks: A risk based approach. *Financial Analysts Journal*, 2004, Vol. 60, p. 65-80.

[5] R. Kalaba, L. Tesfatsion. Time-varying linear regression via flexible least squares. *International Journal on Computers and Mathematics with Applications*, 1989, Vol. 17, pp. 1215-1245.

[6] C. Wells *The Kalman Filter in Finance*. Kluwer Academic Publishers, 1996.

[7] V. Vapnik. Principles of risk minimization for learning theory. In D.S. Lippman, J.E. Moody, and D.S. Touretzky, editors, *Advances in Neural Information Processing Systems 4*, pp. 831-838. Morgan Kaufmann, 1992.

[8] V. Vapnik. *Statistical Learning Theory*. John-Wiley & Sons, Inc. 1998.

[9] M. Markov, V. Mottl, I. Muchnik. *Principles of Nonstationary Regression Estimation: A New Approach to Dynamic Multifactor Models in Finance*. DIMACS Technical Report 2004-47. Rutgers University, USA, October 2004.

[10] D. Allen. Mean square error of prediction as a criterion for selecting variables. Technometrics, 1971, 13, 469-475.

[11] M. Stone Cross-validatory choice and assessment of statistical predictions. *Journal of Royal Statistical Soc.*, 1974, B, 36, pp.111-133.

[12] D.A.Conway, R.R. Marc. Stable factors in security returns: Identification through cross validation. *Journal of Business & Economic Statistics*, 6, January 1988, pp. 1-15.

[13] P. Lachenbruch. An almost unbiased method of obtaining confidence intervals for the probability of misclassification in discriminant analysis. *Biometrics*, 1967, Vol.23, No. 4, pp.639-645.

[14] A. Kostin, O. Krasotkina, M. Markov, V. Mottl, I. Muchnik. Dynamic programming algorithms for analysis of nonstationary signals. *Computational Mathematics and Mathematical Physics*, 2004, Vol. 44, No. 1, pp. 62-77.

[15] M. Markov, O. Krasotkina, V. Mottl, I. Muchnik. Time-varying regression model with unknown time-volatility for nonstationary signal analysis. *The 8th IASTED International Conference on Signal and Image Process*sing. Honolulu, Hawaii, August 14-16, 2006.

[16] Ph. Jorion. Risk management lessons from long-term capital management. *European Financial Management*, September, 2000.

[17] *Hedge Funds, Leverage, and the Lessons of Long-Term Capital Management*. Report of the President's Working Group on Financial Markets, April 1999, Washington, DC.

[18] *The Wall Street Journal*, 1998-1999.

[19] M. Baroni, et. al. Predictive ability of regression models. Part II: selection of the best predictive PLS model. *Journal of Chemometrics*, 2005, Vol. 6, No. 6, pp. 347-356.