Journal of Statistical Computation and Simulation
iFirst, 2012, 1–21

Detecting structural breaks in multivariate financial time series: evidence from hedge fund investment strategies

Loukia Meligkotsidou[a]* and Ioannis D. Vrontos[b]

[a]*Department of Mathematics, University of Athens, Panepistemiopolis, Athens, Greece;* [b]*Department of Statistics, Athens University of Economics and Business, Pattision 76, Athens 10434, Greece*

(*Received 7 June 2012; final version received 14 October 2012*)

This paper extends the class of asset-based style factor models with multiple structural breaks to the multivariate setting. We propose a model that allows for the presence of common breaks in a system of factor models for individual hedge fund investment strategies, which share common investment characteristics. We develop a Bayesian approach to inference for the unknown number and positions of the structural breaks, based on a set of filtering recursions similar to those of the forward–backward algorithm. Furthermore, we identify relevant risk factors, common among the series of hedge funds, using a Bayesian model comparison approach. We apply our method to a set of correlated hedge fund strategies, which are mainly characterized by equity related bets. Multiple common breaks are identified, consistent with well-known market events, which reveal evidence for structural changes in the risk exposures as well as in the correlation structure of the analysed series.

Keywords: Bayesian inference; forward–backward algorithm; multivariate models; risk factors; structural breaks

1. Introduction

Hedge funds have recently attracted great interest from investors and researchers as flexible alternative investment vehicles which differ substantially from traditional regulated investments, such as mutual funds. Hedge funds are administrated by professional investment managers who usually follow complex trading strategies which are not restricted to the use of leverage, short-selling and derivatives. Hence, they have great flexibility with respect to the types of securities they hold and can take large, often extreme, positive and negative positions in bonds, currencies and other asset classes. Hedge fund strategies are highly dynamic and they often change their positions. Due to the special features of hedge funds, the hedge fund return-generating process may exhibit a high degree of non-normality, fat tails and skewness [1,2] and be characterized by nonlinearities [3]. The nonlinearities in the hedge fund return-generating process can be attributed to market events and financial crises having occurred over the last few decades.[1]

*Corresponding author. Email: meligots@math.uoa.gr

ISSN 0094-9655 print/ISSN 1563-5163 online

http://dx.doi.org/10.1080/00949655.2012.740671
http://www.tandfonline.com

In this paper, we are interested in constructing multivariate risk factor models for jointly analysing the return series of a number of hedge fund strategies which share common investment characteristics. The proposed models are developed with the aim to take account of the specific features of hedge fund returns, that is, dynamic strategies, nonlinear nature and occurrence of market events, and also to exploit information from multiple return series. To this end, we introduce a class of multivariate break-point risk factor models, which capture nonlinearities in the data-generating process and allow for time-varying risk exposures to market risk factors, as well as for time-varying correlations. These models are useful for analysing multiple correlated series of returns, which share common risk–return characteristics and in which common structural breaks occur due to certain events. The proposed risk factor models have the form of seemingly unrelated regressions that capture the correlation structure of the data and allow us to combine information from multiple series, rather than analyse the return series independently. This combined information is used for identifying the number and the positions of multiple structural breaks, corresponding to major market events (shocks) which affected the analysed strategies. These market shocks will be detected through changes in the exposures of the hedge fund strategies in various risk factors and/or through changes in the correlation structure of the series, reflecting the divergence or unification of strategies.

The proposed class of models is a generalization of the univariate break-point model of Fung and Hsieh [3], which is a linear to asset-based style (ABS hereafter) factor model with dynamic coefficients. Extending previous work on univariate break-point risk factor models for hedge funds [3,7,8], we use ABS factors to construct multivariate linear risk factor models in which the risk exposures and the correlation/covariance matrix are subject to multiple structural breaks of unknown number and occurrence times.

A vast amount of work has been done in the econometric/financial literature on the study of structural change in univariate linear models. Several tests for a single structural break of unknown time have been proposed, based on F-statistics [9,10] or cumulative and moving sums of recursive or ordinary least-squares residuals (see e.g. [11–13]). Recently, sophisticated approaches to testing for multiple structural breaks in linear regression have been proposed. Bai and Perron [14,15] considered theoretical and computational issues regarding the estimation and testing for multiple breaks. They proposed a dynamic programming approach for the global minimization of the overall sum of squared residuals and a sequential testing procedure for selecting the optimal number of breaks. Perron and Qu [16] extended this work to models with linear restrictions on the coefficients. Meligkotsidou and Vrontos [8] presented a Bayesian approach to detecting multiple structural breaks in linear factor models. This approach provides exact posterior distributions for the number and the positions of the breaks and can be extended to identifying the most relevant factors in the regression model. The method has been found to be able to identify breaks that could not be captured by previous methods (see [8]).

A much smaller amount of work has been done on the study of structural breaks in a multivariate setting. Bai *et al.* [17] considered quasi-maximum-likelihood estimation for a single break in multivariate regression models. Bai [18] also used a quasi-maximum-likelihood approach for the estimation of break-points in a stationary vector autoregressive model. Finally, Qu and Perron [19] recently considered theoretical and computational issues related to inference on multiple structural breaks of unknown occurrence times in systems of equations. They proposed a quasi-maximum-likelihood estimation method for the break-dates and a likelihood ratio testing procedure for selecting the optimal number of breaks.

In this paper, we extend the Bayesian approach of Meligkotsidou and Vrontos [8] in order to provide a method for inference on multivariate break-point risk factor models. The proposed models are appealing for modelling multiple correlated series of single strategy hedge fund returns that in principle should share some common investment characteristics, that is, a major exposure in a particular asset class, and therefore could be driven by common explanatory risk factors and possibly be affected by the same market events. Multivariate break-point models that use information from a group of series are able to capture common structural changes in the risk exposures of strategies, as well as changes in the volatilities and the correlation structure, due to the adjustments of the hedge fund portfolios which naturally follow market events. The aims of this work are several. First, we aim at detecting common structural breaks, associated with important market events, in a group of hedge fund return series and at inferring their number and their positions. Second, we are interested in understanding the correlation structure of the analysed data and examining whether there exist significant differences in the correlations among different subsets of the sample. Third, we aim at identifying a common set of risk factors which are most likely to explain the risk–return behaviour of all the analysed strategies. In terms of model fit, there will be a tradeoff between the number of factors and the number of structural breaks in the model, which suggests that, under the Bayesian model comparison approach, the most relevant risk factors will be included and the breaks driven by major market events will be identified. Finally, it is of interest to investigate if the proposed approach of jointly modelling multiple correlated series that share common investment characteristics is economically useful.

We follow a probabilistic approach to modelling multiple breaks which lies within the Bayesian framework. Under the Bayesian approach to inference all the unknown quantities in the model are treated as parameters. Therefore, the joint posterior distribution summarizes all the available information about the number and the positions of the breaks as well as about the model parameters. Furthermore, the uncertainty associated with the positions of the breaks is conveyed by their posterior distribution. Finally, under the Bayesian approach, variable selection can be easily incorporated in the model comparison exercise by simply extending the set of competing models to include model specifications with different sets of factors.

Our approach is based on a set of efficient recursions recently introduced by Fearnhead [20], which are similar to those of the forward–backward algorithm (for a description of the forward–backward algorithm, see [21], and for recent extensions, see [22,23]). The recursions have been previously applied to signal segmentation problems by Fearnhead [24]. The forward recursion is used to calculate the marginal likelihoods of the break-point models, provided that the segment-specific model parameters can be integrated with respect to their prior distributions. These marginal likelihoods can be used for inference on the number of breaks in a Bayesian model comparison setting. The backward recursion enables us to simulate from the marginal posterior distribution of the positions of the breaks. The advantage of this approach over other Bayesian approaches which are based on Markov chain Monte Carlo (MCMC) methods is that it produces exact inferences on the number and the positions of the breaks.

First, we show how the forward–backward recursions can be used for inference on the number and the positions of multiple structural breaks in multivariate risk factor models, conditional on a specific set of factors. Next we consider the full model comparison problem where the uncertainty about which factors should be included in the model is taken into account together with the uncertainty about the number of structural breaks. If there are K factors available, then there are 2^K different possible sets of factors each of which defines a different model specification. For each model specification we use a separate set of recursions to compute the marginal likelihoods of the respective break-point models. Then, it is straightforward to compute the posterior model probabilities for all of these models. These can be used either for selecting the most probable model (i.e. making inference conditional on the number of breaks and risk factors) or for Bayesian model averaging (i.e. producing inferences averaged over the different models).

We have applied our approach to a group of four hedge fund single strategy indices which have similar investment characteristics, in particular exposures to equity risk factors. The four series of returns are correlated, and independent univariate analyses have shown that they are affected by some common market events. The multivariate analysis revealed that common structural breaks exist in these series, with break-dates close to important market events, such as the long-term capital management (LTCM) crisis and the Technology and Internet bubble and the events of September 2001. The structural breaks usually correspond to substantial changes in the risk exposures of all series, the risk-adjusted performances of the series, as well as their correlations. Finally, we have conducted an out-of-sample exercise with the aim to show potential impacts and benefits of our approach over approaches that ignore either the presence of structural breaks in the series (i.e. simple multivariate regression models) or the correlation structure of the data (i.e. independent univariate analyses).

The contributions of our work are several. First, we employ an exact and efficient Bayesian approach to inference with the aim to detect structural breaks in multivariate series of hedge fund returns. Identifying common structural breaks can be useful for understanding the performance and the correlation/covariance structure of certain hedge funds in time and investigating whether the risk exposures and the correlations change with market conditions. Second, we search for the relevant risk factors which best describe the multivariate series of returns, taking into account the possible presence of one or multiple structural breaks in the series. Third, working in a Bayesian model comparison framework, we are able to compute posterior model probabilities for a large number of competing models, as well as to account for model misspecification by calculating the marginal posterior distributions of the risk factor set and the number of breaks. Finally, the proposed Bayesian approach enables us to obtain the posterior distributions of the model parameters, that is, the alphas, the betas and the correlation matrices in different segments. Thus, we are able to compute reliable estimates for the model parameters and assess the parameter uncertainty.

The paper proceeds as follows. In Section 2, we present our multivariate break-point risk factor models, and in Section 3, we describe the proposed Bayesian approach to inference on the number and the positions of the breaks. In Section 4, we report the results of our empirical application using hedge fund single strategy indices. Finally, we conclude in Section 5.

