Mining financial time series: New insights from model-based clustering methods

José G. Dias*

Department of Quantitative Methods & UNIDE, ISCTE – Lisbon University Institute, Portugal

Jeroen K. Vermunt

Department of Methodology and Statistics, Tilburg University, The Netherlands

Sofia Ramos

Department of Finance & UNIDE, ISCTE – Lisbon University Institute, Portugal

Abstract

In recent years large amounts of financial data have become available for analysis. We propose to explore returns from 21 European stock markets by model-based clustering of regime switching models. These models allow the relaxation of traditional assumptions such as conditional Gaussian returns. The data mining approach handles simultaneously the heterogeneity across stock markets and time, i.e., time-constant and time-varying discrete latent variables capture unobserved heterogeneity between and within stock markets, respectively. The results show a clear distinction between groups of stock markets, each one characterized by different regime switching dynamics that correspond to different expected return-risk patterns.

Keywords: Data mining, hidden Markov model, stock indexes, cluster analysis, model-based clustering, latent class model

*Corresponding author. Tel.: 351 217 903 208 Fax.: 351 217 903 004

Email addresses: `jose.dias@iscte.pt` (José G. Dias), `J.K.Vermunt@uvt.nl` (Jeroen K. Vermunt), `sofia.ramos@iscte.pt` (Sofia Ramos)

1. Introduction

In recent years dealing with unobserved heterogeneity has become a predominant topic in many research areas. As Heckman emphasized in his Nobel lecture, one of the most important discoveries in microeconometrics is the pervasiveness of heterogeneity and diversity in economic life: "When a full analysis of heterogeneity in response was made, a variety of candidate averages emerged to describe the average person, and the long-standing edifice of the representative consumer was shown to lack empirical support. (Heckman, 2001, p. 674)". In finance research, heterogeneity has been mostly assumed observed (e.g., based on countries), where groups and their boundaries are delineated without regarding the intrinsic information on the observed data. However, there are plenty of examples in the academic and professional finance literature that show that heterogeneity exists among capital market participants, business managers, fund managers, among others.

Modeling the dynamics of stock market returns has been an important challenge in modern financial research. The statistics and dynamics of correctly specified distributions provide more accurate and detailed input for financial asset pricing and risk management. Despite of the fact that the dominant approach followed by both academics and practitioners has been to assume that returns follow a normal distribution, skewness and excessive kurtosis have long been recognized for returns of stock market indexes or for returns of many financial assets. A common conclusion is that the normal distribution is inadequate for short period returns of financial assets (Mandelbrot, 1963; Fama, 1965; Praetz, 1972). Several alternative distributions have therefore been suggested for modeling the returns, one of which is the Laplace distribution. These alternatives have in common that they try to accommodate for the excessive kurtosis in the empirical distribution of the returns. Whereas excess kurtosis of financial return distributions has been well addressed in the financial literature, the asymmetry of the distribution has not received much attention, and the few studies available tend to be inconclusive (Simkowitz and Beedles, 1980; Singleton and Wingender, 1986; Peiró, 1999).

Latent class or finite mixture modeling has proven to be a powerful tool for analyzing a wide range of social and behavioral science data (see, for example, Clogg (1995) and Vermunt (2003)). We propose a latent class model for financial data mining that takes into account unobserved heterogeneity by means of time-constant and time-varying discrete latent variables. A feature of latent class modeling that is especially attractive to the typical analysis in finance research is that it yields a model-based clustering of observational units. In finance, one typically aims to identify subpopulations of firms, investors, markets or countries that differ in their propensities to specific characteristics (regulation, company governance characteristics, etc.). Here, this methodology is used to model the dynamics of the returns of 21 European stock market indexes. Stock markets are well-known for presenting cycles, however country idiosyncrasies are also likely to make them differ in their transition among boom and peaks. As is illustrated below, the proposed approach is flexible in the sense that it can deal with the specific features of financial time series data, such as asymmetry, kurtosis, and unobserved heterogeneity, an aspect that tends to be ignored. Because we selected a rather large and heterogeneous sample of countries including both developed and emerging countries and both EMU and non EMU countries, we expect that heterogeneity in market returns due to country idiosyncrasies will show up in the results. For instance, emerging market return distributions show larger deviations from normality; i.e., are more skewed and have fat tails (Harvey, 1995).

Section 2 describes the 21 country financial time series data set that is used throughout this paper. Section 3 presents the statistical framework for the data mining of heterogeneous financial time series. It also discusses parameter estimation by maximum likelihood and model selection issues. Section 4 reports the results obtained for the data set at hand. The paper concludes with a summary of the main findings and a description of possible implications.

2. Description of the data set

The data used in this article are daily closing prices from 27 January 1998 to 8 June 2010 for 21 European stock market indexes drawn from Datastream database.[1] The series are denominated in US dollars. In total, we have 3193

[1]Observations from different time zones can create problems of non-synchronization on the analysis, to eliminate such problems we focus on European markets.

end-of-the-day observations per country. Let P_{it} be the observed daily closing price of market i on day t, $i = 1, ...n$ and $t = 0, ..., T$. The daily rates of return are defined as the percentage rate of return by $y_{it} = 100 \times \log(P_{it}/P_{i,t-1})$, $t = 1, ..., T$, with $T = 3193$. This definition which is commonly used in the literature is justified by the fact that for expected small increases (decreases) of value, say r, $\log(1 + r) \simeq r$.

[Figure 1 about here.]

[Figure 2 about here.]

The 21 stock markets are listed in Table 1. Figures 1 and 2 depict the index and returns time series, respectively. As is well known, stock markets present cycles. In the sample period there were two main periods of global stock market crises. The dot-com bubble collapse that started at the end of 1999 and went on until 2003, and the subprime crisis that had its first signs in the summer of 2007, but made financial markets plummet in 2008 after the Lehman Brothers bankruptcy. Between 2004 and 2007, stock markets registered a strong growth.

Figure 2 depicts the returns of the stock markets. Russia and Turkey showed the high level of instability typical for emerging markets. It is worth to note that in August 1998 Russia defaulted a sovereign bond payment triggering the "ruble crisis" in financial markets. Market features like these seem well-suitable to test our data mining model.

[Table 1 about here.]

Table 1 provides the relevant descriptive statistics for the 21 stock-return series. All markets showed non-negative median returns. However, only 16 out of 21 had positive mean returns. Greece, Ireland, Italy, the Netherlands, Portugal and the United Kingdom provided negative mean returns. Emerging markets such as Czech Republic and Russia showed larger positive mean return rates. This shows that indeed stock market distributions tend to be negatively skewed.

The 21 analyzed markets show very diverse patterns of dispersion, where the largest standard deviations are found for Russia (3.408) and Turkey (3.222) – both emerging markets – which are almost three times as large as for Switzerland, the stock market with the smallest dispersion with a standard deviation of 1.225. Moreover, the excess kurtosis (which equals 0 for normal distributions) shows values above 0, indicating heavier tails and more peakness than the normal. The Jarque-Bera test rejects the null hypothesis of normality for each of the 21 stock markets.

[Figure 3 about here.]

[Figure 4 about here.]

Figures 3 and 4 depict rolling means and standard deviations (30-day window) for these markets. Although moving averages tend to smooth trends, we still see the booms and peaks of stock market returns and volatility. The ruble crisis is visible mainly in Russia, but contaminates neighbor markets such as Poland and Hungary. All stock markets show a volatility peak during the subprime crisis. Other markets whose distinct booms and slumps in volatility came out in the figures are Greece, Finland, Turkey and Sweden.

3. The data mining model

3.1. Model definition

This section introduces a flexible modeling strategy for mining financial time series. Contrary to alternative heuristic clustering techniques introduced for financial time series analysis that operates directly on the correlation between time series (e.g., Basalto et al., 2007), the proposed model is a model-based clustering technique that accommodates for serial dependencies and unobserved heterogeneity.

The proposed model for mining financial time series – the mixture Gaussian hidden Markov model (MGHMM) – contains three types of variables: a time-varying response variable, a time-constant discrete latent variable, and a time-varying discrete latent variable. Let y_{it} represent the metric response of observation i at time point t, where $i \in 1, \ldots, n$, $t \in 1, \ldots, T$. The time-constant and time-varying discrete latent variables are denoted by w and z_t, respectively, where $w \in 1, \ldots, S$ and $z_t \in 1, \ldots, K$. The latter implies that the number of categories of the two types of latent variables equal S and K, respectively. To make as clear as possible distinction between the two types of latent variables, we will refer to w as a latent class and to z_t as a latent state or regime. The time-constant latent classes (w) can be seen as clusters for which the process under study differs. The time-varying latent variable with Markovian transition structure (z_t) is used to flexibly model the distribution of the time-specific responses as well as to capture changes that occur across adjacent time points. Figure 5 provides the graph of the data mining model.