2. Multivariate factor models with multiple breaks

In this section, we present a class of multivariate risk factor models subject to multiple structural breaks occurring at unknown time points. These models will be used in applications in this paper to analyse multivariate series of hedge fund returns. In our analysis, it is of great interest to infer the number and the positions of common break-points in multivariate hedge fund return series, as well as to identify common risk factors associated with hedge fund indices which exhibit similar risk–return characteristics. To this end, we follow a Bayesian approach to inference which takes into account the uncertainty about all the unknown quantities of interest and enables us to calculate posterior model probabilities for a number of break-point multivariate risk factor models.

Suppose that we observe T consecutive realizations on q series of hedge fund returns, $\mathbf{R} = (\mathbf{r}_1, \ldots, \mathbf{r}_q)$, where $\mathbf{r}_l = (r_{l1}, \ldots, r_{lT})'$, $l = 1, \ldots, q$, and that we use k risk factors, $\mathbf{f}_i = (f_{i1}, \ldots, f_{iT})'$, $i = 1, \ldots, k$, to explain the multivariate series of returns. We consider a multivariate linear regression model assuming that m break-points have occurred at unknown times $\tau_0 < \tau_1 < \tau_2 < \cdots < \tau_m < \tau_{m+1}$, and we set $\tau_0 = 0$ and $\tau_{m+1} = T$. These break-points define $m + 1$ disjoint segments. We let $\mathbf{R}_{t:s}$ and $\mathbf{f}_{i,t:s}$ denote the hedge fund returns and the ith factor returns, respectively, from time t to time s inclusive. Also, we denote by $\mathbf{F}_{t:s}$ the design matrix of the linear regression for observations from time t to time s, whose first column is a vector of $(s - t + 1)$ ones and the remaining columns contain the returns of the k factors $\mathbf{f}_{i,t:s}$, $i = 1, \ldots, k$.

The observations in the jth segment, that is, the observations $\boldsymbol{R}_{\tau_{j-1}+1:\tau_j}$, are modelled by the following risk factor model:

$$\mathbf{R}_{\tau_{j-1}+1:\tau_j} = \mathbf{F}_{\tau_{j-1}+1:\tau_j}\mathbf{B}_j + \mathbf{E}_{\tau_{j-1}+1:\tau_j}, \quad \mathbf{E}_{\tau_{j-1}+1:\tau_j} = (\epsilon_{\tau_{j-1}+1}, \dots, \epsilon_{\tau_j})', \tag{1}$$

where $\mathbf{R}_{\tau_{j-1}+1:\tau_j}$ is the $(\tau_j - \tau_{j-1}) \times q$ matrix of hedge fund returns, $\mathbf{F}_{\tau_{j-1}+1:\tau_j}$ is the $(\tau_j - \tau_{j-1}) \times (k+1)$ design matrix, $\mathbf{B}_j$ is the $(k+1) \times q$ matrix of the multivariate regression coefficients and the vectors of errors ϵ_t, for $\tau_{j-1} < t \leq \tau_j$, are assumed to be independent and identically normally distributed with zero mean and segment-specific covariance matrix $\mathbf{\Sigma}_j$, $j = 1, \dots, m+1$. The lth column of the matrix $\mathbf{B}_j$, $l = 1, \dots, q$, contains the intercept and the vector of exposures of the lth series to the risk factors, while the error covariance matrix introduces dependences among the q series of hedge fund returns. The common structural breaks in the hedge fund return series reflect changes in all of the model parameters. Note that in the case where $m = 0$ the model reduces to a standard multivariate linear regression model (since there is a single segment).

Suppose now that there are K available risk factors. An interesting problem is that of selecting the subset of most relevant risk factors to be included in the multivariate regression model. This variable selection problem can be considered as a special case of the model selection problem, where each model corresponds to a different subset of factors f_{kt}, $k = 1, \dots, K$. It is convenient to indicate each of these 2^K possible choices of subsets by the vector $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_K)'$, where $\gamma_i = 1$ if the ith factor is included and $\gamma_i = 0$ if not. Thus, the vector $\boldsymbol{\gamma}$ plays the role of a model identifier. Note that the sum of the elements of $\boldsymbol{\gamma}$ is equal to the number of factors k included in the model.

In our study, we deal simultaneously with two different sources of model uncertainty; there is uncertainty about both the number of structural breaks and the factors included in the regression. Assuming that the factors included in the segment-specific regressions in all the $(m+1)$ segments of an m-breaks model are common (which is an assumption that enables us to produce exact posterior model probabilities), the number of models considered is $(M+1)2^K$, where M is a prespecified maximum number of breaks allowed to occur. We employ the approach introduced by Fearnhead [20] to calculate the marginal likelihoods of all the break-point models and infer the number and the positions of the breaks. Having calculated the marginal likelihoods of all possible models, we are able to compare them by calculating the respective posterior model probabilities.

The approach to model selection followed in this paper is exact, that is, the exact marginal posterior probability (MPP) of each model (i.e. the number of breaks and the set of factors included) is computed. This can be done, under the assumption of conjugate prior distributions, within a set of filtering recursions, at the cost of reducing the complexity of the models being considered. Specifically, the set of factors included in each model is common across series of hedge fund returns and across the different segments of the sample. These assumptions are reasonable in the context considered in this paper, that is, the analysis of sets of hedge fund strategies that share common investment characteristics, but have to be relaxed if more general problems are considered. Some work has been done on multivariate regression models allowing for different sets of predictors to be included in different equations (see e.g. [25,26]). However, these models have to be estimated within MCMC or stochastic search algorithms which will be extremely complicated if the study of structural change is jointly treated with the factor identification. Similarly, the model complexity will increase substantially if we allow different factors to be included in different segments, thus making our exact approach impossible to be used. Such complex models are, therefore, beyond the scope of this paper.

3. Bayesian inference via filtering

In this section, we present our Bayesian approach to testing for multiple structural breaks in multivariate linear regression models. We employ the approach introduced by Fearnhead [20] and base our inferences on filtering recursions similar to those of the forward–backward algorithm, which can be used to calculate the marginal likelihoods of multiple break-point models and infer the number and the positions of the breaks. Furthermore, this approach enables us to perform model selection with respect to the common factors included in the model. For any subset of the risk factor set at hand we can use a different set of recursions to calculate the marginal likelihoods of the corresponding multiple break-point model. Having calculated the marginal likelihoods of all possible models (i.e. all possible subsets of factors and all possible numbers of breaks), we are able to compare them in a Bayesian framework, by calculating the respective posterior model probabilities. Our approach is a generalization of that developed by Meligkotsidou and Vrontos [8] for the simple linear regression model with multiple structural breaks.

3.1. *Inference for a specific risk factor model with multiple breaks*

First, we consider multivariate regression models with k specific factors (i.e. specific risk factor identifier $\boldsymbol{\gamma}$), allowing for m structural breaks, $m = 0, 1, \dots, M$, where M is the prespecified maximum number of breaks. We show how to calculate the marginal likelihoods of these models via filtering recursions. In order to perform Bayesian inference, we need to specify prior distributions for the model parameters. The prior specification is in general very crucial in model comparison, as the choice of the prior can affect the marginal likelihoods of the different models considered for having generated the data. As a general principle, flat priors tend to penalize more the models which are more complex and, therefore, it is advisable to use weekly informative priors in model comparison problems [27, Chapter 6].

We assume independent conjugate prior distributions for the parameters of the multivariate regressions in different segments. This choice of prior enables us to integrate out all of the model parameters, and therefore, the calculation of the marginal likelihoods, using the set of filtering recursions, is exact. For each segment-specific matrix of coefficients, $\mathbf{B}_j$, $j = 1, \dots, m+1$, we assume a matrix-variate normal prior distribution, which will be denoted by $N_{(k+1)\times q}(\mathbf{M}_j, \mathbf{V}_j, \boldsymbol{\Sigma}_j)$, where $\mathbf{M}_j$ is the $(k+1) \times q$ matrix of means, $\mathbf{V}_j$ is a $(k+1) \times (k+1)$ symmetric positive definite matrix, and $\boldsymbol{\Sigma}_j$ is the error covariance matrix associated with the jth segment. The probability density function (pdf) of the matrix-variate normal prior distribution for $\mathbf{B}_j$ is given by

$$\pi(\mathbf{B}_j|\boldsymbol{\Sigma}_j) = (2\pi)^{-(k+1)q/2}|\mathbf{V}_j|^{-q/2}|\boldsymbol{\Sigma}_j|^{-(k+1)/2}\exp\left\{-\frac{1}{2}\,\mathrm{tr}[\boldsymbol{\Sigma}_j^{-1}(\mathbf{B}_j-\mathbf{M}_j)'\mathbf{V}_j^{-1}(\mathbf{B}_j-\mathbf{M}_j)]\right\}.$$

Now, for each $\boldsymbol{\Sigma}_j, j = 1, \dots, m+1$, we assume an inverted Wishart prior distribution, which will be denoted by $\mathrm{IW}(\mathbf{S}, v)$, where $\mathbf{S}$ is a $q \times q$ hyperparameter matrix and v denotes the degrees of freedom. The pdf of the inverted Wishart prior distribution for $\boldsymbol{\Sigma}_j$ is given by

$$\pi(\boldsymbol{\Sigma}_j) = \left[2^{vq/2}\pi^{q(q-1)/4}\prod_{l=1}^{q}\Gamma\left(\frac{v+1-l}{2}\right)|\mathbf{S}|^{-v/2}\right]^{-1}|\boldsymbol{\Sigma}_j|^{-(v+q+1)/2}\exp\left[-\frac{1}{2}\,\mathrm{tr}(\boldsymbol{\Sigma}_j^{-1}\mathbf{S})\right],$$

where $v \geq q$.

Model selection problems require weakly informative prior distributions to be specified for the model parameters. To this end, we choose $\mathbf{M}_j = \mathbf{M} = \mathbf{0}$, $j = 1, \dots, m+1$, which reflects prior ignorance about the location of the means of the regression coefficients, and $\mathbf{V}_j = \mathbf{V} = c(\mathbf{F}'_{1:T}\mathbf{F}_{1:T})^{-1}$, which, for the standard multivariate regression model with $m = 0$, replicates the covariance structure of the factors and is akin to the g-prior proposed by Zellner [28], while for $m > 0$ it is still a sensible choice as an approximation to the covariance structure of the factors. The scalar c must be carefully specified; it should be large enough so that the prior is relatively flat over the range of plausible values of the $\mathbf{B}_j$'s, but not too large to avoid putting increasing weight on the null model in the Bayesian model comparison setting. In this paper, we take $c = T$, as in the g-prior of Zellner [28]. For the hyperparameter matrix $\mathbf{S}$, we take $\mathbf{S} = kI_q$, where the value of k is chosen to be comparable with the variance of the data (here we take $k = 0.0003$), while we take $v = q + 2$ so that the prior expectation of $\mathbf{\Sigma}_j$ to be $E(\mathbf{\Sigma}_j) = \mathbf{S}$. A similar prior specification for the parameters of multivariate linear regression models in the context of Bayesian model comparison was used by Brown *et al.* [29].

For the number of structural breaks, we assume a Poisson Po(λ) prior distribution with rate λ, truncated over the range $\{0, 1, \dots, M\}$. Then, we specify a prior for the positions of the breaks conditional on their number. This conditional prior is defined as

$$\pi_m(\tau_1, \dots, \tau_m) = \pi_m(\tau_m)\pi_m(\tau_{m-1}|\tau_m)\dots\pi_m(\tau_1|\tau_2),$$

where $\pi_m(\cdot)$ denotes conditioning on m breaks, $\pi_m(\tau_m)$ is the prior on the position of the last break-point and $\pi_m(\tau_j|\tau_{j+1})$ is the prior on the position of the jth break-point given the position of the $(j+1)$th. In this paper, we take $\lambda = 2$, and we adopt discrete uniform priors for the positions of the break-points. The value of $\lambda = 2$ is chosen to reflect our prior belief that the number of breaks is not too large. This is consistent with the findings of previous studies dealing with structural changes in univariate hedge fund return series (see the discussion of the empirical analysis in Section 4).

As in [8], for all $t \leq s$, we define $P^{(\gamma)}(t, s)$ to be the marginal probability of the data $\mathbf{R}_{t:s}$ under a multivariate linear regression model with k factors and risk factor identifier γ, given that times t and s are in the same segment. The likelihood of the data $\mathbf{R}_{t:s}$ under a multivariate linear regression model parameterized by $\mathbf{B}$ and $\mathbf{\Sigma}$, with risk factor identifier $\boldsymbol{\gamma}$, is given by

$$f(\mathbf{R}_{t:s}|\mathbf{B}, \mathbf{\Sigma}, \boldsymbol{\gamma}) = (2\pi)^{-(s-t+1)q/2}|\mathbf{\Sigma}|^{-(s-t+1)/2} \exp\left\{-\frac{1}{2}\,\mathrm{tr}[\mathbf{\Sigma}^{-1}(\mathbf{R}_{t:s} - \mathbf{F}_{t:s}\mathbf{B})'(\mathbf{R}_{t:s} - \mathbf{F}_{t:s}\mathbf{B})]\right\}.$$

Then, the marginal probability $P^{(\gamma)}(t, s)$ is obtained by integrating the likelihood $f(\mathbf{R}_{t:s}|\mathbf{B}, \mathbf{\Sigma}, \boldsymbol{\gamma})$ with respect to the prior distribution on the model parameters $\pi(\mathbf{B}, \mathbf{\Sigma}) = \pi(\mathbf{B}|\mathbf{\Sigma})\pi(\mathbf{\Sigma})$, that is,

$$\begin{aligned} P^{(\gamma)}(t, s) &= \Pr(\mathbf{R}_{t:s}|t, s \text{ in the same segment}, \boldsymbol{\gamma}) \\ &= \iint f(\mathbf{R}_{t:s}|\mathbf{B}, \mathbf{\Sigma}, \boldsymbol{\gamma})\pi(\mathbf{B}|\mathbf{\Sigma})\pi(\mathbf{\Sigma})\,\mathrm{d}\mathbf{B}\,\mathrm{d}\mathbf{\Sigma} \\ &= \pi^{-(s-t+1)q/2}|\mathbf{V}^*|^{q/2}|\mathbf{V}|^{-q/2}|\mathbf{S}|^{v/2}|\mathbf{S}^*|^{-(s-t+v+1)/2}\prod_{l=1}^{q}\left[\frac{\Gamma((v^*+1-l)/2)}{\Gamma((v+1-l)/2)}\right], \end{aligned}$$

where

$$\begin{aligned} \mathbf{V}^* &= (\mathbf{V}^{-1} + \mathbf{F}'_{t:s}\mathbf{F}_{t:s})^{-1}, \\ \mathbf{M}^* &= \mathbf{V}^*(\mathbf{V}^{-1}\mathbf{M} + \mathbf{F}'_{t:s}\mathbf{R}_{t:s}), \\ \mathbf{S}^* &= \mathbf{S} + \mathbf{R}_{t:s}{}'\mathbf{R}_{t:s} + \mathbf{M}'\mathbf{V}^{-1}\mathbf{M} - (\mathbf{M}^*)'(\mathbf{V}^*)^{-1}(\mathbf{M}^*), \\ v^* &= v + (s - t + 1). \end{aligned}$$

We will show now how the marginal likelihood of model (1), given the number of structural breaks m, can be calculated via a set of recursions. For $j = 1, \dots, m$ and for $t = j + 1, \dots, T -$

$m+j$, we define $Q_j^{(m,\boldsymbol{\gamma})}(t)$ to be the probability of $\mathbf{R}_{t:T}$ under a multivariate linear regression model with factors specified by $\boldsymbol{\gamma}$, conditional on m break-points and given that the position of the jth break is at time $t-1$, that is,

$$Q_j^{(m,\boldsymbol{\gamma})}(t) = \Pr(\mathbf{R}_{t:T}|\tau_j = t-1, m \text{ break-points}, \boldsymbol{\gamma}).$$

The calculation of the marginal likelihood of the model with m breaks, as well as inference on the positions of the breaks, will be based on the following set of recursions. For $t = m+1, \ldots, T$,

$$Q_m^{(m,\boldsymbol{\gamma})}(t) = P^{(\boldsymbol{\gamma})}(t,T)\pi_m(\tau_m = t-1).$$

For $j = 1, \ldots, m-1$ and $t = j+1, \ldots, T-m+j$,

$$Q_j^{(m,\boldsymbol{\gamma})}(t) = \sum_{s=t}^{T-m+j} P^{(\boldsymbol{\gamma})}(t,s)Q_{j+1}^{(m,\boldsymbol{\gamma})}(s+1)\pi_m(\tau_j = t-1|\tau_{j+1} = s),$$

(for the proof, see [20]). The marginal likelihood of the multivariate risk factor model with m break-points is given by

$$\Pr(\mathbf{R}_{1:T}|m,\boldsymbol{\gamma}) = \sum_{s=1}^{T-m} P^{(\boldsymbol{\gamma})}(1,s)Q_1^{(m,\boldsymbol{\gamma})}(s+1).$$

Note that the computational complexity of the above algorithm is quadratic in T. Details regarding the computational cost of the forward–backward recursions, as well as suggestions of how the cost can be reduced by introducing some approximation to the approach can be found in [20].

For any value of m, the marginal likelihood for the model can be evaluated using the recursions. The probabilities $P^{(\boldsymbol{\gamma})}(t,s)$ need to be calculated only once and stored in an upper triangular matrix to be used within any set of recursions. Then, the posterior probabilities of the models with $m = 0, \ldots, M$ break-points are given by

$$\Pr(m|\mathbf{R}_{1:T},\boldsymbol{\gamma}) = \frac{\pi(m)\Pr(\mathbf{R}_{1:T}|m,\boldsymbol{\gamma})}{\sum_{j=0}^{M}\pi(j)\Pr(\mathbf{R}_{1:T}|j,\boldsymbol{\gamma})}, \quad (2)$$

where the calculation of the marginal likelihood of the standard multivariate regression model with no breaks ($m=0$) is given by $P^{(\boldsymbol{\gamma})}(1,T)$. Note that $\Pr(m|\mathbf{R}_{1:T},\boldsymbol{\gamma})$ is the posterior distribution of the number of breaks, and samples from this discrete distribution can be easily drawn.

Conditional on m structural breaks, samples from the joint posterior distribution of their positions can be also obtained, given the values of $Q_j^{(m,\boldsymbol{\gamma})}(t)$, $j = 1, \ldots, m$, $t = j+1, \ldots, T-m+j$. The posterior distribution of the position of the first break is

$$\Pr(\tau_1|\mathbf{R}_{1:T},m,\boldsymbol{\gamma}) \propto P^{(\boldsymbol{\gamma})}(1,\tau_1)Q_1^{(m,\boldsymbol{\gamma})}(\tau_1+1), \quad (3)$$

$\tau_1 = 1, \ldots, T-m$. Then, the posterior distribution of τ_j given τ_{j-1} is given by

$$\Pr(\tau_j|\tau_{j-1},\mathbf{R}_{1:T},m,\boldsymbol{\gamma}) \propto P^{(\boldsymbol{\gamma})}(\tau_{j-1}+1,\tau_j)Q_j^{(m,\boldsymbol{\gamma})}(\tau_j+1)\pi_m(\tau_{j-1}|\tau_j), \quad (4)$$

for $\tau_j = \tau_{j-1}+1, \ldots, T-m+j$. Note that the model paremeters, that is, the $\mathbf{B}_j$'s and the Σ_j's, have been integrated out from the joint posterior distribution of the positions of the breaks.