[Figure 5 about here.]

The model is defined as:

$$f(\mathbf{y}_i; \boldsymbol{\varphi}) = \sum_{w=1}^{S} \sum_{z_1=1}^{K} \sum_{z_2=1}^{K} \cdots \sum_{z_T=1}^{K} f(w, z_1, \ldots, z_T) f(\mathbf{y}_i | w, z_1, \ldots, z_T) \quad (1)$$

with

$$f(w, z_1, \ldots, z_T) = f(w) f(z_1 | w) \prod_{t=2}^{T} f(z_t | z_{t-1}, w) \quad (2)$$

and

$$f(\mathbf{y}_i | w, z_1, \ldots, z_T) = \prod_{t=1}^{T} f(y_{it} | z_t). \quad (3)$$

Equation 1 describes $f(\mathbf{y}_i; \boldsymbol{\varphi})$, the (probability) density function associated with the time series of index return rates for stock market i. The right-hand side of this equation shows that we are dealing with a mixture model containing one time-constant latent variable and T time-varying latent variables. The total number of mixture components (or latent classes) equals SK^T, which is the product of the number of categories of w and z_t for $t = 1, 2, ..., T$. As in any mixture model, $f(\mathbf{y}_i; \boldsymbol{\varphi})$ results from the marginalization over the latent variables that being discrete is the weighted summation class-specific probability densities – here $f(\mathbf{y}_i|w, z_1, \ldots, z_T)$ – where the (prior) class membership probabilities or mixture proportions – here $f(w, z_1, \ldots, z_T)$ – serve as weights (McLachlan and Peel, 2000). The model described in Equations 1 – 3 is strongly related to the mixed latent Markov model proposed in the social sciences literature (van de Pol and Langeheine, 1990; Vermunt et al., 1999). Differences are that their model was for categorical instead of continuous responses and for (short) panel data rather than long time series like ours.

Equations 2 and 3 and Figure 5 show the conditional independence assumption implied by the MGHMM that simplify the form of the mixture proportion $f(w, z_1, \ldots, z_T)$ and the class-specific densities $f(\mathbf{y}_i|w, z_1, \ldots, z_T)$. More specifically, the equation for $f(w, z_1, \ldots, z_T)$ shows that within latent classes w, z_t is associated only with z_{t-1} and z_{t+1} and thus not with the latent states occupied at the other time points – the well-known first-order Markov assumption. The equation for $f(\mathbf{y}_i|w, z_1, ..., z_T)$ shows that conditionally on z_t, the response at occasion t (y_{it}) is independent of responses at other time points – usually referred to as the local independence assumption – and independent of the latent classes and the latent states at the other time points.

Two remarks should be made about the first-order Markov assumption for the latent states. First, after marginalizing over w, the process for the sequence z_t is no longer Markovian. Second, the Markov assumption for z_t conditionally on w does not imply a first-order Markov structure for the responses y_{it}. This shows that the first-order Markov assumption is not as restrictive as one may initially think (see Figure 5).

The three key elements of the data mining model described in Equations 1–3 relevant for the mining of heterogeneous time series are that it can take into account: 1) time-invariant unobserved heterogeneity in the process under study, 2) autocorrelation, and 3) flexible distributions that deviate in terms of skewness and kurtosis from normality (see, e.g., Dias and Wedel (2004)). Unobserved heterogeneity is captured by the time-constant latent variable w, autocorrelations are captured by the first-order Markov transition process in which the state at time point t may depend on the state at time point $t-1$, and flexible distributions of the returns are possible because of the time-specific mixture distribution for the response variable.

As can be seen from Equations 2 and 3, the model of interest is characterized by four sets of probability functions:

- $f(w)$ is the probability of belonging to a particular latent class w and $\pi_w = P(W = w)$;

- $f(z_1|w)$ is an initial-state probability; that is, the probability of having a particular latent initial state conditional on belonging to latent class w: $\lambda_{kw} = P(Z_1 = k|W = w)$;

- $f(z_t|z_{t-1}, w)$ is a latent transition probability; that is, the probability of being in a particular latent state at time point t conditional on the latent state at time point $t-1$ and class membership; assuming a timehomogeneous transition process, we have $a_{jkw} = P(Z_t = k|Z_{t-1} = j, W = w)$;

- $f(y_{it}|z_t)$ is the Gaussian density function for the observed response, which is the probability density of having a particular observed stock return in index i at time point t conditional on the latent state occupied at time point t. This distribution is characterized by the vector $\boldsymbol{\theta}_k = (\mu_k, \sigma_k^2)$ containing the means (μ_k) and variances (σ_k^2) for latent state k (and invariant across latent classes). Since the marginal distribution is a mixture of densities it defines a flexible model that takes into account skewness and kurtosis.

The $K(SK+2)-1$ free parameters of the MGHMM ($\boldsymbol{\varphi}$) include the $S-1$ class sizes, the $S(K-1)$ initial-state and $SK(K-1)$ transition probabilities and the $2K$ conditional means and variances of the observed variables.

3.2. Restricted special cases of the model

Various special cases of the MGHMM defined in Equations 1–3 can be obtained by eliminating one or more of its three main elements, the autocorrelation structure, the time-varying latent variables, or the time-constant latent variable. For example, if we assume that there is no autocorrelation between the time-varying discrete latent variables - that $p(z_t|z_{t-1}, w) = p(z_t|w)$ – we obtain a model that is called a multilevel or hierarchical mixture model (Vermunt, 2003, 2007). This shows that the MGHMM can be seen as a hierarchical Gaussian mixture model that is expanded with an autocorrelation structure.

The hidden Markov or Markov switching model (Baum et al., 1970; Hamilton, 1989) is the special case of the MGHMM that is obtained by eliminating the time-constant latent variable w from the model, that is, by assuming that there is no unobserved heterogeneity at the upper level of analysis. This model can be obtained without modifying the formulae, but by simply assuming that $S = 1$; that is, that all stock markets belong to the same group.

The mixture Gaussian model can be seen as a restricted variant of the MGHMM that is obtained by removing the time-varying latent variables z_t, resulting in $f(\mathbf{y}_i; \boldsymbol{\varphi}) = \sum_{w=1}^{S} f(w) \prod_{t=1}^{T} f(y_{it}|w)$. Note that this model is equivalent to a latent class model for T response variables which are assumed to be conditionally independent within latent classes w. For $T = 1$ it yields the mixture of (univariate) Gaussian distributions (Dias and Wedel, 2004).

3.3. Parameter estimation by maximum likelihood

Maximum likelihood (ML) estimation of the parameters of the MGHMM involves maximizing the log-likelihood function: $\ell(\boldsymbol{\varphi}; \mathbf{y}) = \sum_{i=1}^{n} \log f(\mathbf{y}_i; \boldsymbol{\varphi})$, a problem that can be solved by means of the Expectation-Maximization (EM) algorithm (Dempster et al., 1977).

In the E step, we compute $f(w, z_1, \ldots, z_T|\mathbf{y}_i) = f(w, z_1, \ldots, z_T, \mathbf{y}_i)/f(\mathbf{y}_i)$, which is the joint conditional distribution of the $T + 1$ latent variables given the data and the current provisional estimates of the model parameters. In the M step, standard complete data ML methods are used to update the unknown model parameters using an expanded data matrix with $f(w, z_1, \ldots, z_T|\mathbf{y}_i)$ as weights. Since the EM algorithm needs to compute and store the $S{\cdot}K^T$ entries of $f(w, z_1, \ldots, z_T|\mathbf{y}_i)$ for each subject or for each unique data pattern, computation time and computer storage increases exponentially with the number of time points, which makes this algorithm impractical or even impossible to apply with more than a few time points. This explains why models like the ones proposed here have been used in social science applications only with very short time series (van de Pol and Langeheine, 1990; Vermunt et al., 1999) and has hampered its application for longer time series (Schmittmann et al., 2006).

However, for hidden Markov models, a special variant of the EM algorithm has been proposed that is usually referred to as the forward-backward or Baum-Welch algorithm (Baum et al., 1970; Hamilton, 1989). This special algorithm is needed because the model for our data set contains a huge number of entries in the joint posterior latent distribution $f(w, z_1, \ldots, z_T|\mathbf{y}_i)$. Recall that in our application $T = 3193$. This means that even for $S = 2$ and $K = 2$, the number of entries in the joint posterior distribution is too large to process and store for all n stock markets as has to be done within a standard EM algorithm. The Baum-Welch algorithm circumvents the computation of this joint posterior distribution making use of the conditional independencies implied by the model. Whereas this algorithm was originally proposed for hidden Markov models, here we expand it to deal with mixture (Gaussian) hidden Markov models.