Given a sample from this joint posterior distribution we can easily simulate draws from the conditional posterior distributions of the model parameters. For the jth segment, the marginal posterior distribution of Σ_j, after integrating out $\mathbf{B}_j$, is

$$f(\Sigma_j|\mathbf{R}_{1:T}, m, \boldsymbol{\gamma}, \tau_{j-1}, \tau_j) \equiv \mathrm{IW}(\mathbf{S}_j^*, v_j^*), \tag{5}$$

while the conditional posterior distribution of $\mathbf{B}_j$ is given by

$$f(\mathbf{B}_j|\mathbf{R}_{1:T}, m, \boldsymbol{\gamma}, \tau_{j-1}, \tau_j, \Sigma_j) \equiv N(\mathbf{M}_j^*, \mathbf{V}_j^*, \boldsymbol{\Sigma}_j), \tag{6}$$

where

$$\begin{aligned}
\mathbf{V}_j^* &= (\mathbf{V}^{-1} + \mathbf{F}'_{\tau_{j-1}:\tau_j}\mathbf{F}_{\tau_{j-1}:\tau_j})^{-1},\\
\mathbf{M}_j^* &= \mathbf{V}_j^*(\mathbf{V}^{-1}\mathbf{M} + \mathbf{F}'_{\tau_{j-1}:\tau_j}\mathbf{R}_{\tau_{j-1}:\tau_j}),\\
\mathbf{S}_j^* &= \mathbf{S} + \mathbf{R}_{\tau_{j-1}:\tau_j}{}'\mathbf{R}_{\tau_{j-1}:\tau_j} + \mathbf{M}'\mathbf{V}^{-1}\mathbf{M} - (\mathbf{M}_j^*)'(\mathbf{V}_j^*)^{-1}(\mathbf{M}_j^*),\\
v_j^* &= v + (\tau_j - \tau_{j-1} + 1).
\end{aligned}$$

Therefore, a Gibbs sampler can be constructed to perform inference on all the unknown quantities in the model, consisting of the following steps:

(1) Simulate from the marginal posterior distribution of the number of breaks, m, given by Equation (2).
(2) Simulate from the joint posterior distribution of the positions of the breaks, given m, using the recursion in Equations (3) and (4).
(3) For $j = 1, \dots, m + 1$, simulate Σ_j from its marginal posterior distribution given by Equation (5).
(4) For $j = 1, \dots, m + 1$, simulate $\mathbf{B}_j$ from its conditional posterior distribution given by Equation (6).

In this sampler, the number of breaks and their positions are updated as a block by first simulating m from its marginal posterior distribution and then simulating the positions given m. Therefore, it is possible to first obtain large samples of draws from the joint posterior of m and $\tau_1, \dots, \tau_m$, having integrated out the remaining parameters, and then iteratively simulate the Σ_j's and the $\mathbf{B}_j$'s from their posterior distributions given the sampled draws of m and $\tau_1, \dots, \tau_m$. Efficient simulations of large samples from the discrete posterior distribution of m can be performed using the algorithm of Carpenter *et al.* [30]. Furthermore, simulations of large samples from the joint posterior distribution of the positions of the breaks, using the recursive scheme in Equations (3) and (4), can be performed in an efficient way as described in [20].

3.2. *Inference under uncertainty about the risk factors*

Here, we consider the full model comparison problem where the uncertainty about which factors should be included in the multivariate regression model is taken into account together with the uncertainty about the number of structural breaks. In other words, we consider both the number of breaks m and the risk factor identifier $\boldsymbol{\gamma}$ as random. In the Bayesian setting, this means that $\boldsymbol{\gamma}$ is treated as a further parameter in the model and, therefore, has to be assigned a prior distribution. A natural prior distribution for $\boldsymbol{\gamma}$, commonly used in variable selection problems, is based on specifying a prior probability ρ_k for the inclusion of a factor in a model with k factors. Hence, the prior distribution we adopt for $\boldsymbol{\gamma}$ is of the form $\pi(\boldsymbol{\gamma}) = \prod_{k=1}^{K} \rho_k^{\gamma_k}(1 - \rho_k)^{1-\gamma_k}$, where $\rho_k \in [0, 1]$ indicating that each risk factor f_k, $k = 1, \dots, K$, enters the model independently of the others with probability $\pi(\gamma_k = 1) = \rho_k$. We assume that we do not have any prior information for the inclusion of a specific factor in the model, thus we take $\rho_k = 0.5$, $k = 1, \ldots, K$, which amounts to assuming that each vector $\boldsymbol{\gamma}$ is assigned equal prior probability $\pi(\boldsymbol{\gamma}) = 1/2^K$. This implies that the investor is neutral about the factors that will enter in the model.

For each specific $\boldsymbol{\gamma}$, we can perform all the calculations described in the previous section in order to calculate the marginal likelihoods of the break-point models with m structural breaks, for $m = 0, 1, \ldots, M$. Having calculated the marginal likelihoods of these $(M+1)2^K$ models, it is straightforward to compute the posterior model probabilities of the models; the posterior probability of the model with risk factor identifier $\boldsymbol{\gamma}$ and m break-points is given by

$$\Pr(m, \boldsymbol{\gamma}|\mathbf{R}_{1:T}) = \frac{\pi(m)\pi(\boldsymbol{\gamma})\Pr(\mathbf{R}_{1:T}|m, \boldsymbol{\gamma})}{\sum_{\boldsymbol{\delta}\in\boldsymbol{\Gamma}}\sum_{j=0}^{M}\pi(j)\pi(\boldsymbol{\delta})\Pr(\mathbf{R}_{1:T}|j, \boldsymbol{\delta})}, \tag{7}$$

for $m = 0, \ldots, M$, $\boldsymbol{\gamma} \in \boldsymbol{\Gamma}$, where Γ denotes the set of all risk factor identifiers. Conditional on a specific model with risk factor identifier $\boldsymbol{\gamma}$ and m break-points, for example, the most probable model, samples from the posterior distribution of the positions of the breaks can be obtained as described in Section 3.1.

Furthermore, we are able to compute MPPs for the number of breaks or the risk factor identifier. For the former we just need to average over all different risk factor identifiers in Γ, while for the latter we need to average over all different numbers of breaks $m = 0, 1, \ldots, M$. This approach, which is usually referred to as Bayesian model averaging, produces posterior probabilities which are robust to model misspecification. That is, the marginal posterior distribution of the number of breaks

$$\Pr(m|\mathbf{R}_{1:T}) = \frac{\sum_{\boldsymbol{\gamma}\in\boldsymbol{\Gamma}}\pi(m)\pi(\boldsymbol{\gamma})\Pr(\mathbf{R}_{1:T}|m, \boldsymbol{\gamma})}{\sum_{\boldsymbol{\delta}\in\boldsymbol{\Gamma}}\sum_{j=0}^{M}\pi(j)\pi(\boldsymbol{\delta})\Pr(\mathbf{R}_{1:T}|j, \boldsymbol{\delta})}, \quad m = 0, \ldots, M,$$

is robust to misspecification of the risk factor identifiers, since we have integrated over the uncertainty about $\boldsymbol{\gamma}$. Similarly, the marginal posterior distribution of $\boldsymbol{\gamma}$ is obtained by summing the left-hand side of Equation (7) with respect to m.

4. Data and empirical analysis

In this section, we present an empirical application of the proposed Bayesian method to hedge fund single strategy indices from Hedge Fund Research (HFR). Our study is more relevant to style allocation decisions, where managers allocate funds among various hedge fund strategies, as in [31–33] among others. From the wide variety of HFR style indices, we select a group of indices that in principle should share some common investment characteristics, that is, a major exposure in a particular asset class, and therefore could be driven by common explanatory risk factors and possibly are affected by the same market events. We consider four HFR single strategy indices: Equity Hedge (EH), Equity Market Neutral (EMN), Merger Arbitrage (MA) and Distressed Securities (DS). These indices are mainly characterized by equity elated bets. Our study of these hedge funds uses net-of-fee monthly excess returns (in excess of the 3-month US Treasury Bill) from January 1994 to November 2005.[2] This period includes a number of crises and market events which affected hedge fund returns: the abrupt increase in the US interest rates in early 1994, the Asian crisis in July 1997, the LTCM crisis in late 1998, the Technology and Internet bubble in March 2000, the Japanese crisis in March 2001 and the events of September 2001.

Several variables have been used in the hedge fund literature as possible risk factors to explain hedge fund returns. While there is no consensus on the actual variables that are appropriate for pricing hedge fund returns, many studies agree that a complete set should cover all markets that hedge funds invest in (see e.g. [3]). In this analysis, we model the hedge fund returns by using the seven ABS factors of Fung and Hsieh [3] which have been shown to be valuable explanatory variables for fund of funds and hedge funds returns (see, for example, the studies of Fung and Hsieh [34–36], Kosowski *et al.* [37], Fung *et al.* [7], and Meligkotsidou and Vrontos [8]). This set of factors includes two equity factors, namely the excess returns on the S&P 500 index (S&P) and the spread between small-cap and large-cap stock returns (SCMLC), constructed as the difference of the Wilshire Small Cap 1750 index returns and the Wilshire Large Cap 750 index returns, which are the most important risk factors for a large number of hedge funds. In addition, it includes two fixed income factors, the change in the 10-year Treasury yields (10Y) and the difference between the change in the Moody's Baa bonds yields and the change in the 10-year T-bonds yields (CredSpr); the latter factor can be regarded as the excess log return of the Baa bond over the 10-year one. This excess return is included in factor models to capture an additional risk premia factor to that associated with the 10-year government bond. Finally, we use three primitive trend-following (PTF) risk factors, that is, three portfolios of lookback straddles on bonds (PTFSBD), currencies (PTFSFX) and commodities (PTFSCOM),[3] which have the ability to explain the returns of trend-following funds.

4.1. *Analysis of hedge fund strategies*

In this section, we analyse monthly data on four HFR single strategy hedge fund indices over the period from January 1994 to November 2005 using multivariate ABS factor models. A brief discussion on the characteristics of the analysed hedge fund return series follows. Table 1 (Panel A) presents summary statistics for the four series of hedge fund returns. We observe that hedge fund series are heterogeneous; EH has relatively high average returns and high volatility, DS and MA have relatively low average returns and standard deviations, while EMN has low average returns

Table 1. Descriptive statistics of monthly returns for the four analysed HFR single strategy indices.