Because of the conditional independence assumption implied by our data mining model, in the M step of the EM algorithm one needs only the distributions $f(w|\mathbf{y}_i)$, $f(w, z_t|\mathbf{y}_i)$, and $f(w, z_{t-1}, z_t|\mathbf{y}_i)$. The Baum-Welch algorithm obtains these quantities directly rather than first computing $f(w, z_1, \ldots, z_T|\mathbf{y}_i)$ and subsequently collapsing over the remaining dimensions as would be done in a standard EM algorithm (Baum et al., 1970). The two key components of the Baum-Welch algorithm are the forward probabilities α_{iwz_t} and the backward probabilities β_{iwz_t}. Because of our generalization to the mixture case, we need one additional quantity γ_{iw}. These three quantities are defined as follows:

$$\alpha_{iwz_t} = f(z_t, y_{i1}, \ldots, y_{it}|w), \tag{4}$$

$$\beta_{iwz_t} = f(y_{i,t+1}, \ldots, y_{iT}|z_t, w), \tag{5}$$

$$\gamma_{iw} = f(w, \mathbf{y}_i). \tag{6}$$

Using α_{iwz_t}, β_{iwz_t}, and γ_{iw}, one can obtain the relevant marginal posteriors as follows:

$$f(w|\mathbf{y}_i) = \frac{\gamma_{iw}}{f(\mathbf{y}_i)} \tag{7}$$

$$f(w, z_t|\mathbf{y}_i) = \frac{\alpha_{iwz_t}\beta_{iwz_t}}{f(\mathbf{y}_i)} \tag{8}$$

$$f(w, z_{t-1}, z_t|\mathbf{y}_i) = \frac{\gamma_{iw}\alpha_{iwz_{t-1}} f(z_t|z_{t-1}, w) f(y_{it}|z_t, w)\beta_{iwz_t}}{f(\mathbf{y}_i)} \tag{9}$$

where $f(\mathbf{y}_i) = \sum_{w=1}^{S} \gamma_{iw}$, and $f(z_t|z_{t-1}, w)$ and $f(y_{it}|z_t, w)$ are model probabilities. The key element of the forward-backward algorithm is that the α_{iwz_t} and β_{iwz_t} are computed using recursive schemes. The forward recursion for α_{iwz_t} is:

$$\alpha_{iwz_1} = f(z_1|w) f(y_{i1}|z_1, w), \tag{10}$$

$$\alpha_{iwz_t} = \left[\sum_{z_{t-1}=1}^{K} \alpha_{iwz_{t-1}} f(z_t|z_{t-1}, w) \right] f(y_{it}|z_t, w), \tag{11}$$

for $t = 2$ up to $t = T$. The backward recursion for β_{iwz_t} is:

$$\beta_{iwz_T} = 1, \tag{12}$$

$$\beta_{iwz_t} = \sum_{z_{t+1}=1}^{K} \beta_{iwz_{t+1}} f(z_{t+1}|z_t, w) f(y_{i,t+1}|z_{t+1}, w), \tag{13}$$

for $t = T - 1$ down to $t = 1$. The quantity γ_{iw} is obtained as:

$$\gamma_{iw} = \sum_{z_t=1}^{K} f(w) \alpha_{iwz_t} \beta_{iwz_t}, \tag{14}$$

for any t. So, first we obtain α_{iwz_t} and β_{iwz_t} for each time point and subsequently we obtain γ_{iw}. Next, we compute $f(w|\mathbf{y}_i)$, $f(w, z_t|\mathbf{y}_i)$, and $f(w, z_{t-1}, z_t|\mathbf{y}_i)$ using Equations 7–9. In the M step, these quantities are used to obtain new estimates for the model parameters appearing in Equations 2 and 3 using standard methods.

Similar recursive schemes have been proposed for obtaining the gradient vector and the observed information matrix (Lystig and Hughes, 2002). These can be used to maximize the log-likelihood using the Newton-Raphson algorithm.

3.4. Decision on the number of latent classes and latent states

An important issue that remains to be addressed is how to estimate/select S and K, *i.e.*, how to decide about the number of latent classes and states needed. The standard model selection approach when using maximum likelihood estimation is by performing likelihood ratio tests across nested models; here the relevant tests are between models with $S-1$ and S classes and between models with $K-1$ and K states. However, in the context of latent class models this approach is problematic because the null hypothesis under test is defined on the boundary of the parameter space, and consequently the regularity condition of Cramer on the asymptotic properties of the MLE is no longer valid. As an alternative, there has been a recent interest in assessing latent class models fit via information statistics. The basic principle under these criteria is parsimony, which results from the trade-off between model fit and model complexity. A number of model selection criteria have been suggested, the most prominent and widely used being the Bayesian Information Criterion (BIC) of Schwarz (Schwarz, 1978) and the Akaike Information Criterion (AIC) of Akaike (Akaike, 1974). Because simulation studies have shown that in mixture modeling AIC tends to overfit (see, for example, Dias and Vermunt (2007)), the selection is based on the BIC. In our application we fix K *a priori* to $K = 2$, yielding the bull and bear market states commonly assumed in financial econometrics. For instance, Pagan and Sossounov (2003) observe that "bull and bear markets are a common way of describing cycles in equity prices (p. 23)." Therefore, like Ang and Bekaert (2002) and Wilfling (2009), we assume two distinct regimes $z_t \in \{1, 2\}$ for all $t = 1, ..., T$. Changes between the two regimes over time are assumed to follow a first-order Markov process. The value of S to be selected is the one that minimizes the value of BIC, which is defined as

$$BIC_S = -2\ell_S(\hat{\boldsymbol{\varphi}}; \mathbf{y}) + N_S \log n, \tag{15}$$

where N_S is the number of free parameters of the model with S latent classes.

4. Results

This section reports the results obtained when applying the data mining model described in the previous section to the stock return data set at hand. We estimated models with S ranging from one to five, where 500 different set of random starting values were used for the parameters to avoid local maxima (Table 2). A solution with four latent classes ($S = 4$) yielded the lowest BIC value (log-likelihood = -121657.84; number of free parameters = 19, and BIC = 243373.53). This model will therefore be treated as the final model in our analysis.

[Table 2 about here.]

[Table 3 about here.]

We start by characterizing the latent states or regimes. Table 3 provides information on the two latent states; that is, the average proportion of markets in state k over time and the mean and variance of the return in each state. The reported means show that the first state has negative returns and the second positive. In the financial jargon such states are also known by bear and bull markets, respectively. The overall probability of being in regime 1 and 2 equals 0.27 and 0.73, respectively.

The values of volatility are consistent with the stylized fact in financial markets: the asymmetry of the volatility. Bear regimes are associated with larger volatility than bull regimes. Thus, in periods of financial crisis, market prices fall down and have higher volatility. With financial market integration, such events tend to spillover to other markets, affecting returns of other stock markets, the so called *financial contagion*.[2]

[Table 4 about here.]

Table 4 reports the probability of being in one of the two states and the transition probabilities between the two states for each of the four latent classes. It can be observed that the four latent classes are ordered in increasing bearish state probability; that is, Class 1 countries are in the bear regime at only 16.2% of the time points, whereas this number increases to

[2]Some examples are the *Asian Flu Crisis* that affected Southeastern Asian countries and the *Russian Crisis* of 1998.

72.6% for Class 4. The diagonal "transition" probabilities are close to one indicating that there is a tendency of state persistence. This is a well-known phenomenon in financial markets; periods of high volatility are followed by high volatility, and periods of low volatility are followed by low volatility. Models like ARCH and GARCH aim at describing these features. The off-diagonal probabilities show that Class 4 has a much larger probability of switching from the bull to the bear regime (0.187) than countries in Class 1 (0.008).

The sojourn time is the expected number of days that a stock market stays in a given regime. For regime k and latent class w it can be obtained by $1/(1 - p_{kkw})$. As reported in Table 4, Class 4 stays the smallest number of days in both bear and bull markets, Class 1 stays the largest number of days in bull markets, and Class 3 stays the largest number of days in bear markets.

Table 5 summarizes the results related to the distribution of stock market across latent classes. The estimated prior class membership probability is largest for Class 1 (0.61), indicating that the probability that a randomly selected stock market belongs to Class 1 is 0.61. These probabilities are smaller for Class 2 (0.23), Class 3 (0.10), and Class 4 (0.06).

[Table 5 about here.]

A more detailed interpretation of the nature of the stock market latent variable is obtained by investigating the posterior class membership probabilities, conditional on the observed data (Table 5). As can be seen, 13 countries are assigned to Class 1, five to Class 2, two to Class 3, and just one to Class 4. Except for Czech Republic, the class assignment probability is always very close to one, indicating that the classes are almost fully non-overlapping. Note that even for Czech Republic the misclassification probability is low, assuming that we assign each stock market to the most likely latent class.

The way stock markets are divided into four groups by the model is empirically sound. Broadly speaking, the model distinguishes European developed markets (Class 1) from European emerging markets, highlighting that emerging market returns have different statistical properties than developed market returns. This is sustained by the so-called fundamentals as emerging markets are well-known to offer higher returns because of the faster pace of economic growth, but returns are also more volatile because of the fragility of the political and social structure.

As can be seen, Class 1 is composed mainly by EU countries: Austria, Belgium, Denmark, France, Germany, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, the UK and a non EU country, Switzerland. Class 2 includes Czech Republic, Greece, Hungary, Poland and Sweden. While this second group is mainly composed of Eastern European countries, it also includes two countries whose stock markets have been severely affected by economic events: Greece and Sweden. Class 3 has two neighbor countries: Finland and Russia. Turkey is alone in Class 4. Note that the countries outside Class 1 are the ones with more pronounced booms and slumps in volatility. The next paragraphs give more explanations about of the reasons behind the clustering.

[Figure 6 about here.]

Figures 6, 7, 8 and 9 show the regime-switching dynamics for countries belonging to the same latent class. They depict the posterior probability of being in a bear regime at period t (the shaded area shows probabilities of being in bear regime of above 0.5, i.e., a higher probability of being in a bear regime than in a bull regime). The figures show that the dynamic patterns of the return indices are remarkably similar for countries within each class and substantially different between the four classes.

Figure 6 shows that Class 1 stock markets showed little propensity to switch to the bear regime. We see a long period of bull markets and two main periods of bear markets: The first ranging from the end of 1999 until 2003, and the second representing the subprime crises. Episodes of bear markets were not very frequent and did not last long.

[Figure 7 about here.]

Countries in Class 2 had higher propensity to switch to bear regimes (Figure 7). Differently from Class 1, these stock markets seem to have been more affected by the 2000 crisis. But the subprime crisis affected these countries in a similar way as the countries from Class 1. Greece and Hungary, two countries close to bankruptcy in the aftermath of subprime crisis, spent a long period in bear markets in the beginning of 2010. The bursting of the dot-com bubble had serious economic implications for Sweden. Ericsson, one of the main listed Swedish securities, was the largest producer of mobile telecommunications equipment and suffered heavy losses after the telecommunication crash in the earlies 2000's. It shed thousands of jobs, affecting many other consulting companies and start-ups in the IT sector.

[Figure 8 about here.]

Stock markets in Class 3 had long periods of bear markets (Figure 8). The stocks markets of the two neighbor countries Russia and Finland, although economically different, showed similar dynamics. However, between 2003 and 2007, Finland had a period of bull market, while Russia tended to switch regularly between bull and bear markets. Finally, the sole Cluster 4 country Turkey (Figure 9) tended to switch very quickly to bear regimes; that is, it did not experience long periods of bull markets.

[Figure 9 about here.]

Table 6 provides details about the estimated duration of the regimes for each country; that is, for each of the 21 markets, it reports the mean, first quartile, median, third quartile, and inter-quartile range of the estimated number of days that the market concerned stayed (most likely regime computed from the estimated posterior probabilities) in a given regime before switching into the other regime. From the fact that the means are much larger than the medians, it can be concluded that both for bull and bear markets the duration distributions are asymmetric.

The average duration of bear regimes ranges from 18 in Czech Republic to 52.2 in Russia. Contrary to the bear regimes, large country differences are encountered in bull market durations. In Turkey the average duration is 7.9 and for Switzerland is 236 days.

Contrary to some myths, bear markets can last long for developed markets as can be seen for Switzerland, the Netherlands, Sweden, and Finland, countries with means over 30. But note that those countries have long periods of bull markets as well. The results are consistent with the encountered sojourn times for Class 1.

Looking at the medians, countries of Class 1 have long periods of bull markets while the others have shorter periods of bull markets because they have higher probabilities of switching to bear markets.

[Table 6 about here.]

The last important question to address is the analysis of market synchronization. Measurement of synchronization of stock markets using cross-correlations of returns is rather popular. Crisis periods may yield very large outliers in returns that introduce so much noise and exaggerate similarities between markets. To measure synchronization and co-movement in the 21 stock markets, we compute the association between markets using the posterior probability of being in a bear regime. In other words, synchronization is measured by the similarity of markets with respect to the likelihood of being in the bear regime.

Let $\hat{\alpha}_{it}$ be the estimated probability that market i in period t will be in a bear regime. To obtain a number in the full range of real numbers, this probability is expressed using the logit transformation:

$$\text{logit}_{it} = \log\left(\frac{\hat{\alpha}_{it}}{1-\hat{\alpha}_{it}}\right). \tag{16}$$

Synchronization is quantified as the product-moment correlation between two markets' posterior bear regime logits. This transformation is needed as the probabilities are between 0 and 1, and the Pearson correlation operates on the real line. Our logit-based measure does not suffer from distortion caused by outliers because it filters out extreme observations of returns. The measure gets close to 1 if markets share the same regime state.

Table 7 shows the correlation between stock market regimes using this measure. Countries belonging to Class 1 show a high level of mutual synchronization, but they are less synchronized with countries in the other latent classes. Turkey has low correlations with all other markets: these range from 0.18 to 0.29. It is worth to note that even if countries are not in the same class, they can be synchronized. For instance, Russia's stock market is largely synchronized with Czech Republic, Poland and Hungary, more than with Finland which is in the same class as Russia. The same happens for

Finland and Sweden, which are highly correlated despite not being in the same group. Greece has high correlation with all EU countries, despite of being in another group.

[Table 7 about here.]

5. Conclusion

The identification of groups or classes of financial assets is key in the field of money management. One of the first decisions in the portfolio construction process is the asset allocation decision, which consists of dividing investments among different classes of assets in order to optimize the risk/reward trade-off. Our data mining methodology gives a strong contribution in that as it identifies groups of financial assets.

This paper introduced a model-based clustering technique as a data analytic tool for mining financial time series. The proposed model takes into account both time-constant unobserved heterogeneity between and hidden regimes within time series. Moreover, the flexible modeling of observed responses using a mixture of normal distributions makes it straightforward to capture almost any departure from the normality. For parameter estimation using maximum likelihood, a generalization of the Baum-Welch algorithm for the HMM to the mixture HMM was used.

In the analysis of a sample of 21 stock markets providing observations for a period of 3193 days the best fitting model was the one with four groups. They clearly distinguished very different stock market dynamics, that is, in the switching patterns between the regimes referred to as bear and bull market.

The results obtained with our approach seem to have a one-to-one correspondence with the real market behavior. First, they are consistent with the asymmetry of volatility: The bear state presents higher volatility than the bull state. Second, periods of market crises are clearly distinguished from periods of market stability. Third, less developed markets are assigned with a higher probability to the latent states associated with higher volatility. Fourth, the transition probability matrices are consistent with the persistency of volatility. Periods of large (low) volatility are followed by large (low) volatility. Overall, this data mining approach is coherent with many stylized facts in finance.

Moreover, specifying the correct distribution for returns has important implications. The fact that risk models do not account for non-zero higher moments might cause bias in hedging strategies and concomitantly great losses for financial institutions. For instance, it is widely acknowledged that financial institutions have gradually been adopting market risk models that rely more on historic or back simulations instead of symmetric distributions. Portfolio decisions also need to incorporate information about those higher moments. Bekaert et al. (1998) analyzed the economic impact of taking into account skewness and kurtosis on asset allocation. Their results based on simulations show that investment weights are increased toward the asset with positive skewness (everything constant) and with higher kurtosis (holding skewness positive and constant). The semiparametric nature of the proposed MHMM allows a flexible specification of the distribution of the observed returns that goes beyond the Gaussian distribution.

Results of the application show that our data mining methodology performs well in capturing the different regime dynamics of stock markets. It clearly distinguishes four groups of countries with different observed patterns. It also goes beyond the traditional approaches of categorizing, for instance, countries in developed and emerging market based on GDP per capita (followed by the World Bank) or credit risk ratings as it is directly based on daily market returns. As we have seen, for some markets, because of specific idiosyncrasies, this data mining methodology shows that they have distinct dynamics. Standard approaches fail to incorporate this phenomenon in asset allocation decisions.

Acknowledgement

The authors would like to thank the Fundação para a Ciência e Tecnologia (Portugal) for its financial support (PTDC/EGE-GES/103223/2008).