Hedge fund index	Mean	SD	Median	Skewness	Kurtosis	Minimum	Maximum
Panel A: January 1994–November 2005							
EH	0.0083	0.0258	0.0091	0.2434	4.8028	−0.0806	0.1045
EMN	0.0032	0.0086	0.0029	0.1417	3.9163	−0.0208	0.0322
DS	0.0065	0.0159	0.0076	−1.5155	11.3114	−0.0891	0.0468
MA	0.0045	0.0101	0.0061	−2.4285	14.7142	−0.0610	0.0200
Panel B: January 1994–September 1998							
EH	0.0090	0.0232	0.0092	−0.7714	5.4034	−0.0806	0.0526
EMN	0.0042	0.0078	0.0043	−0.5084	3.7915	−0.0208	0.0176
DS	0.0046	0.0175	0.0075	−2.9301	16.1817	−0.0891	0.0357
MA	0.0057	0.0116	0.0087	−3.4789	20.1175	−0.0610	0.0200
Panel C: November 1998–March 2000							
EH	0.0301	0.0345	0.0276	0.5976	2.8927	−0.0280	0.1045
EMN	0.0027	0.0135	0.0015	0.3847	2.5112	−0.0172	0.0322
DS	0.0083	0.0157	0.0049	0.8308	3.3275	−0.0140	0.0468
MA	0.0094	0.0067	0.0096	−0.0691	1.7083	−0.0014	0.0195
Panel D: April 2000–November 2005							
EH	0.0020	0.0221	0.0038	−0.2892	2.4507	−0.0482	0.0483
EMN	0.0025	0.0076	0.0025	0.4417	4.4389	−0.0199	0.0254
DS	0.0076	0.0146	0.0080	−0.0814	2.6455	−0.0301	0.0391
MA	0.0022	0.0089	0.0042	−1.1561	4.7946	−0.0292	0.0199

Notes: Panel A contains descriptive statistics based on period January 1994–November 2005, Panel B contains descriptive statistics based on period January 1994–September 1998, Panel C contains descriptive statistics based on period November 1998–March 2000, Panel D contains descriptive statistics based on period April 2000–November 2005. The summary statistics include mean, standard deviation (SD), median, skewness, kurtosis, minimum and maximum.

and low standard deviation. Differences are also apparent in higher order moments. In particular, MA and DS have negative skewness and high kurtosis, while EH and EMN have low positive skewness and relatively small kurtosis.

The four series of returns are correlated. Table 2 (Panel A) presents the correlation coefficients between pairs of single strategy hedge fund returns. We can see that the pairwise correlations are low to medium, ranging from a minimum of 0.21 between EMN and DS to a maximum of 0.63 between EH and DS. The average correlation is 0.435. Note that all the correlation coefficients are statistically significant at 5% level of significance.

It has been noted in the literature (see e.g. [3]) that the characteristics of hedge fund strategies may be nonlinear and possibly affected by market episodes. Therefore, their sample characteristics are likely to change over time due to market events. In a multivariate setting, the correlation/covariance structure may also be time-varying since hedge fund managers may follow different dynamic investment strategies which, however, reflect the same changes in market conditions. Over the last few decades, several crises and market episodes occurred, which affected certain hedge fund returns. It has been found that two important crises that caused substantial changes in the risk–return characteristics of many single strategy hedge fund indices are the LTCM crisis in late 1998 and the Technology and Internet bubble in March 2000. The study of structural change in univariate hedge fund return series captured at least one structural break during each of the periods around the above crises (see [8]).

To investigate whether the sample characteristics of the analysed series, as well as their correlation structure, change due to these events, we compute the descriptive statistics and the correlation coefficients for the series over subperiods of the sample defined by the known dates when the crises occurred, September 1998 and March 2000. In Table 1 (Panels B–D), the summary statistics for the data in the three segments of the sample are shown. We observe differences in the mean and median returns among segments, some of which are substantial (see, for example, the

Table 2. Sample correlations of monthly returns for the four analysed HFR single strategy indices.

		(1)	(2)	(3)	(4)
		Panel A: January 1994–November 2005			
EH	(1)	1.00			
EMN	(2)	0.34*	1.00		
DS	(3)	0.63*	0.21*	1.00	
MA	(4)	0.58*	0.33*	0.52*	1.00
		Panel B: January 1994–September 1998			
EH	(1)	1.00			
EMN	(2)	0.51*	1.00		
DS	(3)	0.73*	0.53*	1.00	
MA	(4)	0.71*	0.49*	0.74*	1.00
		Panel C: November 1998–March 2000			
EH	(1)	1.00			
EMN	(2)	0.54*	1.00		
DS	(3)	0.63*	0.09	1.00	
MA	(4)	0.20	0.21	0.12	1.00
		Panel D: April 2000–November 2005			
EH	(1)	1.00			
EMN	(2)	0.09	1.00		
DS	(3)	0.66*	−0.01	1.00	
MA	(4)	0.54*	0.26*	0.39*	1.00

Notes: Panel A contains correlations based on period January 1994–November 2005, Panel B contains correlations based on period January 1994–September 1998, Panel C contains correlations based on period November 1998–March 2000, Panel D contains correlations based on period April 2000–November 2005.
*Indicates that the respective correlation coefficient is statistically significant at 5% level of significance.

differences in EH and MA). Significant differences can be also seen in the kurtosis of all series, which in most cases are much higher in the first segment and reduce in the last two segments. Finally, some interesting differences are found between the correlations in different subperiods (Table 2, Panels B–D). Obviously, all the correlations are quite high and statistically significant in the period from January 1994 to September 1998, while they reduce substantially, and most of them become insignificant, in the subperiod between the two crises. Most of the correlation coefficients become again significantly different from zero in the last segment after the Technology and Internet bubble in March 2000. These findings show that the differences in the correlation structure of the data are even more striking than the differences in the means, thus justifying the need for jointly modelling correlated series of returns using multivariate break-point models.

In this study, we are interested in detecting common structural breaks of unknown occurrence times in the four analysed series using a probabilistic approach to inference. Furthermore, it is of interest to identify a common set of explanatory risk factors that are most likely to affect the series of hedge fund returns. In particular, we use the Bayesian approach described in Section 3 to compare simultaneously various multivariate break-point models of different model specifications (i.e. various numbers of breaks and different subsets of factors). We consider all the possible subsets in the set of the seven ABS factors (i.e. $2^7 = 128$ factor model specifications) and 0–10 break-points. In the following analysis, we consider multivariate break-point risk factor models, which simultaneously identify relevant risk factors and detect changes in the risk exposures and the correlations/covariances of the data due to market events.

We report results obtained from the simultaneous analysis of the hedge fund single strategy indices. Table 3 presents the five most probable models. The most probable model, with posterior probability of 0.542, is the four-breaks model which includes the S&P and SCMLC factors. It can be seen that the four most probable models include only the two equity ABS factors. Also, this table provides evidence for at least three structural breaks having occurred at the multivariate series of returns. Robust inference about the number of structural breaks can be made using the respective marginal posterior distribution, that is, integrating out the uncertainty about the set of most relevant factors to be included in the model. The marginal distribution of the number of breaks, averaged over all possible factor model specifications, is given in Table 4. The highest posterior probability, achieved by the four-breaks model, is 0.566. The marginal posterior distribution of the different model specifications is extremely concentrated. The two equity factor specification ranks first according to this marginal distribution, with respective probability of 0.963. This clearly indicates that the four analysed series share common risk–return characteristics and, therefore, can be adequately explained by a small set of explanatory risk factors, specifically by the factors that belong to the equity class.

Table 3. Five most probable multivariate regression models obtained from the simultaneous comparison of all possible models with different numbers of breaks and subsets of factors.

Factors included in model	m	PMP
S&P, SCMLC	4	0.542
S&P, SCMLC	5	0.329
S&P, SCMLC	6	0.047
S&P, SCMLC	3	0.040
S&P, SCMLC, CredSpr	4	0.024

Notes: m denotes the number of breaks and PMP denotes the posterior model probability. S&P is the excess returns on the S&P 500 index, SCMLC is the spread between small-cap and large-cap stock returns and CredSpr is the change in the yield spread between 10-year T-bonds and Moody’s Baa bonds.

Table 4. Marginal posterior distribution of the number of breaks for the analysed hedge fund indices.

m	0	1	2	3	4	5	6	7	8	9	10
MPP	0.000	0.000	0.000	0.050	**0.566**	0.330	0.048	0.005	0.001	0.000	0.000

Note: *m* denotes the number of breaks and MPP denotes the marginal posterior probability.

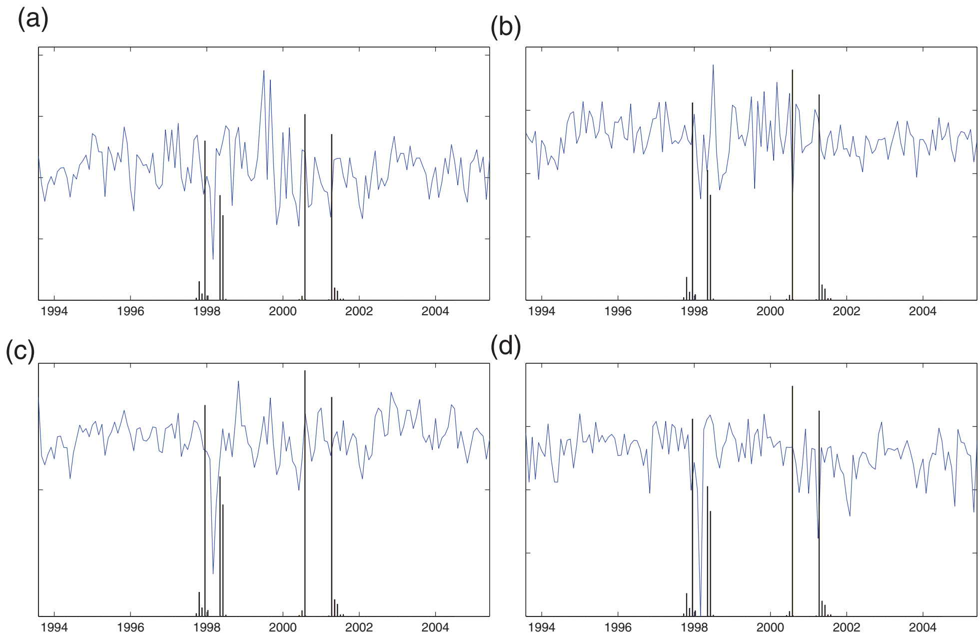

Figure 1. Plots of the hedge fund single strategy indices: (a) EH, (b) Emerging Markets Neutral, (c) DS and (d) MA. Also, the posterior distribution of the break-dates obtained under the most probable multivariate regression model, that is, the four-breaks model including the S&P and SCMLC factors, is shown.