References

Akaike, H., 1974. A new look at statistical-model identification. IEEE Transactions on Automatic Control AC19 (6), 716–723.

Ang, A., Bekaert, G., 2002. International asset allocation with regime shifts. Review of Financial Studies 15 (4), 1137–1187.

Basalto, N., Bellotti, R., De Carlo, F., Facchi, P., Pantaleo, E., Pascazio, S., 2007. Hausdorff clustering of financial time series. Physica A-Statistical Mechanics and Its Applications 379 (2), 635–644.

Baum, L. E., Petrie, T., Soules, G., Weiss, N., 1970. A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains. Annals of Mathematical Statistics 41 (1), 164–171.

Bekaert, G., Erb, C. B., Harvey, C. R., Viskanta, T. E., 1998. Distributional characteristics of emerging market returns and asset allocation. Journal of Portfolio Management 24 (2), 102–116.

Clogg, C., 1995. Latent class models. In: Arminger, G., Clogg, C., Sobel, M. (Eds.), Handbook of Statistical Modeling for the Social and Behavioral Sciences. Plenum, New York, pp. 311–359.

Dempster, A. P., Laird, N. M., Rubin, D. B., 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society Series B-Methodological 39 (1), 1–38.

Dias, J., Vermunt, J., 2007. Latent class modeling of website users' search patterns: Implications for online market segmentation. Journal of Retailing and Consumer Services 14 (6), 359–368.

Dias, J. G., Wedel, M., 2004. An empirical comparison of EM, SEM and MCMC performance for problematic Gaussian mixture likelihoods. Statistics and Computing 14 (4), 323–332.

Fama, E. F., 1965. The behavior of stock-market prices. Journal of Business 38 (1), 34–105.

Hamilton, J. D., 1989. A new approach to the economic-analysis of nonstationary time-series and the business-cycle. Econometrica 57 (2), 357–384.

Harvey, C. R., 1995. Predictable risk and returns in emerging markets. Review of Financial Studies 8 (3), 773–816.

Heckman, J. J., 2001. Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture. Journal of Political Economy 109 (4), 673–748.

Lystig, T. C., Hughes, J. P., 2002. Exact computation of the observed information matrix for hidden Markov models. Journal of Computational and Graphical Statistics 11 (3), 678–689.

Mandelbrot, B., 1963. The variation of certain speculative prices. Journal of Business 36 (4), 394–419.

McLachlan, G., Peel, D., 2000. Finite Mixture Models. John Wiley & Sons, New York.

Pagan, A. R., Sossounov, K. A., 2003. A simple framework for analysing bull and bear markets. Journal of Applied Econometrics 18 (1), 23–46.

Peiró, A., 1999. Skewness in financial returns. Journal of Banking & Finance 23 (6), 847–862.

Praetz, P. D., 1972. Distribution of share price changes. Journal of Business 45 (1), 49–55.

Schmittmann, V. D., Visser, I., Raijmakers, M. E. J., 2006. Multiple learning modes in the development of performance on a rule-based category-learning task. Neuropsychologia 44 (11), 2079–2091.

Schwarz, G., 1978. Estimating the dimension of a model. Annals of Statistics 6 (2), 461–464.

Simkowitz, M. A., Beedles, W. L., 1980. Asymmetric stable distributed security returns. Journal of the American Statistical Association 75 (370), 306–312.

Singleton, J. C., Wingender, J., 1986. Skewness persistence in common-stock returns. Journal of Financial and Quantitative Analysis 21 (3), 335–341.

van de Pol, F., Langeheine, R., 1990. Mixed Markov latent class models. Sociological Methodology 20, 213–247.

Vermunt, J. K., 2003. Multilevel latent class models. Sociological Methodology 33, 213–239.

Vermunt, J. K., 2007. A hierarchical mixture model for clustering three-way data sets. Computational Statistics & Data Analysis 51 (11), 5368–5376.

Vermunt, J. K., Langeheine, R., Bockenholt, U., 1999. Discrete-time discrete-state latent Markov models with time-constant and time-varying covariates. Journal of Educational and Behavioral Statistics 24 (2), 179–207.

Wilfling, B., 2009. Volatility regime-switching in European exchange rates prior to monetary unification. Journal of International Money and Finance 28 (2), 240–270.

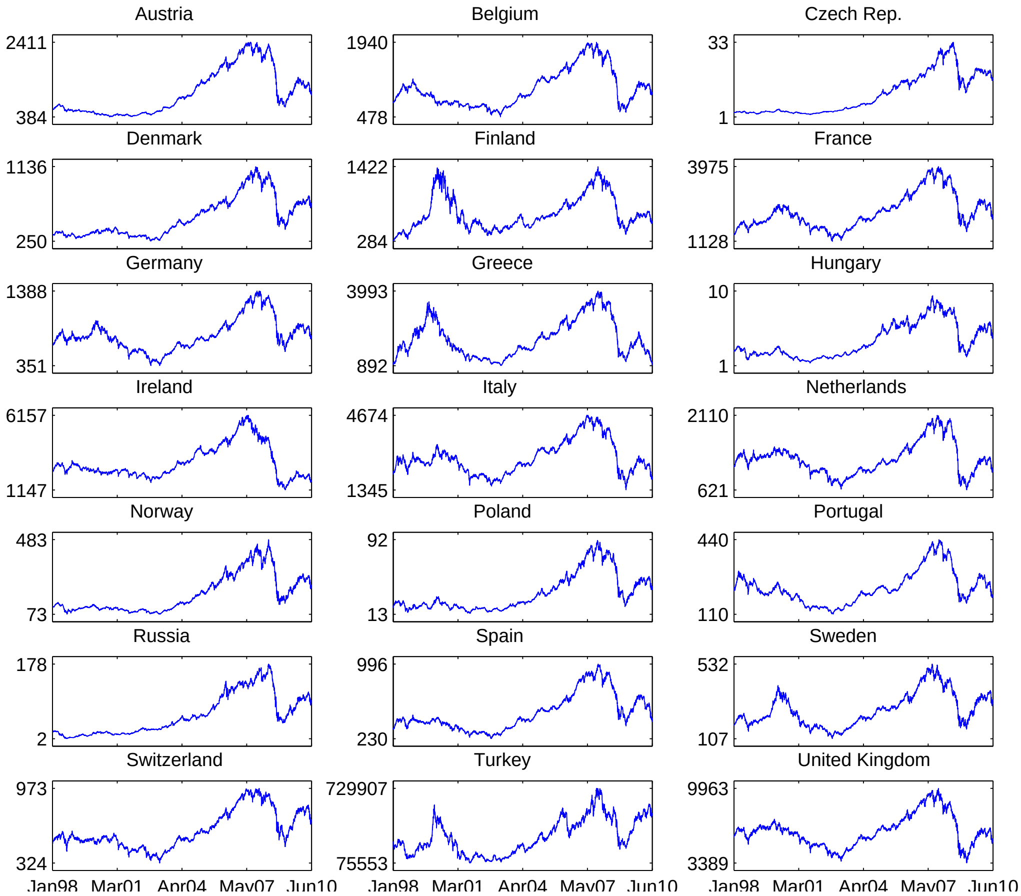

Figure 1: Time series of price indexes for 21 European stock markets

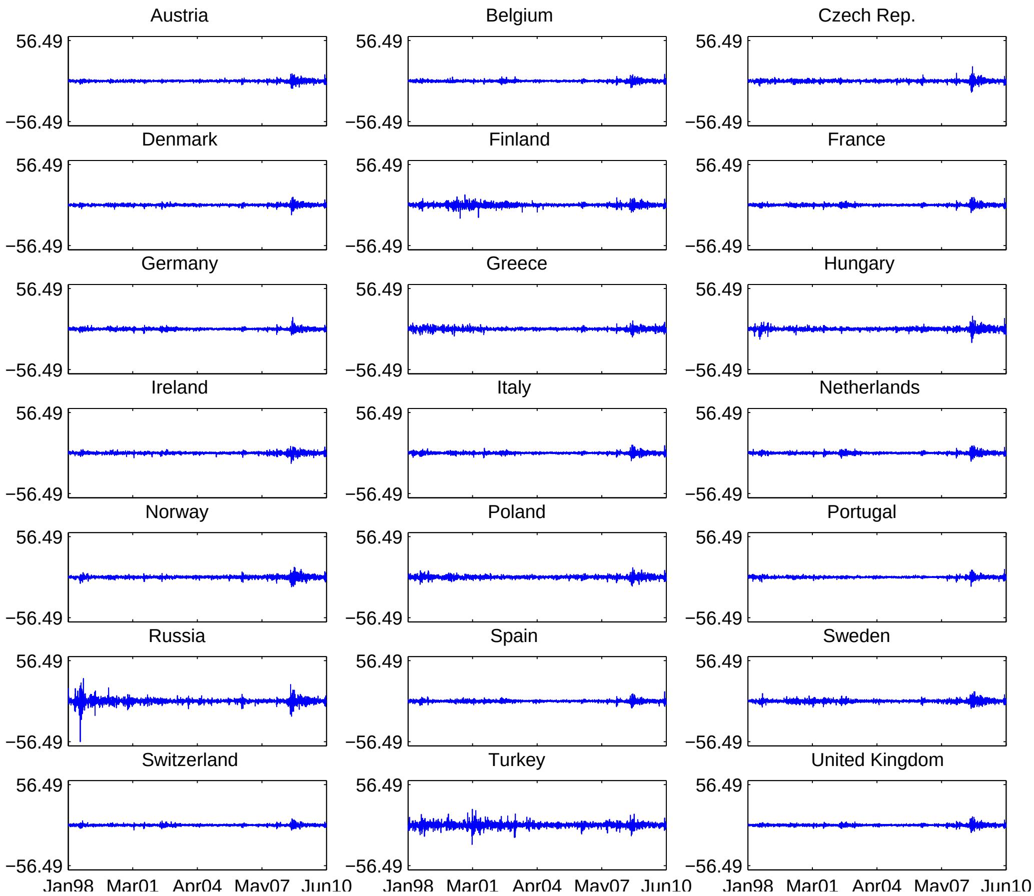

Figure 2: Time series of returns for 21 European stock markets

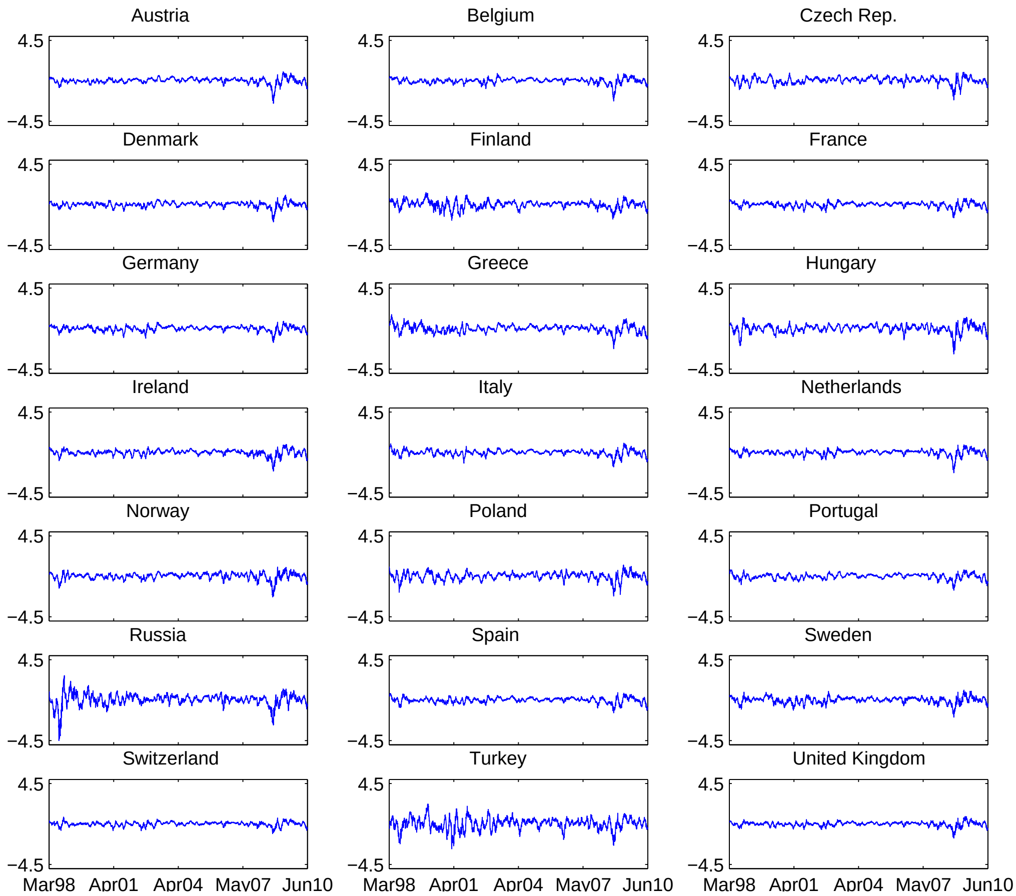

Figure 3: Time series of rolling means for 21 European stock markets (30-day window)

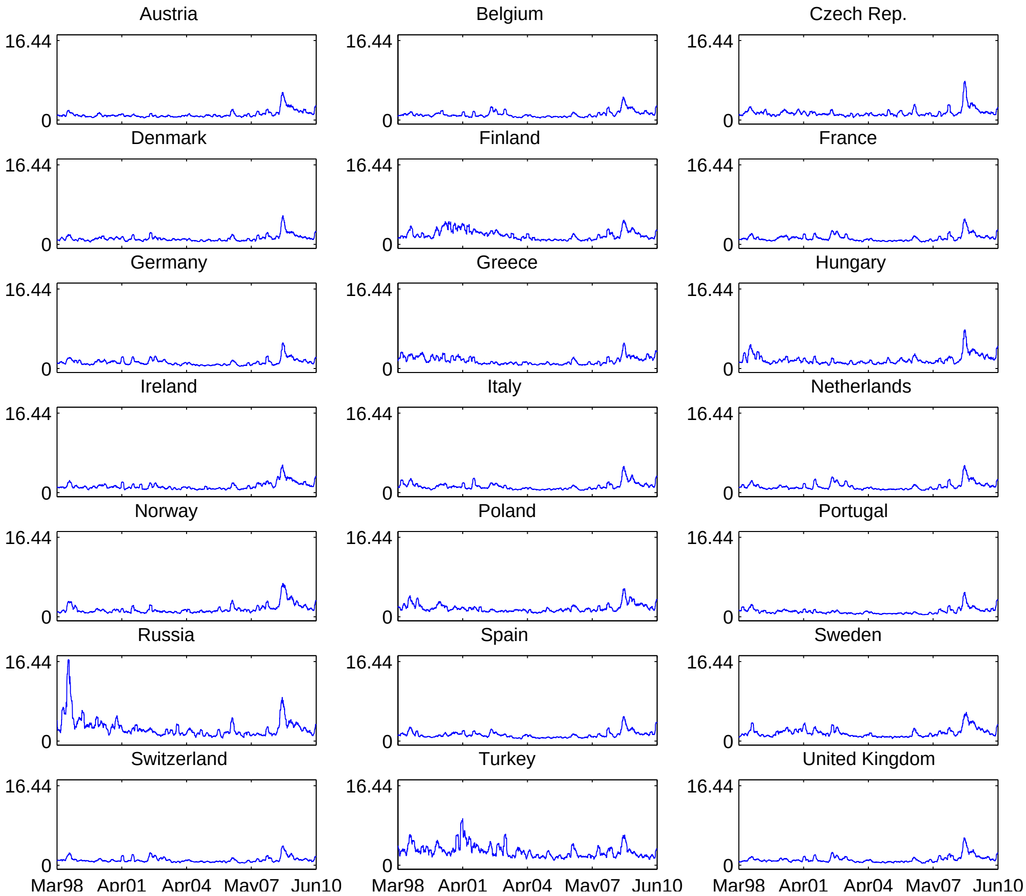

Figure 4: Time series of rolling standard deviations for 21 European stock markets (30-day window)