In Figure 1, the plots of the four hedge fund return series are shown. Also, the histogram of draws from the posterior distribution of the positions of the structural breaks, obtained under the most probable model, that is, the four break-points model including the S&P and SCMLC factors, is shown in each plot. It can be seen that the posterior distributions of the break-dates are highly concentrated with posterior modes at June 1998, October 1998, January 2001 and September 2001. Obviously, the first two breaks have occurred in the period around the LTCM crisis. The third break in early 2001 can be associated with the Technology and Internet boom after which the variability of the returns and potentially the risk characteristics of the series have changed. Finally, the last break coincides with the time of the events in the USA. Looking at Figure 1, it can be seen that in the subperiod between the first two breaks most of the funds (EH, DS and MA) experienced great losses. Also, we may observe that in the period from October 1998 to January 2001 (the period between the LTCM crisis and the Technology and Internet bubble), the variability of the returns of EH, EMN and DS is larger than in the first and last segments of the sample. On the other hand, the variability of the MA returns increases after September 2001 (i.e. in the last segment). Independent univariate analyses have also captured structural breaks at dates close to the LTCM crisis, the Technology and Internet bubble and the events of September 2001 (see [8]). Two of the breaks were identified in the DS and EH series, three breaks in the MA series and four breaks in the EMN series. The multivariate analysis combines information from multiple return series and therefore is able to capture a larger number of structural breaks. Evidence for

Table 5. Estimated regression coefficients of the most probable multivariate break-point model, that is, the four-breaks model including the S&P and SCMLC factors, for the analysis of the four HFR indices.

	EH	EMN	MA	DS
	First segment: January 1994–June 1998			
α	0.0067 (0.0017)	0.0039 (0.0011)	0.0062 (0.0011)	0.0052 (0.0013)
$\beta_{S\&P}$	0.4288 (0.0488)	0.0607 (0.0348)	0.0922 (0.0295)	0.1982 (0.0359)
β_{SCMLC}	0.4928 (0.0707)	−0.0043 (0.0408)	0.1181 (0.0436)	0.1578 (0.0504)
	Second segment: July 1998–October 1998			
α	−0.0003 (0.0077)	−0.0016 (0.0052)	−0.0034 (0.0060)	−0.0251 (0.0143)
$\beta_{S\&P}$	0.4454 (0.0757)	0.0612 (0.0512)	0.3272 (0.0611)	0.3180 (0.1428)
β_{SCMLC}	0.1689 (0.2187)	0.1417 (0.1500)	0.0848 (0.1692)	0.0036 (0.4112)
	Third segment: November 1998–January 2001			
α	0.0100 (0.0031)	0.0031 (0.0025)	0.0077 (0.0015)	0.0014 (0.0030)
$\beta_{S\&P}$	0.6914 (0.0739)	0.0791 (0.0573)	−0.0081 (0.0374)	0.2359 (0.0537)
β_{SCMLC}	0.5406 (0.0677)	0.0941 (0.0512)	0.0427 (0.0287)	0.2031 (0.0553)
	Fourth segment: February 2001–September 2001			
α	−0.0031 (0.0032)	0.0026 (0.0026)	−0.0009 (0.0044)	0.0058 (0.0043)
$\beta_{S\&P}$	0.2985 (0.0573)	−0.1194 (0.0452)	0.0777 (0.0719)	0.0218 (0.0728)
β_{SCMLC}	0.1623 (0.0842)	0.0070 (0.0683)	0.1365 (0.1124)	0.1958 (0.1120)
	Fifth Segment: October 2001–November 2005			
α	0.0019 (0.0011)	0.0006 (0.0008)	0.0003 (0.0008)	0.0084 (0.0017)
$\beta_{S\&P}$	0.2910 (0.0272)	−0.0007 (0.0193)	0.1065 (0.0201)	0.1380 (0.0414)
β_{SCMLC}	0.3321 (0.0421)	0.0759 (0.0290)	0.1456 (0.0297)	0.1393 (0.0619)

Note: The posterior means and standard deviations (in parenthesis) are presented.

structural breaks in the multivariate model may also reflect changes in the correlation structure of the data.

It is of particular interest to investigate how the multivariate break-point model can reveal differences in the risk exposures of the analysed hedge fund single strategy indices, as well as changes in the covariance/correlation matrix. Given the sample from the posterior distribution of the break-dates, we have simulated draws from the posterior distributions of the model parameters α, $\beta_{S\&P}$ and β_{SCMLC}, for the four series and Σ, that is, the managers' skill (alphas), the risk exposures (betas) and the error covariance matrices in all segments. The posterior means and standard deviations of the regression coefficients are given in Table 5, while those of the variances, covariances and correlation coefficients are given in Table 6. Furthermore, plots of the posterior densities of the parameters in different segments of the sample are illustrated in Figures 2 and 3. We discuss the results in the first, third and last subperiods, since the other two segments are quite small for accurate inferences on the model parameters to be drawn from monthly observations. However, in the tables and figures, we present the full results for completeness.

It can be seen (Table 5 and Figure 2) that the differences in the risk-adjusted performance (alphas) are substantial for all series of returns. In particular, all the alphas in the first segment are positive and statistically significant.[4] In the third subperiod, the alphas of EH and MA increase, while the alphas of EMN and DS decrease and become insignificant. Finally, in the last segment, all the alphas are insignificant, except for the DS risk-adjusted performance that increases significantly. With respect to the risk exposures to the two equity factors, we observe that they also change in time. Note that for the EH, EMN and DS series, the equity factor loadings are positive with larger values in the third subperiod, from November 1998 to January 2001, when the above strategies experienced high return variability. This reflects the fact that over the period before the Technology and Internet bubble many individual hedge funds were heavily investing on high technology stocks, while they successfully avoided much of the downturn that followed the

Table 6. Estimates of the error variances, covariances and correlations for the most probable multivariate break-point regression model, that is, the four-breaks model including the S&P and SCMLC factors, for the analysis of the four HFR indices.

	EH	EMN	MA	DS
	First segment: January 1994–June 1998			
EH	*0.1131* (*0.0235*)	*0.0269* (*0.0113*)	*0.0157* (*0.0105*)	*0.0416* (*0.0139*)
EMN	0.3574 (0.1219)	*0.0497* (*0.0109*)	*0.0068* (*0.0068*)	*0.0160* (*0.0085*)
MA	0.2569 (0.1262)	0.1499 (0.1279)	*0.0420* (*0.0133*)	*0.0150* (*0.0081*)
DS	0.4831 (0.1062)	0.2797 (0.1246)	0.2851 (0.1212)	*0.0651* (*0.0154*)
	Second segment: July 1998–October 1998			
EH	*0.1595* (*0.1315*)	*0.0355* (*0.0565*)	*0.0634* (*0.0736*)	*0.1039* (*0.1641*)
EMN	0.3098 (0.3094)	*0.0765* (*0.0569*)	*0.0158* (*0.0419*)	*0.0424* (*0.0995*)
MA	0.4907 (0.2683)	0.1747 (0.3277)	*0.0961* (*0.0698*)	*0.0899* (*0.1175*)
DS	0.3294 (0.3127)	0.1951 (0.3271)	0.3713 (0.2944)	*0.5529* (*0.3816*)
	Third segment: November 1998–January 2001			
EH	*0.2366* (*0.0669*)	*−0.0049* (*0.0385*)	*−0.0085* (*0.0198*)	*0.0458* (*0.0360*)
EMN	−0.0229 (0.1870)	*0.1614* (*0.0461*)	*0.0202* (*0.0160*)	*−0.0861* (*0.0353*)
MA	−0.0874 (0.1880)	0.2510 (0.1673)	*0.0387* (*0.0126*)	*−0.0014* (*0.0166*)
DS	0.2572 (0.1708)	−0.5907 (0.1371)	−0.0238 (0.1834)	*0.1307* (*0.0466*)
	Fourth segment: February 2001–September 2001			
EH	*0.0626* (*0.0343*)	*−0.0082* (*0.0202*)	*0.0829* (*0.0336*)	*0.0121* (*0.0315*)
EMN	−0.0143 (0.2869)	*0.0467* (*0.0268*)	*0.0036* (*0.0276*)	*−0.0244* (*0.0281*)
MA	0.2470 (0.2744)	0.0407 (0.2841)	*0.1307* (*0.0689*)	*0.0157* (*0.0463*)
DS	0.1277 (0.2801)	−0.3053 (0.2626)	−0.1181 (0.2813)	*0.1276* (*0.0656*)
	Fifth Segment: October 2001–November 2005			
EH	*0.0560* (*0.0115*)	*0.0176* (*0.0062*)	*0.0229* (*0.0067*)	*0.0385* (*0.0134*)
EMN	0.4470 (0.1115)	*0.0274* (*0.0055*)	*0.0118* (*0.0043*)	*0.0184* (*0.0088*)
MA	0.5583 (0.0970)	0.4129 (0.1147)	*0.0297* (*0.0060*)	*0.0293* (*0.0097*)
DS	0.4545 (0.1089)	0.3106 (0.1239)	0.4756 (0.1065)	*0.1258* (*0.0258*)

Notes: The estimated posterior means and standard deviations (in parenthesis) are presented. The variances $\times 10^3$ (on the diagonal) and the covariances $\times 10^3$ are given in italics, while the correlations are given in normal type.

bubble. The risk–return characteristics of MA are slightly different; the exposures to the equity ABS factors are insignificant in the third segment while taking their largest values in the last period of the sample. Note that similar estimates of the regression coefficients are obtained by fitting independent univariate break-point risk factor models to these series (see [8]).

The joint modelling of the funds can be also useful for studying the changes in the pairwise correlations/covariances across time. Looking at Table 6 and Figure 3 we can see that there are differences in the estimated error variances in different segments (see the diagonal entries of the table and figure), but even more striking are the differences in the estimated correlations (entries below the diagonal). In the first subperiod of the sample most of the correlations are significant (apart from the correlation between EMN and MA) although their positive values are moderate, ranging from 0.2569 to 0.4831. On the other hand, in the third segment most of the correlations are insignificant (except for the negative correlation between EMN and DS), which suggests that during the volatile period that followed the LTCM crisis the analysed series are uncorrelated. Finally, the four series become substantially correlated again in the last period of the sample, after the events of September 2001 in the US. All of the correlations in this segment are significantly positive and take their largest values (they range from 0.3106 to 0.5583). This substantial change in the correlation structure justifies the presence of the last structural break in the multivariate analysis of the series, which is not apparent if we look at the plots of the series and can not be captured by the univariate break-point models for most series.