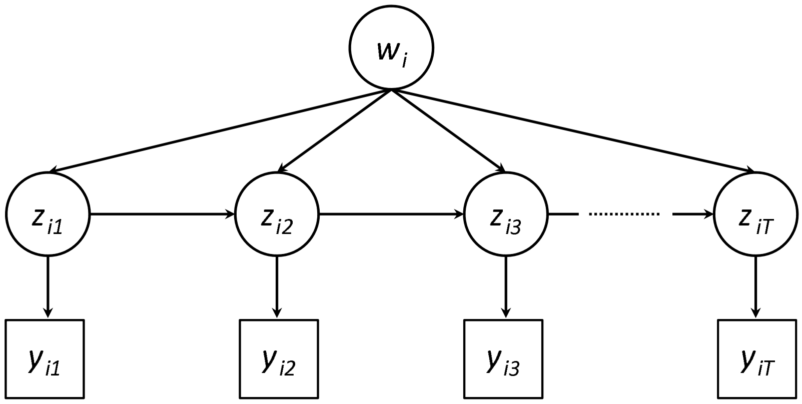

Figure 5: The graph of the proposed model

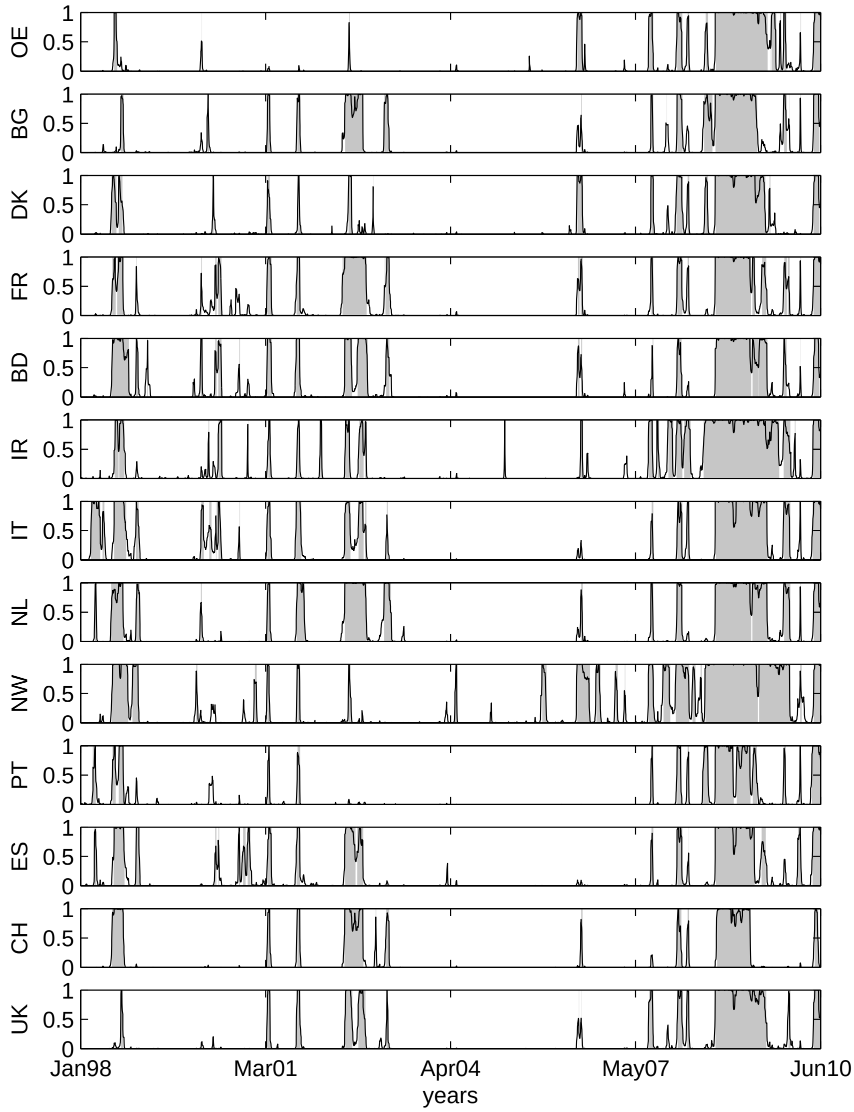

Figure 6: Estimated posterior latent state probabilities and modal state in latent class 1

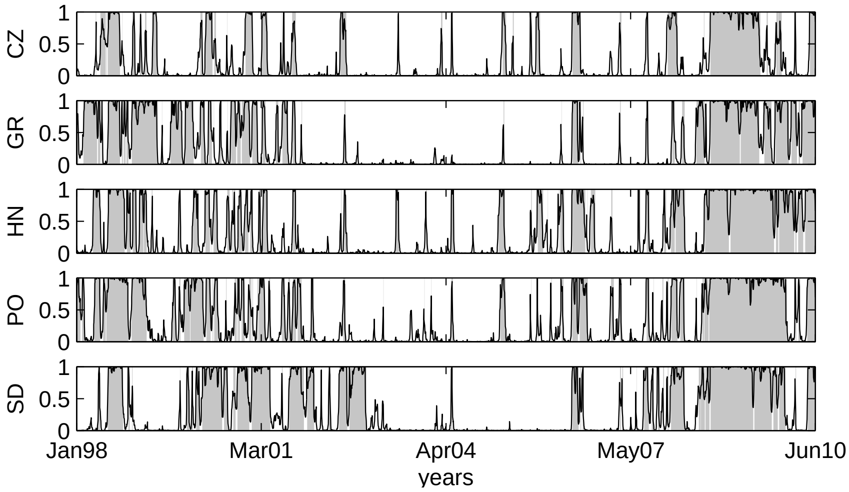

Figure 7: Estimated posterior latent state probabilities and modal state in latent class 2

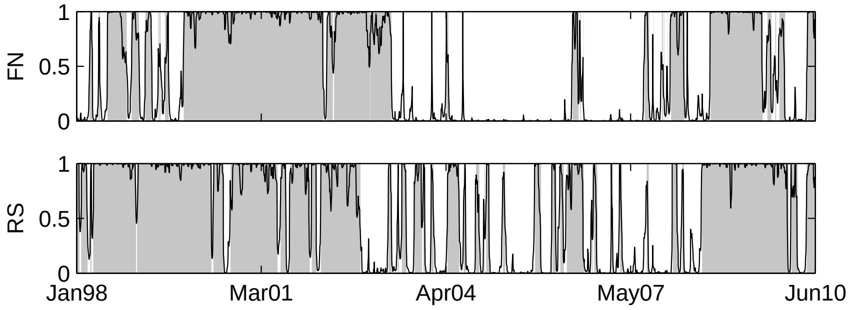

Figure 8: Estimated posterior latent state probabilities and modal state in latent class 3

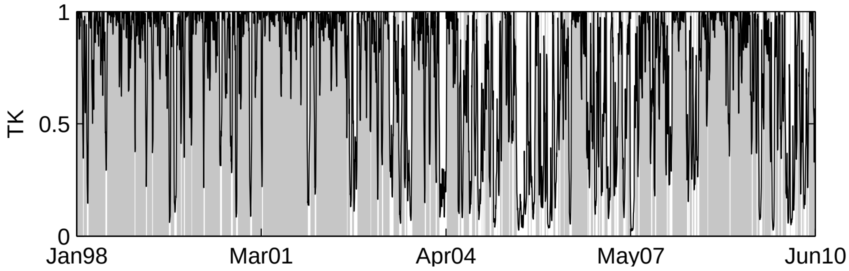

Figure 9: Estimated posterior latent state probabilities and modal state in latent class 4

Table 1: Summary statistics

Stock market	Mean	Median	Standard deviation	Skewness	Kurtosis	Jarque-Bera test	
						statistics	p-value
Austria (OE)	0.016	0.055	1.411	-0.397	11.537	9727.96	0.000
Belgium (BG)	0.006	0.081	1.377	-0.168	8.921	4652.36	0.000
Czech Rep. (CZ)	0.056	0.000	1.762	-0.104	16.141	22865.83	0.000
Denmark (DK)	0.022	0.065	1.427	-0.383	11.242	9066.71	0.000
Finland (FN)	0.022	0.051	2.193	-0.318	8.658	4288.99	0.000
France (FR)	0.012	0.064	1.492	-0.024	9.203	5090.53	0.000
Germany (BD)	0.005	0.077	1.459	0.239	11.437	9450.20	0.000
Greece (GR)	-0.002	0.061	1.837	-0.132	6.963	2085.80	0.000
Hungary (HN)	0.013	0.000	2.106	-0.196	12.208	11241.19	0.000
Ireland (IR)	-0.009	0.059	1.579	-0.601	10.228	7105.78	0.000
Italy (IT)	-0.004	0.052	1.517	-0.054	9.638	5831.67	0.000
Netherlands (NL)	-0.003	0.060	1.518	-0.233	10.072	6647.45	0.000
Norway (NW)	0.021	0.088	1.900	-0.480	10.033	6666.59	0.000
Poland (PO)	0.022	0.065	1.981	-0.131	6.862	1980.57	0.000
Portugal (PT)	-0.005	0.039	1.316	-0.152	12.649	12335.84	0.000
Russia (RS)	0.052	0.104	3.408	-1.262	36.977	153728.66	0.000
Spain (ES)	0.006	0.066	1.489	0.004	9.894	6288.85	0.000
Sweden (SD)	0.016	0.056	1.937	0.073	7.293	2439.62	0.000
Switzerland (CH)	0.011	0.034	1.225	-0.023	7.773	3013.11	0.000
Turkey (TK)	0.019	0.013	3.222	-0.129	9.177	5057.35	0.000
United Kingdom (UK)	-0.002	0.043	1.408	-0.125	11.846	10365.40	0.000