The above results reveal vital information about how the risk exposures and the correlations/covariances change in time. The differences in the parameters can certainly not

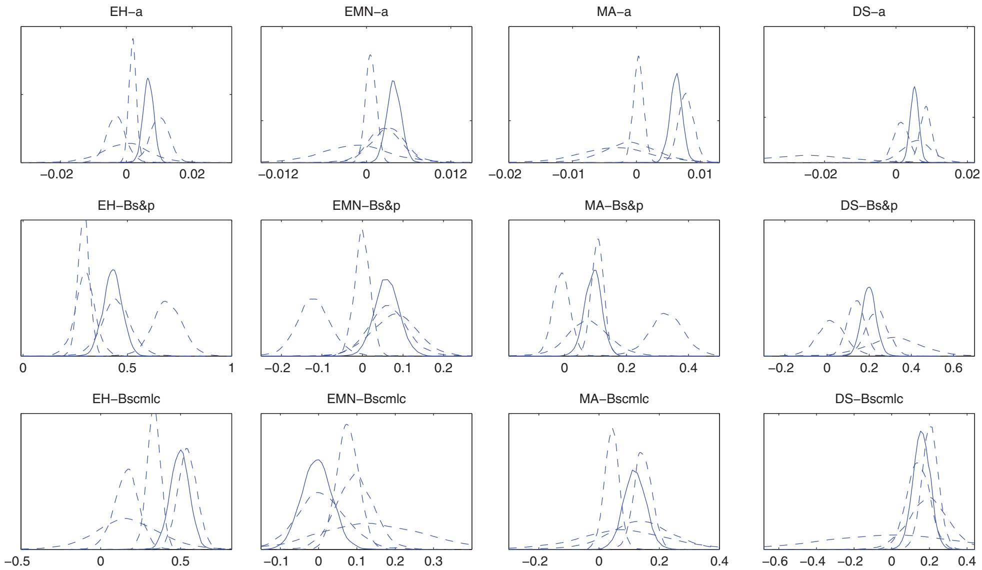

Figure 2. Posterior distribution of the parameters of the most probable multivariate break-point model, that is, the four-breaks model including the S&P and SCMLC factors. In each subplot, the posterior density of the respective parameters in the five different segments is illustrated.

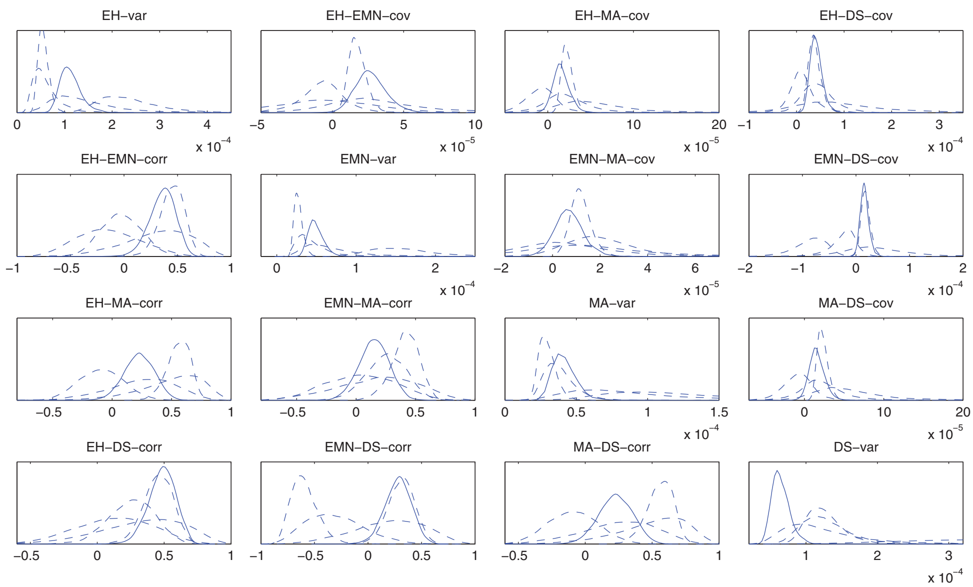

Figure 3. Posterior distribution of the error variances, covariances and correlations based on the most probable multivariate break-point model, that is, the four-breaks model including the S&P and SCMLC factors. In each subplot, the posterior density of the respective parameters in the five different segments is illustrated.

be captured by a static multivariate regression model. In Table 7, we present the parameter estimates of the most probable simple multivariate regression model. This model includes the S&P, SCMLC, CredSpr and PTFSBD factors. It is of particular interest to note that the estimated correlations in this model have quite low positive values; two of them are insignificant, while the values of the remaining four range from 0.2179 to 0.3104 (Table 7, Panel B). This

Table 7. Estimates of the regression coefficients (Panel A) and the variances, covariances and correlations (Panel B) for the most probable simple multivariate regression model which includes the factors S&P, SCMLC, CredSpr and PTFSBD.

	EH	EMN	MA	DS
		Panel A: Regression coefficients		
α	0.0052 (0.0011)	0.0029 (0.0007)	0.0036 (0.0007)	0.0053 (0.0009)
$\beta_{\text{S\&P}}$	0.4308 (0.0257)	0.0326 (0.0160)	0.1228 (0.0156)	0.1543 (0.0225)
β_{SCMLC}	0.4203 (0.0338)	0.0710 (0.0208)	0.1118 (0.0204)	0.1569 (0.0289)
β_{CredSpr}	0.5797 (0.9043)	1.8365 (0.5514)	0.1161 (0.5499)	−3.1035 (0.7608)
β_{PTFSBD}	0.0013 (0.0069)	−0.0060 (0.0042)	−0.0051 (0.0042)	−0.0277 (0.0059)
		Panel B: Variances, covariances and correlations		
EH	*0.1655 (0.0197)*	*0.0276 (0.0088)*	*0.0107 (0.0085)*	*0.0441 (0.0124)*
EMN	0.2694 (0.0764)	*0.0630 (0.0074)*	*0.0151 (0.0054)*	*0.0131 (0.0075)*
MA	0.1060 (0.0822)	0.2423 (0.0786)	*0.0610 (0.0071)*	*0.0188 (0.0073)*
DS	0.3104 (0.0744)	0.1488 (0.0812)	0.2179 (0.0784)	*0.1214 (0.0145)*

Notes: The posterior means and standard deviations (in parenthesis) are presented. In Panel B, the variances $\times 10^3$ (on the diagonal) and the covariances $\times 10^3$ are given in italics, while the correlations are given in normal type.

happens because the model that ignores the presence of structural breaks uses information from the whole sample to estimate the model parameters. Therefore, the resulting estimates of the correlations are rough averages of the corresponding correlation estimates in different subperiods of the sample. Similarly, the estimates of the regression coefficients in the simple multivariate regression model (Table 7, Panel A) are rough averages of the respective parameter estimates in different segments. Therefore, we may conclude that if the structural breaks are ignored, inferences on the risk exposures and the correlations/covariances will be inaccurate, since the nonlinearities in the multivariate return-generating process are not taken into account.

4.2. *Out-of-sample evaluation of the model*

In this section, we investigate potential impacts and benefits of our approach over approaches that ignore either the presence of the structural breaks in the series or the correlation structure of the data. We conduct an out-of-sample exercise with the aim to compare our multivariate regression break-point model with the simple multivariate regression as well as with four independent univariate break-point models. In order to evaluate these models, we use the logarithmic scoring rule which is based on the calculation of the conditional predictive ordinate under each model (see [8,38]). We calculate the predictive log score under each modelling approach. This is a measure of the predictive ability and consequently of the efficacy of a pricing model. Models which assign large predictive probability to the values that actually occur return a large log score.

In this evaluation exercise, we use the data from January 1994 to November 2003 for inference (estimation period) and consider the time interval from December 2003 to November 2005 as our 2-year out-of-sample period. We apply the methods of Section 3 to the estimation sample in order to obtain the most probable model, that is, the subset of factors $\boldsymbol{\gamma}$ and the number of breaks m. This model is used for the construction of 1 year of out-of-sample evaluation exercise. Then, the estimation period is redefined iteratively; it grows by one observation at each step. Let T' denote the initial estimation sample size. At the ith iteration, the data $\mathbf{R}_{1:T'+i-1}$, $\mathbf{F}_{1:T'+i-1}$, $i = 1, \ldots, 12$, are used for inference on the positions of the breaks and the model parameters, conditional on the most probable model specification. The return in the next period, $R_{T'+i}$, can be estimated using the estimates of the parameters in the last segment and the realized values of the factors at time

$T' + i, f_{T'+i}$. Furthermore, we can estimate the conditional predictive density at time $T' + i$ as

$$\hat{p}(R_{T'+i}|\mathbf{R}_{1:T'+i-1}, f_{T'+i}, m, \boldsymbol{\gamma}) = \frac{1}{B}\sum_{s=1}^{B} p(R_{T'+i}|\mathbf{R}_{1:T'+i-1}, f_{T'+i}, m, \boldsymbol{\gamma}, \mathbf{B}^{(s)}_{m+1,T'+i-1}, \Sigma^{(s)}_{m+1,T'+i-1}),$$

where $R_{T'+i}$ is the vector of realized returns at time $T' + i$, and $\mathbf{B}^{(s)}_{m+1,T'+i-1}$ and $\Sigma^{(s)}_{m+1,T'+i-1}$, $s = 1, \ldots, B$, are draws from the posterior distribution of the regression parameters and the covariance matrix in the last segment. The above procedure is repeated once again, after 1 year, to obtain the most probable model $(m, \boldsymbol{\gamma})$ and then estimate the returns $R_{T'+i}$ and the the conditional predictive densities $\hat{p}(R_{T'+i}|\mathbf{R}_{1:T'+i-1}, f_{T'+i}, m, \boldsymbol{\gamma})$, for $i = 13, \ldots, 24$. The same procedure is also used to calculate the estimates of the predictive densities under the most probable specifications of the simple multivariate regression model and, in a univariate setting, independently under the most probable univariate break-point risk factor model for each series. Then, the predictive log scores, $\sum_{i=1}^{24} \log \hat{p}(R_{T'+i}|\mathbf{r}_{1:T'+i-1}, f_{T'+i}, m, \boldsymbol{\gamma})$, can be used to measure the efficacy of the pricing models.