Table 2: Model selection

Number of latent classes (S)	LL	No. of parameters	BIC
1	-121799.23	7	243619.76
2	-121710.66	11	243454.81
3	-121667.34	15	243380.35
4	-121657.84	19	243373.53
5	-121653.33	23	243376.69

Table 3: Estimated marginal probabilities of the latent states and within Gaussian parameters

$P(Z)$		Return (mean: $\hat{\mu}_k$)		Risk (variance: $\hat{\sigma}^2_k$)	
Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2
0.270	0.730	-0.175	0.079	10.264	1.129
(0.036)	(0.036)	(0.025)	(0.005)	(0.146)	(0.011)

Table 4: Estimated transition probabilities

	Latent class 1		Latent class 2		Latent class 3		Latent class 4	
	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2
$P(Z\|W)$	0.162	0.838	0.324	0.676	0.534	0.466	0.726	0.274
	(0.012)	(0.012)	(0.020)	(0.020)	(0.039)	(0.039)	(0.025)	(0.025)
Regime 1	0.960	0.040	0.946	0.054	0.975	0.026	0.930	0.070
	(0.003)	(0.003)	(0.005)	(0.005)	(0.004)	(0.004)	(0.012)	(0.012)
Regime 2	0.008	0.992	0.026	0.974	0.029	0.971	0.187	0.813
	(0.001)	(0.001)	(0.003)	(0.003)	(0.004)	(0.004)	(0.028)	(0.028)
Sojourn time	24.752	126.582	18.484	38.462	39.216	34.247	14.205	5.350

Table 5: Estimated prior probabilities, posterior probabilities, and modal class

	Latent classes				
	1	2	3	4	Modal
Prior probabilities	0.61	0.23	0.10	0.06	
Posterior probabilities					
Austria (OE)	1.00	0.00	0.00	0.00	1
Belgium (BG)	1.00	0.00	0.00	0.00	1
Czech Rep. (CZ)	0.10	0.90	0.00	0.00	2
Denmark (DK)	1.00	0.00	0.00	0.00	1
Finland (FN)	0.00	0.00	1.00	0.00	3
France (FR)	1.00	0.00	0.00	0.00	1
Germany (BD)	1.00	0.00	0.00	0.00	1
Greece (GR)	0.00	1.00	0.00	0.00	2
Hungary (HN)	0.00	1.00	0.00	0.00	2
Ireland (IR)	1.00	0.00	0.00	0.00	1
Italy (IT)	1.00	0.00	0.00	0.00	1
Netherlands (NL)	1.00	0.00	0.00	0.00	1
Norway (NW)	0.99	0.01	0.00	0.00	1
Poland (PO)	0.00	1.00	0.00	0.00	2
Portugal (PT)	1.00	0.00	0.00	0.00	1
Russia (RS)	0.00	0.00	1.00	0.00	3
Spain (ES)	1.00	0.00	0.00	0.00	1
Sweden (SD)	0.00	0.99	0.01	0.00	2
Switzerland (CH)	1.00	0.00	0.00	0.00	1
Turkey (TK)	0.00	0.00	0.00	1.00	4
United Kingdom (UK)	1.00	0.00	0.00	0.00	3

Table 6: Estimated regime durations.

Countries	Bear regime					Bull regime				
	Mean	Q1	Median	Q3	IQR	Mean	Q1	Median	Q3	IQR
Austria (OE)	25.7	3.0	10.0	22.0	19.0	187.2	17.0	61.0	287.0	270.0
Belgium (BG)	25.2	2.5	12.0	24.0	21.5	153.6	45.0	92.0	193.0	148.0
Czech Rep. (CZ)	18.0	3.0	7.0	23.0	20.0	66.1	18.0	40.5	83.0	65.0
Denmark (DK)	26.3	7.0	12.0	24.0	17.0	174.9	27.5	109.5	220.5	193.0
Finland (FN)	48.4	2.0	11.0	32.0	30.0	54.6	14.0	21.0	60.0	46.0
France (FR)	22.1	6.0	12.0	18.0	12.0	116.7	7.0	57.0	137.0	130.0
Germany (BD)	23.7	5.0	14.0	28.0	23.0	116.2	24.0	72.0	136.0	112.0
Greece (GR)	22.3	3.0	10.5	31.0	28.0	48.1	8.0	18.0	32.0	24.0
Hungary (HN)	21.1	5.0	12.0	20.5	15.5	44.1	8.0	20.0	66.0	58.0
Ireland (IR)	28.6	7.0	15.5	22.0	15.0	116.5	19.0	66.0	115.0	96.0
Italy (IT)	25.5	7.0	11.5	23.0	16.0	119.6	23.0	47.5	110.0	87.0
Netherlands (NL)	35.2	7.5	24.5	43.5	36.0	164.4	58.5	89.5	218.5	160.0
Norway (NW)	35.3	8.0	15.5	31.0	23.0	109.9	24.0	49.0	136.0	112.0
Poland (PO)	19.7	3.0	8.5	18.0	15.0	38.0	8.0	19.0	44.0	36.0
Portugal (PT)	21.6	7.0	14.0	24.0	17.0	191.3	27.0	61.0	118.0	91.0
Russia (RS)	52.2	9.0	17.0	42.5	33.5	35.1	12.5	20.5	46.5	34.0
Spain (ES)	24.8	8.0	14.0	26.0	18.0	143.3	28.0	75.0	127.5	99.5
Sweden (SD)	25.6	4.0	10.0	35.0	31.0	48.7	8.0	14.0	37.0	29.0
Switzerland (CH)	32.8	6.0	13.0	51.0	45.0	236.0	48.0	128.5	342.5	294.5
Turkey (TK)	22.8	5.0	12.5	26.0	21.0	7.9	4.0	6.0	11.0	7.0
United Kingdom (UK)	29.3	8.0	14.0	30.0	22.0	198.8	91.0	109.5	194.0	103.0

Table 7: Synchronization of stock markets (posterior probabilities correlations)

Countries	OE	BG	CZ	DK	FN	FR	BD	GR	HN	IR	IT	NL	NW	PO	PT	RS	ES	SD	CH	TK
Austria (OE)																				
Belgium (BG)	0.74																			
Czech Rep. (CZ)	0.71	0.60																		
Denmark (DK)	0.79	0.79	0.69																	
Finland (FN)	0.36	0.44	0.41	0.44																
France (FR)	0.69	0.85	0.60	0.76	0.55															
Germany (BD)	0.68	0.78	0.62	0.75	0.54	0.86														
Greece (GR)	0.48	0.41	0.47	0.45	0.37	0.41	0.44													
Hungary (HN)	0.63	0.51	0.62	0.59	0.36	0.52	0.55	0.45												
Ireland (IR)	0.74	0.72	0.62	0.71	0.40	0.68	0.67	0.41	0.58											
Italy (IT)	0.71	0.75	0.62	0.75	0.52	0.82	0.80	0.52	0.60	0.68										
Netherlands (NL)	0.67	0.84	0.57	0.75	0.48	0.90	0.84	0.38	0.52	0.65	0.80									
Norway (NW)	0.75	0.67	0.67	0.75	0.36	0.63	0.65	0.41	0.64	0.69	0.67	0.64								
Poland (PO)	0.60	0.49	0.60	0.53	0.38	0.49	0.53	0.50	0.62	0.54	0.58	0.47	0.60							
Portugal (PT)	0.74	0.70	0.64	0.72	0.42	0.68	0.66	0.54	0.60	0.69	0.77	0.64	0.62	0.56						
Russia (RS)	0.34	0.29	0.42	0.33	0.34	0.30	0.35	0.38	0.42	0.29	0.38	0.29	0.37	0.45	0.33					
Spain (ES)	0.71	0.75	0.64	0.75	0.52	0.83	0.80	0.48	0.62	0.68	0.83	0.78	0.66	0.56	0.78	0.36				
Sweden (SD)	0.63	0.68	0.56	0.68	0.64	0.75	0.74	0.40	0.53	0.65	0.71	0.68	0.65	0.52	0.59	0.35	0.71			
Switzerland (CH)	0.63	0.77	0.60	0.75	0.46	0.82	0.79	0.39	0.50	0.62	0.78	0.79	0.61	0.44	0.69	0.31	0.79	0.64		
Turkey (TK)	0.20	0.20	0.28	0.25	0.29	0.25	0.26	0.25	0.23	0.18	0.27	0.22	0.20	0.28	0.23	0.28	0.27	0.27	0.26	
United Kingdom (UK)	0.75	0.84	0.63	0.82	0.46	0.86	0.79	0.41	0.55	0.73	0.78	0.84	0.70	0.52	0.68	0.29	0.77	0.71	0.79	0.23