The model with the highest log predictive score should be preferred in terms of its predictive ability. In our example, the multivariate break-point model with predictive log score of 354.98 outperforms both the simple multivariate regression (predictive log score of 335.48) and the independent univariate break-point models (predictive log score of 340.39). The difference between the predictive log scores of two models, respectively, denoted by $\widehat{\text{pls}}_1$ and $\widehat{\text{pls}}_2$, can be interpreted on a 'per month' basis as follows. It represents an improvement in the predictive performance by a factor of $\exp\{(\widehat{\text{pls}}_1 - \widehat{\text{pls}}_2)/24\}$. In this example, the difference in log score between the multivariate break-point risk factor model and the simple multivariate regression model represents an improvement in the predictive performance by a factor of $\exp\{(354.98 - 335.48)/24\} = 2.25$ or by about 125%. Furthermore, the multivariate break-point risk factor model outperforms the independent univariate break-point models by 83.66%. Obviously, ignoring either the presence of structural breaks or the correlation structure of the data leads in substantial reductions in the efficacy of the model.

5. Conclusions

In this paper, we have developed a multivariate break-point risk factor model for the joint analysis of correlated series of hedge fund returns. We have extended the filtering method of Fearnhead [20] to the multivariate setting, in order to provide an exact approach to inference on the number and the positions of multiple common structural breaks in systems of regressions. Our method, which is based on a Bayesian model comparison approach, is able to detect the presence of structural breaks and also to identify the most relevant factors to explain the multivariate series of returns, given a set of available risk factors. The proposed approach is appealing for modelling multivariate series of hedge fund returns since it is able to capture the nonlinearities in the return-generating process. Nonlinearities, that is, changes in the risk exposures of the funds in certain market factors, as well as changes in the correlation structure of the data, are due to the highly dynamic trading strategies followed by fund managers which usually follow the changes in market conditions.

We have applied our approach to four correlated hedge fund single strategy indices, which share common investment characteristics and therefore are likely to be affected by the same risk factors and to experience common structural changes. In particular, we have used single strategy indices mainly characterized by equity related bets. Univariate analyses of these series have shown that there are some structural breaks which are common among the series (see [8]). In this study, we have found that several breaks occurred in the multivariate return series; we detected two structural breaks in 1998, that is, during and after the LTCM crisis, one break in January 2001, after the Technology and Internet boom, and finally one break in September 2001 when the events in the USA happened. The most relevant common explanatory risk factors identified by our Bayesian model comparison approach were the two equity ABS factors, namely the S&P 500 index and the spread between small-cap and large-cap stock returns.

The proposed nonlinear multivariate risk factor model has revealed vital information about how hedge fund risk exposures and correlations/covariances vary in time due to market events and episodes. We have found evidence that there are some substantial differences in the risk-adjusted performances, the risk exposures as well as the correlations/covariances of hedge funds among the different segments of our break-point model. Ignoring the presence of structural breaks is likely to result in misleading conclusions about the model parameters. Finally, from an out-of-sample exercise we have conducted, we have found that our multivariate break-point risk factor model has much better predictive ability compared to the standard multivariate regression model (which ignores the presence of structural breaks) and to independent univariate break-point analyses (which ignore the correlation structure of the data).

Acknowledgements

We thank Elias Tzavalis for many helpful discussions and comments.

Notes

1. Edwards [4], Brown *et al.* [5] and Brunnermeier and Nagel [6] described and studied hedge fund behaviour during well-known financial events, such as the collapse of long-term capital management, the Asian currency crisis and the Technology and Internet bubble, respectively.
2. Details can be found in HFR, www.hedgefundresearch.com.
3. See [34] for a detailed description of the construction of the three PTF factors. The seven ABS factors have been downloaded from David A. Hsieh's Data Library on Hedge Fund Risk Factors at http://faculty.fuqua.duke.edu/d̃ah7/HFRFData.htm
4. Certainly, significance tests do not match with the Bayesian approach to inference. Throughout this paper we use the term 'significant parameter' to refer to a parameter whose posterior distribution does not include zero in its highest posterior density region.

References

[1] G.S. Amin and H.M. Kat, *Hedge fund performance 1990–2000: Do the 'money machines' really add value?* J. Financ. Quant. Anal. 38 (2003), pp. 251–274.
[2] R. Kouwenberg and W.T. Ziemba, *Incentives and risk taking in hedge funds*, J. Bank. Finance 31 (2007), pp. 3291–3310.
[3] W. Fung and D.A. Hsieh, *Hedge fund benchmarks: A risk-based approach*, Financ. Analyst J. 60 (2004), pp. 65–80.
[4] F.R. Edwards, *Hedge funds and the collapse of long-term capital management*, J. Econ. Perspect. 13 (1999), pp. 189–210.
[5] S.J. Brown, W.N. Goetzmann, and J. Park, *Hedge funds and the Asian currency crisis of 1997*, J. Portfolio Manage. 26 (2000), pp. 95–101.
[6] K.M. Brunnermeier and S. Nagel, *Hedge funds and the technology bubble*, J. Finance LIX (2004), pp. 2013–2040.
[7] W. Fung, D.A. Hsieh, N.Y. Naik, and T. Ramadorai, *Hedge funds: Performance, risk and capital formation*, J. Finance 8 (2008), pp. 1777–1803.
[8] L. Meligkotsidou and I.D. Vrontos, *Detecting structural breaks and identifying risk factors in hedge fund returns: A Bayesian approach*, J. Bank. Finance 32 (2008), pp. 2471–2481.
[9] D.W.K. Andrews, *Tests for parameter instability and structural change with unknown change point*, Econometrica 61 (1993), pp. 821–856.
[10] D.W.K. Andrews and W. Ploberger, *Optimal tests when a nuisance parameter is present only under the alternative*, Econometrica 62 (1994), pp. 1383–1414.
[11] R.L. Brown, J. Durbin, and J.M. Evans, *Techniques for testing the constancy of regression relationships over time*, J. R. Statist. Soc. Ser. B 37 (1975), pp. 149–163.
[12] W. Ploberger and W. Kramer, *The CUSUM test with OLS residuals*, Econometrica 60(2) (1992), pp. 271–285.

[13] C.-S.J. Chu, K. Hornik, and C.-M. Kuan, *MOSUM tests for parameter constancy*, Biometrika 82(3) (1995), pp. 603–617.
[14] J. Bai and P. Perron, *Estimating and testing linear models with multiple structural changes*, Econometrica 66 (1998), pp. 47–78.
[15] J. Bai and P. Perron, *Computation and analysis of multiple structural change models*, J. Appl. Econometrics 18 (2003), pp. 1–22.
[16] P. Perron and Z. Qu, *Estimating restricted structural change models*, J. Econometrics 134 (2006), pp. 373–399.
[17] J. Bai, R.L. Lumsdaine, and J.H. Stock, *Testing for and dating common breaks in multivariate time series*, Rev. Econom. Stud. 65 (1998), pp. 395–432.
[18] J. Bai, *Vector-autoregressive models with structural changes in regression coefficients and variance–covariance matrices*, Ann. Econ. Finance 1 (2000), pp. 303–339.
[19] Z. Qu and P. Perron, *Estimating and testing structural changes in multivariate regressions*, Econometrica 75 (2007), pp. 459–502.
[20] P. Fearnhead, *Exact and efficient Bayesian inference for multiple changepoints problems*, Stat. Comput. 16 (2006), pp. 203–213.
[21] S.L. Scott, *Bayesian methods for hidden Markov models: Recursive computing in the 21st century*, J. Am. Statist. Assoc. 97 (2002), pp. 337–351.
[22] P. Fearnhead and L. Meligkotsidou, *Exact filtering for partially-observed continuous-time models*, J. R. Statist. Soc. Ser. B 66 (2004), pp. 771–789.
[23] P. Fearnhead, *Exact filtering for partially-observed queues*, Stat. Comput. 14 (2004), pp. 261–266.
[24] P. Fearnhead, *Exact Bayesian curve fitting and signal segmentation*, IEEE Trans. Signal Process. 53 (2005), pp. 2160–2166.
[25] E.I. George, D. Sun, and S. Ni, *Bayesian stochastic search for VAR model restrictions*, J. Econometrics 142 (2008), pp. 553–580.
[26] H. Wang, *Sparse seemingly unrelated regression modelling: Applications in finance and econometrics*, Comput. Stat. Data Anal. 54 (2010), pp. 2866–2877.
[27] J.M. Bernardo and A.F.M. Smith, *Bayesian Theory*, Wiley Series in Applied Probability and Statistics, John Wiley & Sons, Chichester, 1994.
[28] A. Zellner, *On assessing prior distributions and Bayesian regression analysis with g-prior distributions*, in *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti*, P.K. Goel and A. Zellner, eds., North-Holland, Amsterdam, 1986, pp. 233–243.
[29] P.J. Brown, M. Vannucci, and T. Fearn, *Multivariate Bayesian variable selection and prediction*, J. R. Statist. Soc. Ser. B 60 (1998), pp. 627–641.
[30] J. Carpenter, P. Clifford, and P. Fearnhead, *Improved particle filter for nonlinear problems*, IEEE Proc. Radar Sonar Navig. 146 (1999), pp. 2–7.
[31] N. Amenc and L. Martellini, *Portfolio optimization and hedge fund style allocation decisions*, J. Altern. Investments 5 (2002), pp. 7–20.
[32] V. Agarwal and N.Y. Naik, *Risks and portfolio decisions involving hedge funds*, Rev. Financ. Stud. 17(1) (2004), pp. 63–98.
[33] D. Giamouridis and I.D. Vrontos, *Hedge fund portfolio construction: A comparison of static and dynamic approaches*, J. Bank. Finance 31(1) (2007), pp. 199–217.
[34] W. Fung and D.A. Hsieh, *The risk in hedge fund strategies: Theory and evidence from trend followers*, Rev. Financ. Stud. 14 (2001), pp. 313–341.
[35] W. Fung and D.A. Hsieh, *The risk in fixed-income hedge fund styles*, J. Fixed Income 12 (2002), pp. 6–27.
[36] W. Fung and D.A. Hsieh, *The risk in hedge fund strategies: Theory and evidence from long/short equity hedge funds*, J. Empir. Finance 18 (2011), pp. 547–569.
[37] R. Kosowski, N.Y. Naik, and M. Teo, *Do hedge funds deliver alpha? A Bayesian and bootstrap analysis*, J. Financ. Econ. 84 (2007), pp. 229–264.
[38] S.D. Vrontos, I.D. Vrontos, and D. Giamouridis, *Hedge fund pricing and model uncertainty*, J. Bank. Finance 32 (2008), pp. 741–753.